

Björn Goerke

Essays zur Förderung und Bewertung von Innovationen

Essays zur Förderung und Bewertung von Innovationen

Inaugural-Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Wirtschafts- und Sozialwissenschaften der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von
Dipl.-Kfm. Björn Goerke
aus Kiel

Hamburg, 2016

Gedruckt mit Genehmigung der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

Dekan: Prof. Dr. Achim Walter

Erstbegutachtung: Prof. Dr. Dr. h.c. Sönke Albers

Zweitbegutachtung: Prof. Dr. Stefan Hoffmann

Tag der Abgabe der Arbeit: 23.02.2016

Tag der mündlichen Prüfung: 30.05.2016

Danksagung

Die vorliegende Arbeit ist im Rahmen und teilweise mit Förderung des DFG-Graduiertenkollegs „Betriebswirtschaftliche Aspekte lose gekoppelter Systeme und Electronic Business“ entstanden, das sich in einem Schwerpunkt mit dem Management von Innovationen und der damit verbundenen Arbeit in Netzwerken beschäftigt hat. Die vorliegenden Studien spiegeln jeweils Facetten dieses Forschungsfeldes wider, das für die durch Innovationen getriebene Wirtschaft von großer Relevanz ist.

Es hat etwas mehr als zehn Jahre gedauert, diese Studien durchzuführen und sie zu Papier zu bringen. Das ist eine sehr lange Zeit. Über die Jahre habe ich von vielen einzelnen Personen auf ganz unterschiedliche Art und Weise Unterstützung erfahren, für die ich sehr dankbar bin. Zunächst möchte ich hier meinen Doktorvater Professor Dr. Dr. h.c. Sönke Albers nennen, der mit mir immer wieder konstruktiv die einzelnen Projekte diskutiert und vorangetrieben hat. Ihm gebührt mein tiefer Dank. Insbesondere auch seine Bereitschaft, sich immer wieder über die Jahre tief in die Fragestellungen hineinzuarbeiten, hat es erst ermöglicht, diese Arbeit fertigzustellen.

Ebenso möchte ich mich bei Professor Dr. Hoffmann bedanken, der das Zweitgutachten erstellte. Außerdem gilt mein Dank Professor Dr. Schultz für die Übernahme des Prüfungsvorsitzes.

Ohne meine Diplomarbeitsbetreuer, Professor Dr. Michel Clement und Professor Dr. Marc Fischer, die heute beide ihre eigenen Lehrstühle bekleiden, hätte ich vielleicht nie den Weg in die Wissenschaft gefunden. Da ich sehr glücklich darüber bin, dass ich einige Jahre der Forschung widmen konnte, möchte ich Euch beiden für die Inspiration danken und dafür, dass Ihr mir den Weg geebnet habt. Unvergessen bleibt unser Retreat in Oppenheim (oder wo auch immer wir genau waren), der in Erinnerung gleich in mehrfacher Hinsicht die Tränen in die Augen treibt. Auch wenn aus dem Retreat nie ein Artikel entstanden ist – möchte ich Euch dafür danken, dass ich bei anderen Projekten Euer Ko-Autor sein durfte.

Ein Highlight war auch die Teilnahme an der Marketing Science 2007 in Singapur. Unvergessen bleiben Vorträge, die beinahe zu Tumulten im Auditorium führten, und der Ausflug nach Sentosa.

Besonders bedanken möchte ich mich bei meinem ehemaligen Arbeitgeber, der Wirtschaftsförderung und Technologietransfer Schleswig-Holstein (WTSH). Ohne die Hilfe der Kollegen hätte ich die Studie über die Wirkungsmessung in der Wirtschaftsförderung nicht durchführen können. An dieser Stelle möchte ich gern einzelnen Personen auch noch meinen besonderen Dank direkt aussprechen. Zum einen gilt der Dank Dr. Bernd Bösche, der mir die Freiräume

gewährte, die ich brauchte, um die Arbeit durchführen zu können. Und zum anderen gilt der Dank Hajo Klar, der für mich natürlich immer Chef, aber auch ein toller Mentor war und mich unterstützt hat. Zwei weitere Kollegen haben mit mir Kaltakquise für die Datenerhebung bei den Mittelständlern in Schleswig-Holstein gemacht: Kolja C. Scepanik und Niels Leininger. Ich denke, wir haben dabei alle etwas fürs Leben gelernt. Vielen Dank für die Unterstützung!

Während der Promotion haben Dr. Dennis Proppe, ebenfalls Doktorand von Prof. Dr. Dr. h.c. Albers, und ich uns mit dem Wissen aus der Forschung selbständig gemacht. Ziemlich schnell haben wir festgestellt, dass es sich um zwei unterschiedliche Welten handelt, auch wenn die Wissensbasis die gleiche ist. Dennis, vielen Dank für die zahlreichen Diskussionen über methodische Fragestellungen, die jede für sich ein Baustein dieser Arbeit waren.

Und schließlich möchte ich mich bei meiner Familie bedanken. Ohne die Unterstützung meiner Eltern hätte ich mich vielleicht nie für das Studium an der Uni entschieden. Und nur durch die großzügige Unterstützung war das Studium überhaupt in dieser Form möglich. Und während der Promotion ein offenes Ohr für die Sorgen zu haben, wenn nicht alles ganz geradlinig verläuft, hilft sehr. Hierfür danke ich vor allem Panja (korrekterweise Dr. Panja Goerke) von ganzem Herzen für die Hilfe! Noch viel mehr danke ich Dir aber für das unendliche Verständnis, das Du immer wieder aufgebracht hast, wenn ich Abend um Abend und dann wieder an den Wochenenden meine Zeit mit der Forschung verbracht habe!! Mein Dank ist nicht in Worte zu fassen!

Björn Goerke

Inhaltsverzeichnis

Synopsis	IX
A. Event-Studies	1
B. Ausreißerwerte.....	23
C. Long-term effects of subsidies on firm growth: introducing the concept of outcome additionality	44
D. Neuprodukteinführungen in der Filmindustrie: Wie reagieren Kapitalmarktinvestoren auf den Umsatzerfolg neuer Kinofilme?	74
E. The Effect of Unexpected Chart Positions on the Firm Value of Music Labels. An Event Study of Album Success.....	75
Eidesstattliche Erklärung.....	XVIII

Synopsis

1 Einleitung

Die vorliegende Arbeit ist sowohl extern, im Stipendium der Deutschen Forschungsgemeinschaft (DFG) als auch während der Zeit als Mitarbeiter des Lehrstuhls für Innovation, Neue Medien und Marketing von Prof. Dr. Dr. h.c. Sönke Albers entstanden. Den Rahmen gab dabei zunächst das Graduiertenkolleg Betriebswirtschaftliche Aspekte lose gekoppelter Systeme und Electronic Business“. Zu dem Forschungsschwerpunkt des Graduiertenkollegs sind die folgenden fünf Aufsätze verfasst worden.

Übersicht der Fachartikel

Methodenartikel

Event-Studies	Björn Goerke erschieden in: Albers, Sönke, Daniel Klap- per, Udo Konradt, Achim Walter und Joachim Wolf (Hrsg.): Methodik der empiri- schen Forschung, 3. Auflage, Wiesbaden, 467-484.
Ausreißerwerte	Björn Goerke erschieden in: Albers, Sönke, Daniel Klap- per, Udo Konradt, Achim Walter und Joachim Wolf (Hrsg.): Methodik der empiri- schen Forschung (http://www.bwl.uni-kiel.de/grad-kolleg/methodenbuch/)

Empirische Studien

Long-term effects of subsidies on firm growth: introducing the concept of outcome additionality	Björn Goerke und Sönke Albers wird zu Begutachtung eingereicht bei <i>Research Policy</i>
Neuprodukteinführungen in der Filmindustrie: Wie reagieren Kapitalmarktinvestoren auf den Umsatzerfolg neuer Kinofilme?	Michel Clement, Marc Fischer und Björn Goerke erschieden in: <i>Die Betriebswirtschaft</i> , 64. Jg., Heft 4, 418-444.

The Effect of Unexpected Chart Positions on the Firm Value of Music Labels. An Event Study of Album Success.

Nima Mehrafshan, Björn Goerke und Michel Clement

Unter Begutachtung beim

Journal of Media Economics

Dabei finden sowohl grundlegende methodische Aspekte (in zwei Methodenartikeln) als auch empirische Studien (in drei empirischen Studien) Berücksichtigung. In der folgenden kurzen Einführung sollen der theoretische Hintergrund und die thematischen Grundlagen der Beiträge aufgezeigt werden.

2 Förderung und Bewertung von Innovationen

Es gilt als vorteilhaft für die Gesellschaft, wenn alle Innovations-Projekte mit einem positiven Erwartungswert auch tatsächlich durchgeführt werden (Arrow, 1962; Nelson, 1959). Was so einfach klingt, ist in der Realität nicht immer so eindeutig. Gerade wenn es um die Entwicklung neuer Produkte, geht, tragen diese Projekte substanzielle Risiken, die nicht von allen Marktteilnehmern unbedingt gleich bewertet werden bzw. bewertet werden sollten (Merito, Giannangeli & Bonnacorsi, 2010).

Die zukünftige Entwicklung von Unternehmen hängt vielfach davon ab, wie gut sie in der Entwicklung neuer Produkte sind (David, Hall & Toole, 2000). Entsprechend hoch ist das Interesse sowohl aus privatwirtschaftlicher als auch aus öffentlicher Sicht, das wirtschaftliche Potenzial von Innovationen richtig einzuschätzen. Im öffentlichen Bereich versucht der Staat mit Hilfe von Subventionen diejenigen Innovationsprojekte zu fördern, die einen besonders hohen Beitrag für die Wohlfahrt versprechen (David, Hall & Toole, 2000; Wolff & Reinthaler, 2008). Im privatwirtschaftlichen Bereich hat sich mit dem Forschungsfeld der Event Study ein Gebiet entwickelt, in dem mittlerweile auch intensiv der Einfluss des Erfolgs von Innovationen auf den Unternehmenswert beleuchtet wird (Chaney, Devinney & Winer, 1991; Kelm, Narayanan & Pinches, 1995; Sorescu, Shankar & Kushwaha, 2007). Im Rahmen dieser Forschungsarbeiten werden sowohl Investitionen der öffentlichen Hand in die Entwicklung neuer Produkte als auch Investitionen privatwirtschaftlicher Marktteilnehmer in innovations-treibende Unternehmen untersucht.

2.1 Methodenartikel

Die Bewertung von Innovationen stellt sowohl für die Forschung als auch für die Praxis eine Herausforderung dar. Die Fragestellung, wie wertvoll ein neues Produkt für den Unterneh-

mentwert ist, ist zwar sehr relevant, jedoch sind insbesondere unternehmensintern ermittelte Werte nur eingeschränkt vertrauenswürdig. Mit der Event-Study steht eine Methode zur Verfügung, die die Schätzung des Wertes einer Aktion, z.B. der Einführung eines neuen Produktes, anhand der Entwicklung des Aktienkurses ermöglicht. In der Finanzwirtschaft entwickelt (Fama, 1970; Fama, Fisher, Jensen & Roll, 1969), wird die Methode heute auch in der Innovationsforschung eingesetzt und hochrangig publiziert (Agrawal & Kamakura, 1995; Koku, Jagpal & Visnawath, 1997; McWilliams & Siegel, 1997; Tellis & Johnson, 2007). Ziel der Event Study ist es, diejenige Bewegung von Aktienkursen, die durch ein Ereignis ausgelöst wurden, von anderen kursbewegenden Einflüssen zu trennen und zu quantifizieren, um diesem Ereignis dann schließlich einen Wert beimessen zu können. So wird es beispielsweise möglich, den Wert eines neuen Produktes aus Sicht der Investoren zu bestimmen. Der Beitrag zu *Event-Studies*, der in dem von Albers et al. herausgegebenen Buch „Methodik der empirischen Forschung“ erschienen ist, liefert einen Einstieg in die Anwendung in der eigenen Forschung und bietet darüber hinaus einen Überblick über die Erweiterungen, die in den vergangenen Jahren entwickelt wurden. In zwei der drei nachfolgend vorgestellten empirischen Studien findet die Event Study Anwendung.

Der Beitrag *Ausreißerwerte* bezieht sich auf ein grundsätzlich relevantes empirisches Thema. In der empirischen Forschung werden vielfach Daten „aus dem wahren Leben“ verwendet. Die Datensätze enthalten dadurch nicht durchgehend wünschenswerte bzw. erwartete Variablenausprägungen, sondern häufig auch extreme Werte, die Einfluss auf die Ergebnisse nehmen. Diese so genannten Ausreißerwerte können die eigentlichen Ergebnisse überlagern oder verzerren oder aber auch den Weg zu neuen Erkenntnissen ebnen (Hill & Rothaermel, 2003).

Wichtig ist der umsichtige und verantwortungsvolle Umgang mit Ausreißerwerten. Der Methodenbuchartikel wurde mit der Motivation verfasst, genau dafür das grundlegende Handwerkszeug zu liefern. Der Artikel bietet einen Einstieg in die Identifikation und Klassifizierung unterschiedlicher Arten von Ausreißerwerten. Neben Verfahren der Bestimmung, wie sehr die eigentliche Untersuchung durch die einzelnen Ausreißerwerte beeinflusst wird, werden unterschiedliche Wege aufgezeigt, wie man mit diesen Werten umgehen kann. Die häufig verwendete Elimination von Ausreißerwerten ist nur eine von mehreren Möglichkeiten und nur unter bestimmten Bedingungen das Mittel der Wahl. Häufig bieten sich Alternativen wie eine andere Modellierung oder sogar die Erweiterung des Untersuchungshorizonts an.

2.2 Empirische Studien

Im Hinblick auf die öffentliche Förderung von privatwirtschaftlichen Investitionen durch die Vergabe von Subventionen stellt sich die Frage, ob das auch tatsächlich eine sinnvolle Verwendung der Steuern ist (Ortega-Argilés, Vivarelli & Voigt, 2009). Dies gilt natürlich auch für die Förderung der Entwicklung neuer Produkte. Die Argumentation dafür folgt der Linie, dass die Investitionen für die Unternehmen zu viel Risiko bedeuten und sie aufgrund ihrer Risikoaversion diese Innovationsprojekte nicht ausführen, obwohl sie bei Risikoneutralität vorteilhaft wären. Die Gesellschaft als Ganzes führt so viele Projekte aus, dass sie risikoneutral entscheiden kann und muss deswegen ein Interesse haben, dass diese Art von Innovationsprojekten durchgeführt wird. Durch die Vergabe der Subventionen wird das Risiko teilweise auf die Gesellschaft übertragen in der Hoffnung, dass die oben beschriebenen Projekte dann durchgeführt werden und einen positiven Beitrag zur Wohlfahrt leisten (Aerts & Schmidt, 2008).

Die bisherige Forschung hat sich bei der Untersuchung der Wirksamkeit von Subventionen zu großen Teilen auf die Frage konzentriert, ob es zu einem so genannten Crowding-out kommt (Aerts & Schmidt, 2008; Görg & Strobl, 2007, David, Hall & Toole, 2000). Das bedeutet, dass die Unternehmen ihre eigenen Investitionsmittel einfach durch die Subventionen ersetzen. Kommt es nicht zu einem Crowding-out oder nur teilweise, dann spricht man von Input Additionality (Clarysse, Wright & Mustar, 2009). Das bedeutet, dass nach der Vergabe von Subventionen insgesamt mehr Mittel für die Innovationsprojekte zur Verfügung stehen als vorher.

Diese Frage ist zwar wichtig, sie beantwortet jedoch nicht die Frage, ob tatsächlich ein Nutzen im Sinne von Wohlfahrt für die Gesellschaft entstanden ist. Ein Teil der Forschung geht einen Schritt weiter und untersucht die Fragestellung, ob neben mehr Input auch mehr Output durch die Subventionen entsteht (Hujer & Radic, 2005; Branstetter & Sakakibara, 1998; Griliches & Regev, 1995). Gemessen wird diese so genannte Output Additionality beispielsweise mit der Erhöhung der Anzahl neuer Produkte, neuer Patente und ähnlicher Zielgrößen.

Eine große Forschungslücke befindet sich im Bereich der Frage, wie denn tatsächlich die Auswirkungen auf die Wohlfahrt sind. Erste wichtige Artikel befassen sich mit den Auswirkungen eines umfangreichen US-amerikanischen Subventionsprogramms auf den Umsatz und die Anzahl Beschäftigter der geförderten Unternehmen (Lerner, 1999; Wallsten, 2000). Der vorliegende Artikel *Long-term effects of subsidies on firm growth: introducing the concept of outcome additionality* geht hier weiter und beleuchtet neue Aspekte in dem Feld, für das der

Begriff der *Outcome Additionality* von den Autoren zunächst neu eingeführt wird. Insbesondere wird die langfristige Wirkung der Subventionen sowohl auf den Umsatz als auch auf die Beschäftigtenzahl der geförderten Unternehmen analysiert. Hierbei wird auch erarbeitet, inwiefern die Effekte der Subventionen über die Jahre gegebenenfalls auch wieder nachlassen.

Die Studie zeigt zum einen, dass es möglich und sinnvoll ist, die Frage nach der langfristigen Wirksamkeit der Subventionen zu untersuchen. Und zum anderen wird deutlich, dass die Subventionen nicht zwangsläufig eine positive Wirkung auf die geförderten Unternehmen entfalten. Vielmehr ist es so, dass es je nach Förderinstrument auch zu möglichen Schäden bei den Unternehmen führen kann, wenn diese beispielsweise mit einem Innovationsprojekt überfordert sein könnten, es aber dennoch durchführen, weil sie aufgrund der Subventionen dazu ermutigt wurden.

Einen anderen Blickwinkel nimmt die Betrachtung der Bewertung von Innovationen aus privatwirtschaftlicher Sicht ein. Die beiden nachfolgenden Studien zielen auf die Unternehmenswertentwicklung im Laufe der Neuprodukteinführung ab. Die Studien werden jeweils mit einem Anwendungsfall aus dem Medienbereich durchgeführt.

Neue Kinofilme werden mittlerweile regelmäßig mit Budgets größer 100 Mio. USD ausgestattet. Entsprechend groß ist das Investitionsrisiko für die Produzenten, die aufgrund der Charakteristik des Lebenszyklus eines Films nur wenige Wochen Zeit haben, die Investition wieder zu verdienen und einen Gewinn einzustreichen (Prosser, 2002). Anders als bei gleichmäßiger verlaufenden Geschäftsmodellen bietet sich für einen Investor die Bewertung anhand von Vergangenheitsdaten nicht an (Rappaport, 1998). Eine Möglichkeit für den Investor bietet die Analyse des Einflusses auf den Aktienkurs mit dem Verfahren der Event Study.

Die Forschung hat sich bereits mit der Bewertung von neuen Produkten in Bezug auf den Unternehmenswert auseinandergesetzt. Dabei wurden jedoch immer nur (*Vor-)*Ankündigungen neuer Produkte untersucht (Chaney, Devinney & Winer, 1991; Kelm, Narayanan & Pinches, 1995; Sorescu, Shankar & Kushwaha, 2007). Der *tatsächliche* Produkterfolg im Markt wurde nicht bewertet.

Der Grund hierfür liegt in der Charakteristik der Event-Study. Bei der ersten Ankündigung eines neuen Produktes gibt es noch keine Erwartungen im Markt. Hier kann unmittelbar untersucht werden, ob die reine Produktankündigung einen Einfluss auf den Unternehmenswert hat. Ab diesem Zeitpunkt bestehen dann aber Erwartungen im Markt – und damit auch Bewertungen in den Aktienkursen, die diese Erwartungen widerspiegeln. Wird das Produkt dann

tatsächlich in den Markt eingeführt und werden erste Verkaufserfolge offensichtlich, so muss zunächst untersucht werden, inwiefern die Erwartungen über- oder untererfüllt werden. Gelingt dies nicht, so hat man einen Datensatz, der beispielsweise zu ungefähr gleichen Teilen Beobachtungen zu Über- und Untererfüllungen der Erwartungen beinhalten kann. Ein Test, ob die Informationen zu den neuen Produkten die Kurse bewegt, würde dann fälschlicherweise insignifikant werden, da sich die Beobachtungen gegenseitig aufheben.

Das Umgehen dieses Problem, indem die *erste* Vorankündigung zu einem neuen Produkt untersucht wird, ist bei Filmen nicht möglich, da die Informationen zu einem Film langsam in den Markt diffundieren und sich über die Zeit verdichten. In dem Beitrag *Neuprodukteinführungen in der Filmindustrie: Wie reagieren Kapitalmarktinvestoren auf den Umsatzerfolg neuer Kinofilme?* lösen wir dieses Problem zum ersten Mal, indem wir die Erwartungen im Markt schätzen und so die Beobachtungen in zwei Gruppen aufteilen können: In die Gruppe der Filme, die die Erwartungen übertreffen, und in die Gruppe der Filme, die hinter den Erwartungen zurückbleiben. Mit diesem innovativen Ansatz leisten wir neben dem inhaltlichen Beitrag der Bewertung von Filmerfolgen an der Börse auch einen methodischen Beitrag, der auch international aufgegriffen und hochrangig publiziert wurde (Joshi & Hanssens, 2009).

Der dritte empirische Beitrag *The Effect of Unexpected Chart Positions on the Firm Value of Music Labels. An Event Study of Album Success* untersucht im Kern auch den Einfluss neuer Produkte auf den Unternehmenswert. Untersucht wird der Einfluss des Erfolgs von Produkt-Samples, die einen Teil des eigentlichen Hauptprodukts darstellen und vor eben diesem auf den Markt gebracht werden. Sie vermitteln dem Kunden einen ersten Eindruck. Ziel der Samples ist es, den Kunden, der zunächst nur das Sample gekauft hat, von der Qualität bzw. dem Wert des Hauptprodukts zu überzeugen, damit dieser es dann kauft (Heiman, McWilliams, Shen & Zilberman, 2001).

Natürlich können auch Investoren den Erfolg von Samples beobachten und ihre Erwartungen hinsichtlich des Hauptprodukts bilden. Am Beispiel der Musik-Industrie wird untersucht, wie sich die Erwartungen zum Hauptprodukt bereits durch die Samples entwickeln und welchen Einfluss das Übertreffen bzw. das Verfehlen dieser Erwartungen auf den Wert des jeweiligen Unternehmens hat.

Die Musikindustrie verfährt bei der Einführung neuer Musikalben nach dem Schema, zunächst eine oder mehrere Singles zu veröffentlichen (Yanbin & Min, 2006). Singles dienen als Sample für das Album, das Hauptprodukt. Die Chartpositionen der Singles für die Investoren

ist frei ersichtlich (Bradlow & Fader, 2001). So bilden die Investoren bereits vor der Veröffentlichung des Musikalbums ihre Erwartungen darüber, welche Einnahmen das Unternehmen wohl erzielen müsste (Chellappa & Shivendu, 2005; Hirschman & Holbrook, 1982).

Mit dieser Studie wird das erste Mal die Bedeutung von Samples für den Unternehmenswert untersucht. Unternehmen können damit ihren Veröffentlichungsprozess optimieren und insbesondere versuchen, Einfluss auf die Erwartungsbildung zu nehmen. Zentrales Ergebnis der Studie ist, dass Unternehmen, die die in die Alben gesetzten Erwartungen verfehlen, dafür mit einem Abschlag auf den Unternehmenswert bestraft werden. Sollte das Album besser abschneiden, als dies zu erwarten war, so können die Unternehmen daraus keine Wertsteigerung erzielen. Dieses Ergebnis entspricht den Erwartungen, die sich aus der Theorie des Negativity Bias ergeben (Akhtar, 2012; Kahneman & Tversky, 1979).

3 Hinweise zur Ko-Autorenschaft

Von den fünf eingereichten Artikeln wurden drei in Ko-Autorenschaft erstellt. Bei dem Beitrag „*Neuprodukteinführungen in der Filmindustrie: Wie reagieren Kapitalmarktinvestoren auf den Umsatzerfolg neuer Kinofilme?*“ besteht der Beitrag des Verfassers insbesondere in der Durchführung der zu Grunde liegenden Event-Study sowie in der Mitentwicklung des innovativen Ansatzes zur Bestimmung von Erwartungen im Markt. Bei den übrigen Artikeln spiegelt die Reihenfolge der Autoren den Beitrag zu dem Artikel wieder.

Literatur

- Aerts, K. & Schmidt, T. (2008).** Two for the price of one? Additionality effects of R&D subsidies: A comparison between Flanders and Germany. *Research Policy*, 37, 806-822.
- Agrawal, J. & Kamakura, W. A. (1995).** The Economic Worth of Celebrity Endorsers: An Event Study Analysis. *Journal of Marketing*, 59 (July), 56-62.
- Akhtar, S., Faff, R., Oliver, B. & Subrahmanyam, A. (2012).** Stock Salience and the Asymmetric Market Effect of Consumer Sentiment News. *Journal of Banking & Finance*, 36, 3290-3301.
- Arrow, K.J. (1962).** Economic welfare and the allocation of resources of invention. In: Nelson, R.R. (Ed.), *The Rate and Direction of Inventive Activity: Economic and Social Factors*. Princeton, 361-392.
- Bradlow, E. T. & Fader, P.S. (2001).** A Bayesian Lifetime Model for the "Hot 100" Billboard Songs. *Journal of the American Statistical Association*, 96 (454), 368-81.
- Branstetter, L. & Sakakibara, M. (1998).** Japanese Research Consortia: A Microeconomic Analysis of Industrial Policy. *Journal of Industrial Economics*, 46 (2), 207-233.
- Chaney, P. K., Devinney, T. M. & Winer, R. S. (1991).** The Impact of New Product Introductions on the Market Value of Firms. *Journal of Business*, 64 (4), 573-610.

- Chellappa, R. K. & Shivendu, S. (2005).** Managing piracy: Pricing and sampling strategies for digital experience goods in vertically segmented markets. *Information Systems Research*, 16 (4), 400-17.
- Clarysse, B., Wright, M. & Mustar, P. (2009).** Behavioural additionality of R&D subsidies: A learning perspective. *Research policy*, 38, 1517-1533.
- David, P.A., Hall, B.H. & Toole, A.A. (2000).** Is public R&D a complement or substitute for private R&D? A review of the econometric evidence. *Research Policy*, 29, 497-529.
- Fama, E. F. (1970).** Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, 25 (2), 383-417.
- Fama, E. F., Fisher, L., Jensen, M. C. & Roll, R. (1969).** The Adjustment of Stock Prices to New Information. *International Economic Review*, 10 (1), 1-21.
- Görg, H. & Strobl, E. (2007).** The Effect of R&D Subsidies on Private R&D. *Economica*, 74, 215-234.
- Griliches, Z. & Regev, H. (1995).** Firm productivity in Israeli industry 1979-1988. *Journal of Econometrics*, 65 (1), 175-203.
- Heiman, A., McWilliams, B., Shen, Z.H. & Zilberman, D. (2001).** "Learning and forgetting: Modeling optimal product sampling over time," *Management Science*, 47 (4), 532-46.
- Hill, C. W. L. & Rothaermel, F. T. (2003).** The Performance of Incumbent Firms in the Face of Radical Technological Innovation. *Academy of Management Review*, 28 (2), 257-274.
- Hirschman, E. C. & Holbrook, M.B. (1982).** Hedonic Consumption: Emerging Concepts, Methods and Propositions. *Journal of Marketing*, 46 (3), 92-101.
- Hujer, R. & Radic, D. (2005).** Evaluating the Impacts of Subsidies on Innovation Activities in Germany. *Scottish Journal of Political Economy*, 52 (4), 565-586.
- Joshi, A. M., & Hanssens, D. M. (2009).** Movie Advertising and the Stock Market Valuation of Studios: A Case of "Great Expectations"? *Marketing Science*, 28 (2), 239-250.
- Kahneman, D. & Tversky, A. (1979).** Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47 (2), 263-91.
- Kelm, K. M., Narayanan, V. K. & Pinches, G. E. (1995).** Shareholder Value Creation During R&D Innovation and Commercialization Stages. *Academy of Management Journal*, 38 (3), 770-786.
- Koku, P. S., Jagpal, H. S. & Viswanath, P. V. (1997).** The Effect of New Product Announcements and Preannouncements on Stock Price. *Journal of Market Focused Management*, 2 (2), 183-199.
- Lerner, J. (1999).** The Government as Venture Capitalist: The Long-Run Impact of the SBIR Program. *Journal of Business*, 72 (3), 285-318.
- McWilliams, A. & Siegel, D. (1997).** Event Studies in Management Research: Theoretical and Empirical Issues. *Academy of Management Journal*, 40 (3), 626-657.
- Merito, M., Giannangeli, S. & Bonaccorsi, A. (2010).** Do incentives to industrial R&D enhance research productivity and firm growth? Evidence from the Italian case. *International Journal of Technology Management*, 49 (1), 25-48.
- Nelson, R.R. (1959).** The Simple Economics of Basic Scientific Research. *Journal of Political Economy*, 67 (3), 297-306.
- Ortega-Argilés, R., Vivarelli, M. & Voigt, P. (2009).** R&D in SMEs: a paradox? *Small Business Economics*, 33 (3), 3-11.

- Prosser, E. K. (2002).** How Early Can Video Revenue Be Accurately Predicted? *Journal of Advertising Research*, 42 (2), 47-55.
- Rappaport, A. (1998).** *Creating Shareholder Value: A Guide for Managers and Investors* (2. ed.). New York: Free Press.
- Sorescu, A., Shankar, V. & Kushwaha, T. (2007).** New Product Preannouncements and Shareholder Value: Don't Make Promises You Can't Keep. *Journal of Marketing Research*, 44 (3), 468-489.
- Tellis, G. J. & Johnson, J. (2007).** The Value of Quality. *Marketing Science*, 26 (6), 758-773.
- Wallsten, S.J. (2000).** The effects of government-industry R&D programs on private R&D: the case of the Small Business Innovation Research program. *RAND Journal of Economics*, 31 (1), 82-100.
- Wolff, G.B. & Reinthaler, V. (2008).** The effectiveness of subsidies revisited: Accounting for wage and employment effects in business R&D. *Research Policy*, 37, 1403-1412.
- Yanbin, T. & Min, L. (2006).** An Experimental and Analytical Study of on-Line Digital Music Sampling Strategies. *International Journal of Electronic Commerce*, 10, 39-70.

A.Event-Studies

Björn Goerke

Erschienen als:

Goerke, Björn (2009): Event Studies, in: Albers, Sönke, Daniel Klapper, Udo Konradt, Achim Walter und Joachim Wolf (Hrsg.): Methodik der empirischen Forschung, 3. Auflage, Wiesbaden 2009, 467-484

1 Einleitung

Das Ziel jeder Maßnahme eines Unternehmens ist die Steigerung des Unternehmensgewinns und damit auch des Unternehmenswertes am Markt. Wie sehr eine einzelne Maßnahme der Unternehmenswert beeinflusst, kann intern beispielsweise mit Hilfe von Discounted-Cash-Flow-Modellen ermittelt werden. Dieses Vorgehen erfordert jedoch interne Daten und führt abhängig von der individuellen Risikoeinschätzung zu unterschiedlichen Bewertungen (Brealey, Myers und Allen, 2008, S. 88 ff.). Eine objektive Bewertung ist für Dritte nicht möglich. Einen alternativen Ansatz bietet die Event-Study-Methode, die zur Schätzung von Kapitalmarktreaktionen eingesetzt wird. Der Aktienkurs stellt ein einheitliches Maß für die Bewertung eines Unternehmens dar. Reagieren die Kapitalmärkte auf eine Unternehmensnachricht mit Kursauf- oder abschlägen, so können diese Reaktionen allgemein beobachtet und bewertet werden.

Insofern kann für börsennotierte Unternehmen an den Börsen gemessen werden, welchen Einfluss einzelne Maßnahmen auf den Wert eines Unternehmens haben. Auf diesen wirken viele unterschiedliche Faktoren, so zum einen die allgemeine Marktbewegung und zum anderen die unternehmensspezifischen Informationen. Meistens liegen zu einem einzelnen Unternehmen mehrere relevante Informationen gleichzeitig vor, so dass durch die Überlagerung der Informationen für die Zuordnung von Einflüssen auf die Kursbewegung eine Methode eingesetzt werden muss, um die Zuordnung und Messung zu ermöglichen.

Die Event-Study-Methode wurde entwickelt, um diesen Anforderungen Rechnung zu tragen. Die Methode stammt aus der Finanzwirtschaft und findet mittlerweile vermehrt Anwendung im Marketing und Management (Agarwal und Bayus, 2002; Agrawal und Kamakura, 1995;

Clement, Fischer und Goerke, 2007; Sorescu, Shankar und Kushwaha, 2007). Als Event wird hierbei eine Information bezeichnet, die zu einem Unternehmen bekannt wird. Dies kann beispielsweise die Ankündigung eines neuen Produktes sein. Mit Hilfe dieser Methode können die Einflüsse sowohl in ihrer Richtung als auch in ihrer Stärke ermittelt werden. Gemessen werden die Reaktionen auf Basis der Renditen.

Im Folgenden werden die Methode und ihre Einsatzmöglichkeiten in der betriebswirtschaftlichen Forschung vorgestellt. Da die Methode auf der Annahme effizienter Kapitalmärkte fußt, wird zunächst zum Verständnis hergeleitet, wann mit einer Kursreaktion an einer Börse zu rechnen ist. Danach wird die Methode schrittweise erläutert. Die für eine Anwendung notwendigen Schritte werden präsentiert und an einem Beispiel erläutert. Der Beitrag schließt mit einer Zusammenfassung und konkreten Hinweisen für die praktische Anwendung.

2 Kapitalmarkteffizienz

Unter der Annahme der Kapitalmarkteffizienz reflektiert ein Aktienkurs P_{it} des Unternehmens i alle im Zeitpunkt t zur Verfügung stehenden Informationen zu dem Unternehmen. Eine Arbitragemöglichkeit kann sich durch öffentliche Informationen nicht ergeben. Eine Kursanpassung erfolgt erst wieder, wenn zu dem Unternehmen neue Informationen veröffentlicht werden, die vom Markt als bewertungsrelevant angesehen werden. Die Anpassung erfolgt dann ohne zeitliche Verzögerung (Fama, 1965; Fama et. al, 1969). Diese Annahme wurde mittlerweile in einer Vielzahl empirischer Studien belegt (Fama, 1991). Die Annahme wird formal durch

$$E(P_{i,t+1} | \Phi_t) = [1 + E(R_{i,t+1} | \Phi_t)]P_{it}, \quad (1)$$

beschrieben, wobei Φ_t die bekannten Informationen zum Zeitpunkt t , P_{it} den Kurs von Aktie i zum Zeitpunkt t , $R_{i,t+1}$ die Rendite zum Zeitpunkt $t+1$ und E den Erwartungswert des Wertes von Unternehmen i darstellen (Fama, 1970). Der Erwartungswert des Kurses ist also abhängig von der Menge an Informationen. Ändert sich diese, so erfolgt eine Neubewertung des Unternehmens.

Die prozentuale Änderung des Aktienkurses in einer Periode wird als Rendite bezeichnet. Sie dient der Untersuchung in einer Event-Study. Formal ergibt sich die Rendite aus

$$R_{it} = [P_{it} + d_{it} - P_{i,t-1}] / P_{i,t-1}, \quad (2)$$

wobei R_{it} die Rendite der Aktie i am Tag t , und d_{it} den aus dem untersuchten Event zusätzlich generierten Cashflow pro Aktie i im Zeitpunkt t bezeichnen (Clement, Fischer und Goerke, 2007).

3 Event-Study-Methode

Nachdem mit der Kapitalmarkteffizienz die theoretische Grundlage für die Anwendung der Event-Study-Methode vorgestellt wurde, folgt jetzt die Darstellung der einzelnen Bestandteile.

3.1 Events

Zur Untersuchung von Börsenreaktionen auf neue Informationen muss zunächst ein so genannter Event definiert werden. Dieser Event ist die neue Information, die zu einem Unternehmen bekannt wird. Dieser Event kann beispielsweise die Ankündigung einer Neuproduktentwicklung (Chaney und Devinney, 1992; Koku, Jagpal und Viswanath, 1997) oder die Einführung eines neuen Produktes sein (Chaney, Devinney und Winer, 1991). Es ist nicht möglich, eine bereits bekannte Information als Event zu definieren, da diese aufgrund der Kapitalmarkteffizienz bereits in den Unternehmenswert eingepreist wäre. Die Information muss also für den Markt völlig neu sein (McWilliams und Siegel, 1997).

Weiterhin sollte die als Event definierte Information auch nicht vom Kapitalmarkt antizipiert werden können (McWilliams und Siegel, 1997). Auch antizipierte Informationen werden aufgrund der Kapitalmarkteffizienz in den Kurs integriert. Bei antizipierten Informationen kommt es häufig dazu, dass zunächst nur ein Teil der Marktteilnehmer reagiert, so dass sich die Reaktion über einen längeren Zeitraum hinziehen kann (Campbell, Lo und MacKinlay, 1997).

Zusätzlich sollte der Event weitere Anforderungen erfüllen (Campbell, Lo und MacKinlay, 1997; McWilliams und Siegel, 1997). Der Event muss so exakt wie möglich definiert werden. Durch eine klare Definition des zu untersuchenden Events kann sichergestellt werden, dass die Ergebnisse nicht durch zusätzliche, andere Events verwässert oder verzerrt werden. Darüber hinaus muss sichergestellt sein, dass die Veröffentlichung einer neuen Information eindeutig als Event eingestuft werden kann. Hierzu ist es notwendig, dass die vorliegende Information mit Sicherheit die erste frei verfügbare Veröffentlichung ist, die von den Marktteilnehmern zum Handeln der Aktien genutzt werden kann. Interne Papiere, die den Vorschriften des Insiderhandels unterliegen, können keine Verwendung finden. Die Informationsquellen,

die den Anforderungen an eine Event-Study gerecht werden, werden im kommenden Abschnitt vorgestellt.

Zur Untersuchung der Reaktion auf einen Event bestehen zwei Möglichkeiten. Entweder es wird getestet, ob der Kapitalmarkt überhaupt reagiert oder es wird getestet, ob der Kapitalmarkt in eine Richtung – steigend oder fallend – reagiert. Hierbei gilt es zu beachten, dass die Kapitalmarkttheorie zunächst nur die Argumentation für die sofortige Kursreaktion liefert. Soll auch die Richtung der Kursreaktion getestet werden, so muss die Richtungshypothese aus dem Kontext heraus unter Verwendung einer geeigneten Theorie entwickelt werden.

Zur Veranschaulichung dienen zwei Artikel. Agrawal und Kamakura (1995) untersuchen in Ihrer Studie zu Prominenten, die als Werbeträger für Unternehmen und Produkte gewonnen werden konnten, ob diese Werbekampagnen von den Kapitalmärkten honoriert werden oder ob die Kapitalmärkte in dieser Maßnahme eine kostenintensive und wertlose Maßnahme sehen und die Unternehmen mit einem Kursabschlag bestrafen. Eine Hypothese hinsichtlich der Reaktionsrichtung wird nicht entwickelt. Vielmehr steht im Mittelpunkt, ob eine entsprechende Werbemaßnahme überhaupt einen Effekt erzielt.

Clement, Fischer und Goerke (2007) untersuchen die Kursreaktion auf unerwartet gut und auf unerwartet schlecht gelaufene Kinofilme. Hierzu stellen sie jeweils eine Hypothese einer positiven und einer negativen Kursreaktion auf, die einzeln getestet werden.

3.2 Datensatz

Die Qualität des verwendeten Datensatzes wird insbesondere durch zwei Faktoren bestimmt. Zum einen sind es die Event-Daten, die vor allem hinsichtlich des zeitlichen Bezugs exakt sein müssen, und zum anderen sind es Informationen zu anderen Events, die zu dem gleichen Zeitpunkt oder in zeitlicher Nähe geschehen. Diese Ereignisse werden Confounding Events genannt. Die Überlagerung von Events kann zu Verzerrungen der Ergebnisse führen.

3.2.1 Event-Daten

Für eine Event-Study müssen für einen definierten Untersuchungszeitraum zum einen Daten über die Aktienkurse der untersuchten Unternehmen vorliegen und zum anderen Informationen hinsichtlich der Events. Abhängig vom Untersuchungsrahmen werden zusätzliche Kontrollvariablen benötigt.

Die Börsendaten werden durch eine Reihe von Datenanbietern bereitgestellt. Es ist insbesondere darauf zu achten, dass die verwendeten Daten um Kapitalmaßnahmen wie Splits oder Dividendenausschüttungen bereinigt werden (McWilliams und Siegel, 1997). Anerkannte

Lieferanten für Börsendaten im wissenschaftlichen Umfeld sind beispielsweise das Center of Research in Security Prices (CRSP) oder Compustat. Kostenlose Datensätze können beispielsweise über Google Finance oder Yahoo! Finance erstellt werden.

Die Börsendaten sollten zeitlich sehr genau sein (Fama, 1991). Es bieten sich die Tageschlusskurse der relevanten Börsen an. Spezielle Fragestellungen können andere Spezifikationen wie stundengenaue Kurse erfordern, als Standard hat sich jedoch die taggenaue Untersuchung durchgesetzt. Im Folgenden Abschnitt werden die Vor- und Nachteile kürzerer und längerer Intervalle am Beispiel so genannter Confounding Events erläutert.

Neben den Kursdaten müssen auch die Informationen zu den Events im gleichen Zeitintervall vorliegen. Entscheidend ist die Tatsache, dass es sich bei den vorliegenden Quellen um die erste Veröffentlichung handelt. Für US-Daten stellt die Datenbank der U.S. Securities and Exchange Commission (SEC) eine Basis dar. Die SEC archiviert die Meldungen für alle in den USA gelisteten Unternehmen. Zusätzlich sollten weitere Quellen genutzt werden. Hierzu zählen die Investor-Relations-Informationen der untersuchten Unternehmen und Wirtschaftspublikationen wie beispielsweise das Wall Street Journal. Je nach Kontext können weitere Datenbanken oder Informationsquellen von Bedeutung sein. Die Genauigkeit der Event-Informationen limitiert die Qualität der Studie. Um Synergien zu nutzen, sollte die Suche nach den Event-Informationen direkt mit der Überprüfung auf überlappende Events verknüpft werden. Die Bedeutung der Confounding Events wird im folgenden Abschnitt näher erläutert.

Sollen in die Studie auch Kontrollvariablen einfließen, so muss überprüft werden, ob diese über die Zeit variieren können. Sollte das der Fall sein, so ist auch hier die Erfassung im gleichen Zeitintervall notwendig. Soll die Unternehmensgröße, gemessen in der Marktkapitalisierung, zur Erklärung genutzt werden, ist es notwendig, die Kapitalisierung im Intervall des Events zu erfassen.

3.2.2 Confounding Events

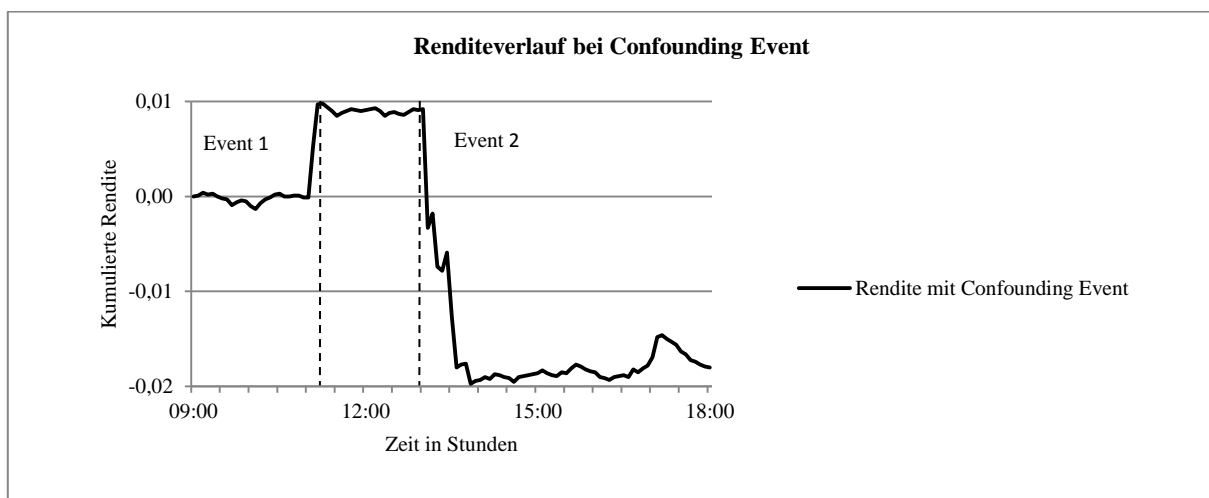
Neben den allgemeinen Marktinformationen, die im Laufe eines Börsentages bekannt werden, und den im Fokus stehenden Event-Daten können zusätzliche Informationen zu einer untersuchten Firma veröffentlicht werden, die Einfluss nehmen können. So ist eine Konstellation denkbar, in der ein Unternehmen vormittags eine Veröffentlichung vornimmt, die den Börsenkurs nach oben treibt und wenig später eine negative Veröffentlichung eines Analysten veröffentlicht wird, z.B. eine „Sell“-Einstufung mit Herabsetzung des Kursziels. Dies wäre

ein Signal, dass der Analyst kein Potenzial in der Aktie sieht und mit fallenden Kursen rechnet. Darauf folgen häufig tatsächlich fallende Kurse.

Abbildung 1 zeigt schematisiert eine entsprechende Konstellation. Zunächst reagiert der Kurs positiv auf die Unternehmensmeldung um 11.00 Uhr. Danach folgt im 13.00 Uhr die Negativreaktion des Marktes auf die Analysteneinschätzung. Sollte die Unternehmensmeldung ein Event im Sinne der Studie sein, so kommt es durch diese überlappende Information der Analysteneinschätzung zu einer Verzerrung des Bildes. Während der Markt die Unternehmensmeldung positiv aufgenommen hat und der Kurs entsprechend um ein Prozent anstieg, würde die Betrachtung des Events auf Basis des Schlusskurses die Schlussfolgerung zulassen, dass der Event einen Abschlag von zwei Prozent verursacht hat.

Zur Reduzierung dieses Problems bietet sich die Wahl kürzerer Zeitintervalle an. So könnte man beispielsweise nur die zwei Stunden nach dem Event betrachten. Durch diese Verkürzung des Intervalls würde die Wahrscheinlichkeit, Confounding Events zu erhalten, sinken. Hier gilt es jedoch zu bedenken, dass Kapitalmärkte in der Realität teilweise nicht sofort reagieren und die Reaktion nicht innerhalb weniger Minuten abgeschlossen ist. Durch eine Verkürzung des Intervalls besteht also das Risiko, Teile der Kursbewegung nicht zu messen (Dies kann auch noch für ein Intervall von einem Tag gelten. Eine Verlängerung des Intervalls steigert dann jedoch wiederum die Wahrscheinlichkeit, Confounding Events zu erhalten). Soll die Untersuchung mit anderen Intervallen als Tagen durchgeführt, so muss zunächst geprüft werden, inwiefern die erforderlichen Informationen über die Zeitpunkte verlässlich vorliegen.

Abbildung 1: Confounding Event



Um Einflüsse von Confounding Events unabhängig von der Wahl des Intervalls aus der Studie zu entfernen, ist es notwendig, jeden einzelnen Event auf überlappende Events zu untersuchen (Campbell, Lo und MacKinlay, 1997; McWilliams und Siegel, 1997). Diese müssen

zunächst markiert werden. Bei wenigen Confounding Events bietet sich die Elimination des betreffenden Events aus der Studie an, damit keine Verzerrungen durch überlappende Events in den Datensatz einfließen. Sollte es zu einer Vielzahl von Confounding Events kommen, sollte geprüft werden, ob man das Problem durch eine Verkürzung des Intervalls oder des Eventfensters lösen kann. Erschwert wird die Untersuchung der überlappenden Events durch den Umstand, dass auch die Informationen von anderen Unternehmen zu betrachten sind. Sollte eine Information zu einem Unternehmen aus der gleichen Branche bekannt werden, die als branchenrelevant einzustufen ist, so muss davon ausgegangen werden, dass auch die Werte der anderen Unternehmen der gleichen Branche an dem Tag hiervon beeinflusst werden, was wiederum bedeutet, dass auch diese verzerrend wirken können. Deshalb müssten auch diese Confounding Events kontrolliert werden, was jedoch in der wissenschaftlichen Literatur selten berichtet wird.

3.3 Abnormale Rendite

Ziel einer Event Study ist es festzustellen, ob an den Tagen des Events Reaktionen zu beobachten sind. Hierzu werden so genannte abnormale Renditen ermittelt (englisch: abnormal returns). In der deutschen Forschung werden auch die Begriffe Über- und Unterrenditen verwendet (Böhmer und Löffler, 1999). Dies sind Renditen, die nicht der erwarteten Kursbewegung im untersuchten Intervall entsprechen. Um die abnormalen Renditen zu bestimmen, sind mehrere Schritte notwendig, die im Folgenden beschrieben werden.

3.3.1 Schätzung der abnormalen Rendite

Zunächst muss ermittelt werden, welche Rendite an dem zu untersuchenden Tag für das Unternehmen zu erwarten ist. Hierfür ist es notwendig, ein Modell für die Rendite R_{it} zu verwenden, das eine Schätzung zulässt. Hierfür stehen diverse Modelle zur Verfügung. Sowohl in Simulationsstudien als auch in der Praxis hat sich das Marktmodell durchgesetzt: (Brown und Warner, 1985):

$$R_{it} = \alpha_i + \beta_i R_{mt} + v_{it}. \quad (3)$$

R_{it} wird gemäß Formel (2) aus den Schlusskursen der untersuchten Aktie ermittelt. Die Parameter α_i und β_i sind dabei in einer linearen Regression im Vorfeld des Events im so genannten Schätzfenster zu schätzen. R_{mt} ist die Tagesrendite eines zu wählenden Benchmarks und v_{it} steht für den Fehlerterm. α_i kann als autonome Rendite interpretiert werden, die regelmäßig durch das Unternehmen erzielt wird. β_i gibt an, inwiefern sich die beobachtete Aktie mit dem Benchmark bewegt. Als Benchmark wird häufig ein geeigneter Aktienindex gewählt. Die

Wahl des Benchmarks wird im folgenden Abschnitt näher erläutert. Aus (4) ergibt sich eine erwartete Rendite $E(R_{it})$ an Tagen ohne Event für Aktie i:

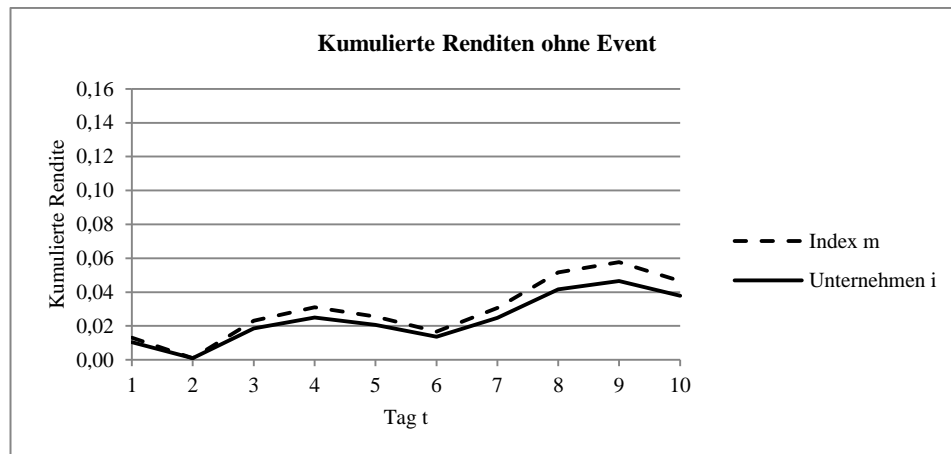
$$E(R_{it}) = \alpha_i + \beta_i R_{mt}, \quad (4)$$

wobei α_i und β_i die geschätzten Parameter aus (3) darstellen. Die erwartete Rendite ist dabei abhängig von der Performance des Benchmarks an dem beobachteten Tag. Abbildung 2 verdeutlicht diesen Zusammenhang exemplarisch. Der Aktienkurs des Unternehmens schwankt mit dem Markt, der in der Abbildung durch den Index repräsentiert wird.

Abbildung 3 hingegen verdeutlicht die Kursreaktion am Eventtag. In diesem Beispiel ist es der fünfte Tag. Bis zum vierten Tag schwankt der Kurs mit dem Markt. Am Eventtag erfolgt die Reaktion und danach schwankt der Kurs wiederum mit dem Markt. In diesem Fall gab es also auch nach dem Event am fünften Tag keine weiteren Events im betrachteten Zeitraum. Ansonsten hätte eine erneute Reaktion erfolgen müssen. Diese hätte wiederum dazu geführt, dass der Aktienkurs sein Niveau verlassen hätte. Um entsprechend die abnormale Rendite AR_{it} zu ermitteln, wird die Differenz aus der tatsächlichen und der erwarteten Rendite gebildet:

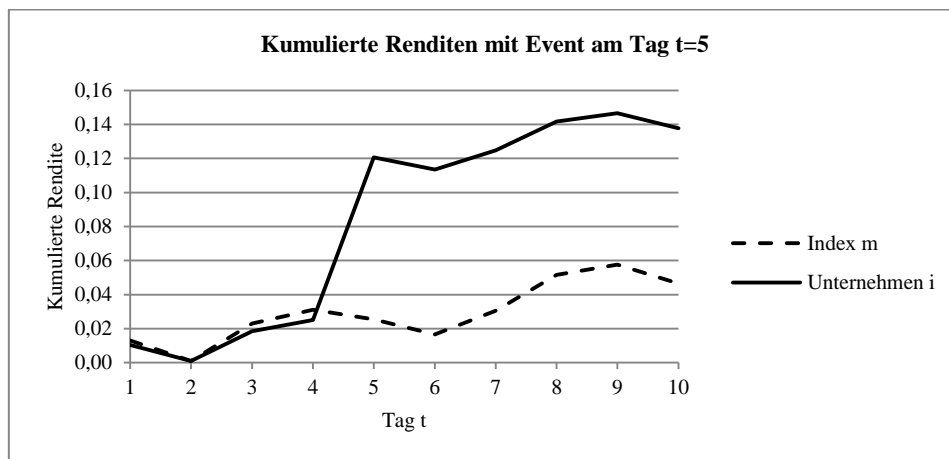
$$AR_{it} = R_{it} - E(R_{it}). \quad (5)$$

Abbildung 2: Kursverlauf ohne bewertungsrelevanten Event



Diese Berechnungsschritte müssen für jeden Event durchgeführt werden. Abschließend werden die abnormalen Renditen über die Events i hinweg aggregiert und können auf den signifikanten Einfluss des Events getestet werden (siehe Abschnitt 3.4).

Abbildung 3: Kursverlauf mit bewertungsrelevantem Event



3.3.2 Benchmarks

Der Wahl des Benchmarks kommt eine hohe Bedeutung zu. Der Benchmark des Marktmodells ist die Basis für die Berechnung der erwarteten Rendite. Nur wenn dieser Benchmark richtig gewählt wird, sind die Ergebnisse unverzerrt.

Etabliert ist die Verwendung eines breit gestreuten Indexes. Für US-bezogene Untersuchungen haben sich insbesondere drei Indizes etabliert. Zum einen ist dies der S&P 500 von Standard and Poor's und zum anderen sind es die CRSP-equal-weighted- und CRSP-value-weighted-Indizes, die vom Center of Research in Security Prices berechnet werden (Campbell, Lo und MacKinlay, 1997, S. 155). Der Vorteil dieser breit gestreuten Indizes liegt in der geringen Beeinflussung der Indizes durch die Events selbst. Der S&P 500 wird aus 500 Aktienkursen in den USA zusammengestellt. Wenn ein zu untersuchender Event Einfluss auf einzelne dieser Aktien nimmt, so wird die Beeinflussung des Indexes gering ausfallen. Anders wäre es bei der Verwendung des Dow Jones Industrial Index, der als Leitindex der USA gilt. Dieser wird aus 30 Aktienkursen berechnet. Ein einzelner Wert kann diesen Index beeinflussen, was wiederum Einfluss auf die berechneten abnormalen Renditen hätte.

Das hieraus folgende Endogenitätsproblem würde dazu führen, dass beispielsweise durch eine positive kursrelevante Nachricht sowohl der Aktienkurs des betrachteten Unternehmens als auch der Index, in dem die Aktie enthalten ist, in positiver Richtung beeinflusst werden. Gemäß (4) würde so auch der Erwartungswert der Rendite steigen. Die abnormale Rendite fällt dann entsprechend geringer aus. Das hat zwei Folgen: Zum einen wird der Effekt kleiner eingeschätzt, als er tatsächlich ist, und zum anderen wird dadurch der t-Wert für den folgenden Signifikanztest geschwächt. Insofern kann ein Event in Folge des beschriebenen Endogenitätsproblems fälschlicherweise als insignifikant eingestuft werden. Denkbar wäre auch die eigenhändige Berechnung des Benchmarks ohne das betreffende Unternehmen. Aufgrund

teilweise dynamischer Gewichtungverschiebungen in den Indizes wäre dieses Vorgehen jedoch fehleranfällig und unter Umständen nicht durchführbar, sofern die Vergangenheitsdaten für die Indexberechnungen nicht erhältlich sein sollten.

Abhängig von der Auswahl der zu untersuchenden Unternehmen ist der Benchmark-Index entsprechend bewusst zu wählen. Für Untersuchungen in anderen Märkten als den USA müssen die üblichen Benchmark-Indizes zunächst der Literatur entnommen werden.

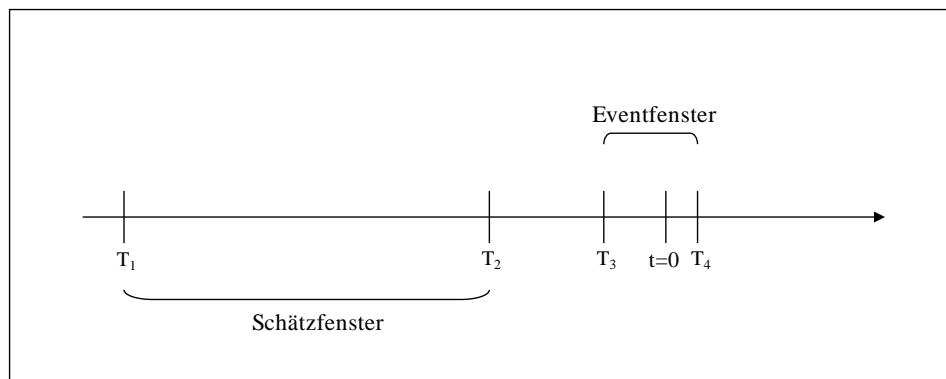
3.3.3 Schätz- und Eventfenster

Eine weitere Voraussetzung für die Durchführung der Event Study ist die Wahl der so genannten Schätz- und Eventfenster. Das Schätzfenster ist der Zeitraum, in dem die Parameter des Marktmodells geschätzt werden. Das Eventfenster stellt den Zeitraum dar, für den die abnormalen Renditen ermittelt werden. Abbildung 4 verdeutlicht den zeitlichen Zusammenhang.

$t=0$ bezeichnet dabei den Eventtag. Die Zeitpunkte T_1 bis T_4 sind alle in Relation zum Eventzeitpunkt zu verstehen. Es handelt sich hierbei nicht um einen Bezug zu Kalendertagen.

Die Wahl der Lage des Schätzfensters sollte mehreren Aspekten Rechnung tragen. Zum einen sollte es ausreichend groß sein, um stabile Parameter für das Marktmodell zu erhalten. Die Verwendung eines Zeitraums von einem Jahr hat beispielsweise neben den stabilen Parametern den Vorteil, dass der Zeitraum auch alle saisonabhängigen Effekte genau einmal abdeckt (Clement, Fischer und Goerke, 2007). Es kommt nicht dazu, dass beispielsweise die so genannte Berichtssaison der Unternehmen einen Einfluss auf die Schätzung der Parameter nimmt. Zusätzlich sollte das Schätzfenster in einem Mindestabstand zum Event gelegt werden. Im Vorfeld des Events kann es beispielsweise aus Gründen der Antizipation des Events zu Kursbewegungen kommen. Um zu vermeiden, dass die geschätzten Parameter bereits Einflüsse des Events enthalten, empfiehlt sich ein Abstand des Schätzfensters zum Event, mit dem sichergestellt ist, dass keine Einflüsse enthalten sein können. Dieser Abstand kann beispielsweise auf 30 Tage gesetzt werden (Geyskens, Gielens und Dekimpe, 2002). Die Entscheidung muss abhängig von der Fragestellung getroffen werden. Je mehr zu erwarten ist, dass der Event antizipiert werden kann, desto größer sollte der Abstand gewählt werden.

Abbildung 4: Umfang von Schätz- und Ereignisfenster



Quelle: in Anlehnung an Clement, Fischer und Goerke (2007)

Von noch höherer Bedeutung ist das Eventfenster. Das Eventfenster ist der Zeitraum um den Event, für den die täglichen abnormalen Renditen berechnet werden. Das Eventfenster wird meistens symmetrisch um den Event gelegt (Clement, Fischer und Goerke, 2007; Geyskens, Gielens und Dekimpe, 2002). Es umfasst häufig zwischen zehn und 20 Tagen. Dadurch wird zwei Aspekten Rechnung getragen. Zwar dürfte aufgrund der Annahme der Kapitalmarkteffizienz keine Reaktion vor dem Event zu verzeichnen sein, jedoch können antizipierte Informationen bereits Einfluss nehmen. Außerdem kann es zu einem so genannten Leakage kommen. Das bedeutet, dass einzelne Marktteilnehmer bereits vor der Veröffentlichung Informationen erhalten. Diese können bereits zu Kursbewegungen führen, die vor dem eigentlichen Event liegen (Brown und Warner, 1980, 1985; McWilliams und Siegel, 1997). Würde die Untersuchung der abnormalen Renditen erst am Tag des Events beginnen, so würde man die dargestellten Effekte nicht berücksichtigen können.

Ebenso sollte das Eventfenster einige Tage nach dem Event umfassen. Dadurch wird es möglich, wiederum zwei Aspekten Rechnung zu tragen. Erstens können Informationen zunächst verzögert zu Kursreaktionen führen, wenn die Marktteilnehmer die Informationen nicht sofort als eindeutiges Signal auffassen, sondern zunächst eine Meinung entstehen muss, und zum anderen wird es so möglich, potenzielle Gegenreaktionen der Marktteilnehmer zu beobachten, die entstehen können, wenn die Event-Information zunächst überbewertet wurde (De Bondt und Thaler, 1985).

3.4 Tests

Zur Beurteilung der Signifikanz des Einflusses eines Events auf den Aktienkurs der betrachteten Unternehmen müssen die errechneten abnormalen Renditen mit Hilfe eines Tests untersucht werden. In einem ersten Schritt werden die abnormalen Renditen, die für den Eventtag t

berechnet wurden, analog (6) addiert und durch die Anzahl N der untersuchten Events i dividiert.

$$AR_t = \sum_{i=1}^N AR_{it} / N \quad (6)$$

AR_t steht also für die durchschnittliche abnormale Rendite am Eventtag t . Zur Ermittlung der Signifikanz der so ermittelten durchschnittlichen abnormalen Rendite muss folgende Teststatistik aufgestellt werden:

$$t_0 = AR_t / \sigma_{AR} \quad (7)$$

mit

$$\sigma_{AR} = AR_t / [1 / N(N-1) \sum_{i=1}^N (AR_{it} - 1/N \sum_{i=1}^N AR_{it})^2]^{1/2} \quad (8)$$

Aufgrund der t-Verteilung der Teststatistik t_0 kann die Signifikanz den entsprechenden Tabellen entnommen werden (Brown und Warner, 1980; McWilliams und Siegel, 1997). Dieser Test auf Signifikanz der abnormalen Renditen an einem einzelnen Tag t kann für jeden einzelnen Tag im Eventfenster ausgeführt werden. Dann sind die Formeln (6) bis (8) entsprechend anzuwenden.

Neben den Tests für einzelne Tage kann auch der Test zusammenhängender Tage von Interesse sein. Dies ist insbesondere dann angezeigt, wenn es Hinweise auf eine allmähliche Durchsetzung der neuen Information am Markt gibt oder beispielsweise im Vorfeld eines Events Leakage beobachtet werden kann. Einige Autoren testen auch ohne a priori Information mehrere unterschiedliche Zeitfenster auf Signifikanz, um zu ermitteln, an welchen Tagen die Marktteilnehmer auf eine Information reagieren (Agrawal und Kamakura, 1995). Die notwendige Ermittlung der kumulierten abnormalen Renditen $CAAR_{[-t_1, t_2]}$ im Zeitfenster zwischen den Tagen t_1 und t_2 erfolgt in zwei Schritten. Zunächst werden alle abnormalen Renditen zu einem Event i gemäß (9) über die Tage im Zeitfenster aufaddiert:

$$CAR_{i[-t_1, t_2]} = \sum_{t=-t_1}^{t_2} AR_{it}. \quad (9)$$

Danach folgt die Ermittlung der durchschnittlichen kumulierten abnormalen Renditen $CAAR_{[-t_1, t_2]}$ über alle Events i gemäß (10):

$$CAAR_{[-t_1, t_2]} = \sum_{i=1}^N CAR_{i[-t_1, t_2]} / N. \quad (10)$$

Um die durchschnittlichen kumulierten abnormalen Renditen $CAAR_{[-t_1, t_2]}$ wiederum auf Signifikanz zu überprüfen, wird folgende Teststatistik t_{CAAR} gebildet (Campbell, Lo und MacKinlay, 1997):

$$t_{CAAR} = CAAR_{[-t_1, t_2]} / \sigma_{CAAR}, \quad (11)$$

mit

$$\sigma_{CAAR} = CAAR / [1 / (N(N-1)) \sum_{i=1}^N (CAR_i - 1/N \sum_{i=1}^N CAR_i)^2]^{1/2}. \quad (12)$$

Wiederum finden die t-Tabellen Anwendung.

3.5 Moderatorenanalyse

Durch die Berechnung der abnormalen Renditen und durch die Tests auf Signifikanz entsteht zunächst ein grobes Bild, ob der untersuchte Event die Aktienkurse bewegt. Erst durch eine anschließende Moderatorenanalyse kann ermittelt werden, welche Details des Events den Effekt verstärken oder abschwächen (McWilliams und Siegel, 1997).

Deswegen sollte für die abnormalen Renditen ein Modell mit erklärenden Faktoren formuliert werden. In diesem Modell stellen die abnormalen Renditen die abhängige Variable dar. Die erklärenden Faktoren sind entsprechend die unabhängigen Variablen (beispielsweise Unternehmensgröße, Branche und andere Charakteristika des Unternehmens oder des Events). Grundsätzlich kann dieses Modell mit der linearen Regression geschätzt werden (Campbell, Lo und MacKinlay, 1997). Sollte das Erklärungsmodell besondere Anforderungen an die Schätzung stellen, so muss auf angemessene Schätzmethoden wie beispielsweise die Strukturgleichungsmodelle (siehe Beitrag von Marisa Schlichthorst in diesem Buch) zurückgegriffen werden. Das Modell hat folgende Form:

$$AR_{it} = \gamma_0 + \sum_{j=1}^k \gamma_j x_{ij} + u_{it}, \quad (13)$$

wobei AR_{it} die abnormalen Renditen der Events i an dem zu erklärenden Tag t , γ_j die $k+1$ Koeffizienten für die erklärenden Variablen und u_{it} den Störterm darstellen. Die x_{ij} stehen für die unabhängigen Variablen im Moderatorenmodell.

Für die zu testenden Erklärungszusammenhänge sind zunächst theoriegeleitet Hypothesen aufzustellen. In ihrer Untersuchung zur Bewertung neuer Internetkanäle von börsennotierten Zeitungsverlagen stellen Geyskens, Gielens und Dekimpe (2002) zunächst sieben Hypothesen auf. Diese beziehen sich auf die Firmencharakteristika, auf die Einführungsstrategie des neuen

Kanals und auf die Charakteristika des Marktes. Während die beiden letztgenannten Kriterien individuell für diese Studie entwickelt wurden, sind die einzelnen Punkte, die die Merkmale der Unternehmen beschreiben, für jede Studie auf Relevanz zu prüfen. Zu diesen Kriterien zählt beispielsweise die Größe des Unternehmens. So kann es sein, dass die Reaktion auf einen Event bei kleineren Unternehmen heftiger ausfällt. Da es sich um prozentuale Reaktionen in Bezug zur Marktkapitalisierung handelt, müssen die Reaktionen bei kleineren Unternehmen höher ausfallen, um die gleiche absolute Wertveränderung abzubilden wie bei einem größeren Unternehmen mit derselben absoluten Wertveränderung. In der Untersuchung von Geyskens, Gielens und Dekimpe (2002) blieb die Firmengröße allerdings ohne signifikanten Effekt.

Neben der Notwendigkeit, die unternehmensbezogenen Effekte zu bestimmen, müssen weitere Aspekte zur Erklärung der Renditen herangezogen werden. Zu dem Zeitpunkt der Studie wurde die Frage der Schnelligkeit des Markteintritts in das Internet-Geschäft intensiv untersucht. Die Autoren tragen dieser Diskussion durch die Erarbeitung der Hypothese, dass die Rangposition des Markteintritts durch einen umgekehrten u-förmigen Verlauf beschrieben wird, Rechnung. Diese bilden sie ab, indem sie die Rangposition zunächst linear und zusätzlich quadriert in das Modell aufnehmen. Die Schätzer sind für beide Faktoren signifikant – für den linearen Schätzer stark positiv, für den quadrierten Schätzer klein und negativ. Dadurch konnte der angenommene Verlauf bestätigt werden. Die Erarbeitung des Moderatorenmodells ist für jede Untersuchung individuell vorzunehmen.

3.6 Robustheit

Die Ergebnisse der Event Study Untersuchung können durch das Design der Studie bzw. durch die Wahl einzelner Untersuchungsparameter beeinflusst werden. Um diesen Effekt auszuschließen bzw. um eine auftretende Beeinflussung beurteilen zu können, ist es notwendig, eine Untersuchung der Ergebnisse auf Robustheit gegen Veränderungen vorzunehmen (Geyskens, Gielens und Dekimpe, 2002).

Zunächst sollte das aufgestellte Modell mit anderen Benchmarks alternativ gerechnet werden. Vorangehend ist die Bedeutung der Wahl des Benchmarks verdeutlicht worden. Abhängig von dem zu Grunde liegenden Markt sollten alternative Indizes gewählt werden, um die Studie mit diesen zu wiederholen (Geyskens, Gielens und Dekimpe, 2002). Für die Untersuchung US-amerikanischer Unternehmen können dies der S&P 500, der Dow Jones Industrial, aber auch Branchenindizes wie der Dow Jones Media sein. Sollten sich hier grundlegend andere

abnormale Renditen ergeben, so ist das Design auf die angemessene Spezifikation hin zu überprüfen.

Neben der Wahl des Benchmarks sollten zusätzlich die Schätzfenster variiert werden (Clement, Fischer und Goerke, 2007). Hinsichtlich der Schätzfenster besteht das Risiko, systematisch Effekte im Vorfeld der Events mit aufzunehmen, die die Koeffizienten des Marktmodells beeinflussen. Sofern die Entscheidung für die Untersuchung für ein kurzes Schätzfenster gefallen ist, sollte die Berechnung zur Überprüfung der Robustheit mit einem deutlich erweiterten Schätzfenster wiederholt werden. Für lange Schätzfenster gilt das umgekehrte Verfahren. Nur wenn die Ergebnisse gegenüber dieser Variation stabil sind, können die Ergebnisse uneingeschränkt berichtet werden.

4 Anwendungsbeispiel

In diesem Beispiel wird exemplarisch dargestellt, wie die Untersuchung einer Neuprodukteinführung mit Hilfe der Event-Study-Methode durchgeführt werden kann. Als Event ist hier die Markteinführung elektronischer Produkte für den privaten Gebrauch definiert. Untersucht wird ein Zeitraum von zehn Jahren (1.1.1998 bis 31.12.2007). Tabelle 1 zeigt beispielhaft den Renditendatensatz, der für die Untersuchung notwendig ist. Für den Untersuchungszeitraum müssen die Renditen der verwendeten Indizes (CRSPw, CRSPe und DJI) als auch der relevanten Unternehmen (beispielhaft hier als Aktie 1 bis Aktie 5 bezeichnet) vorliegen.

Tabelle 1: Renditendatensatz von 1998 bis 2007

Datum	CRSPw	CRSPe	DJI	Aktie 1	Aktie 2	Aktie 3	Aktie 4	Aktie 5
31.12.07	0,0165	0,0125	0,0103	0,0156	0,0191	0,0211	-0,0097	0,0117
30.12.07	0,0214	0,0239	0,0297	0,0462	0,0127	0,0285	-0,0122	0,0106
.
.
.
02.01.98	-0,0087	-0,0103	-0,0162	0,0111	-0,0397	0,0004	0,0410	-0,0008
01.01.98	0,0004	0,0074	-0,0025	0,0009	0,0265	0,0038	0,0002	0,0011

In den restlichen Spalten finden sich zum einen die geschätzten Parameter α und β , die zugeordnete tatsächliche Rendite R_{it} am Eventtag, die erwartete Rendite $E[R_{it}]$ am Eventtag und schließlich die abnormale Rendite AR_{it} . Die Marktrendite ist durch R_{mt} angegeben.

Tabelle 2: Eventinformationen und Ergebnisse der Event-Study-Rechnung

Event	Datum	Firma	α	β	$E[R_{it}]$	R_{mt}	R_{it}	AR_{it}
Produkt 1	12.08.98	1	0,00022	1,15	0,0151	0,0130	0,0210	0,0058
Produkt 2	03.11.00	2	0,00019	1,00	-0,0088	-0,0090	0,0280	0,0368
Produkt 3	21.06.03	3	0,00010	0,80	0,0153	0,0190	0,0240	0,0087
.
.
.

Zusätzlich wird eine Aufstellung über die zu untersuchenden Events benötigt. Tabelle 2 zeigt in den ersten drei Spalten, wie die Informationen vorliegen müssen. Für jedes einzelne Produkt muss der Tag der Markteinführung bekannt sein. Zusätzlich müssen die Events eindeutig einem Unternehmen zugeordnet werden können.

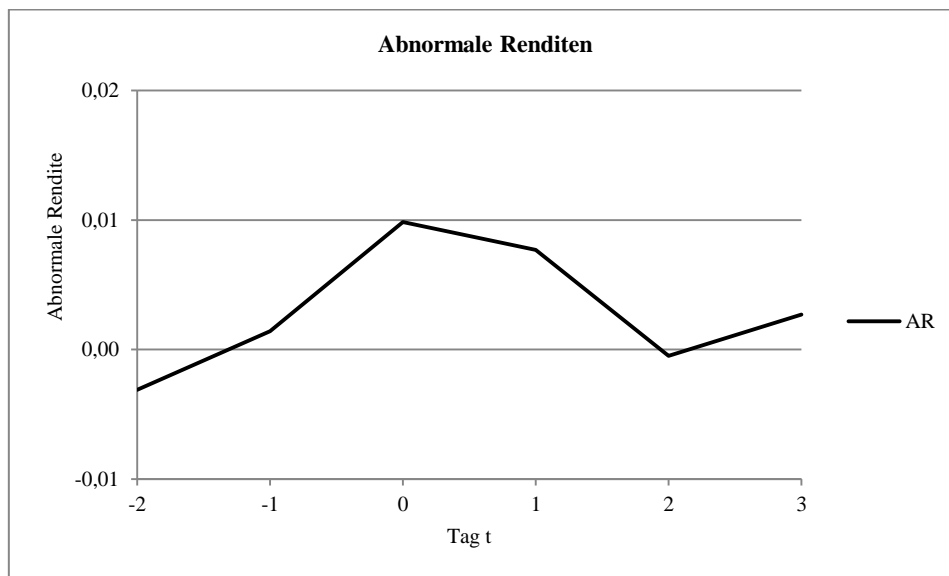
Tabelle 3: Kursverlauf ohne bewertungsrelevanten Event, eigene Erstellung

Tag	AR [%]	t-Wert	CAAR [%]	t-Wert
-2	-0,31	-0,24	-0,31	0,24
-1	0,14	0,20	-0,17	0,24
0	0,98	2,04	0,82	2,14
1	0,77	1,98	1,58	2,01
2	-0,05	-0,10	1,54	1,37
3	0,27	0,31	1,81	0,88

Anhand der Informationen aus Tabelle 2 können nun die durchschnittlichen abnormalen Renditen ermittelt werden. In Tabelle 3 sind die abnormalen Renditen für das Eventfenster von $t = -2$ bis $t = 3$ mit dem t-Wert abgetragen. Zusätzlich sind die kumulierten durchschnittlichen abnormalen Renditen ab Beginn des Eventfensters gebildet worden. Auch hierzu liegen die t-Werte vor. Nach diesem Beispiel ergeben sich signifikante Effekte für den Eventtag und den Folgetag. Die abnormalen Renditen liegen bei 0,98 % und 0,77 %.

Zusätzlich bietet sich die graphische Aufbereitung der Ergebnisse an. Abbildung 5 zeigt die täglichen durchschnittlichen abnormalen Renditen im Eventfenster. Für diese Abbildung werden die Renditen ab dem ersten Tag des Eventfensters addiert.

Abbildung 5: Tägliche abnormale Renditen



Hier sind deutlich die starken Reaktionen an den beiden signifikanten Tagen zu erkennen. Für die Beurteilung des Gesamteinflusses innerhalb des Schätzfensters sollte auch eine Abbildung der kumulierten durchschnittlichen Renditen erstellt werden. Abbildung 6 zeigt den Einfluss kumuliert über die Tage hinweg. Hier wird deutlich, dass die Tage vor und nach dem Event die Kurse nicht treiben. Insbesondere lässt sich auch keine Gegenreaktion feststellen, die ein Hinweis für eine zu heftige Reaktion unmittelbar nach dem Event wäre. Die Rendite schwankt nach dem Event um das neue Niveau.

Um Aufschluss darüber zu erhalten, welche Faktoren die abnormalen Renditen treiben, wird eine zusätzliche Moderatorenanalyse durchgeführt. Hierzu werden drei Variablen auf die CAAR analog (14) regressiert:

$$CAAR_i = \gamma_0 + \gamma_1 Gr\ddot{o}\ddot{e} + \gamma_2 Saison + \gamma_3 Medien + u_i \quad (14)$$

Zum einen sollen Größeneinflüsse und zum anderen Saisoneinflüsse kontrolliert werden. Hierzu wird die Unternehmensgröße durch den Umsatz gemessen. Die Saison wird durch eine Dummy-Variable dargestellt (Weihnachtsgeschäft 1, 0 sonst). Zusätzlich wird durch eine weitere Dummy-Variable abgebildet, ob es sich um ein Produkt zur Verwendung von Medien handelt (1 für MP3-Player etc., 0 für Küchengeräte u.ä.).

Abbildung 6: Kumulierte tägliche abnormale Renditen, eigene Erstellung

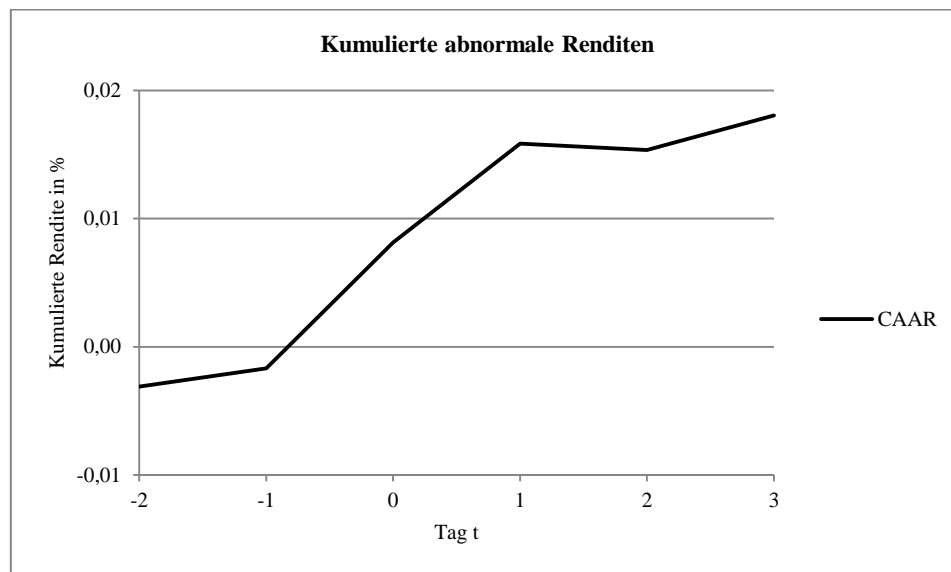


Tabelle 4 zeigt die Ergebnisse. Die Größe hat keinen Einfluss auf die Stärke der abnormalen Renditen. Die Dummy-Variablen jedoch schon. Die Produkteinführung in der Weihnachtssaison steigert den Wert um weitere 0,18 Prozentpunkte und Medienprodukte werden besser bewertet als sonstige elektronische Produkte, so dass diese einen Aufschlag von 0,05 Prozentpunkten bewirken.

Tabelle 4: Kursverlauf ohne bewertungsrelevanten Event

Variable	Schätzer	t-Wert	VIF
Größe	-0,01	-0,67	1,02
Saison	0,18	2,43	1,62
Medien	0,05	1,98	1,33

5 Erweiterungen

5.1 Umgang mit antizipierten Events

Die bislang vorgestellte Methodik kann nicht für jede Art von Event eingesetzt werden. Voraussetzung ist bislang, dass der Event nicht antizipiert werden kann. Das ist jedoch für eine Reihe von Fragestellungen nicht der Fall. Um trotzdem eine Event-Study durchführen zu können, muss die Erwartung, die im Markt zu dem Event vorliegt, geschätzt werden. Danach ist es möglich, die Erwartung über den Event mit der tatsächlichen Realisation abzugleichen und zu bestimmen, ob die Erwartungen über- oder untererfüllt wurden.

Clement, Fischer und Goerke (2007) untersuchen, inwiefern der Kinoerfolg neuer Filme Auswirkungen auf die Unternehmenswerte der Produzenten hat. Da ein Filmstart bereits frühzeitig bekannt ist, kann nicht angenommen werden, dass dieser Event nicht antizipiert wird.

Im Gegenteil ist davon auszugehen, dass die Marktteilnehmer Erwartungen über den Filmerfolg bilden. Diese werden durch eine Reihe von Faktoren getrieben, zu denen beispielsweise das Genre, Stars, Produktionsbudgets, namhafte Regisseure usw. zählen. Mit dem Marktmodell zur Bestimmung des erwarteten Umsatzes von Filmen von Elberse und Eliashberg (2003) bestimmen die Autoren die erwarteten Umsätze an den Kinokassen zum ersten, zweiten und dritten Wochenende.

Mit den erwarteten Umsätzen an den Wochenenden können die realisierten Umsätze, die sehr zeitnah von Variety zur Verfügung gestellt werden, abgeglichen werden. Danach kann das Sample in Filme, die Erwartungen übererfüllen, und in Filme, die die Erwartungen nicht erfüllen, unterteilt werden. Es ist so möglich, einen Event zu wählen, der ohne die Erwartungsschätzung für eine Event-Study nicht hätte verwendet werden können, da durch die ungefähr gleichmäßig verteilten Über- und Untererfüllungen der Erwartungen eine gemeinsame Untersuchung ein verzerrtes Bild ergeben hätte. Die positiven und negativen Abweichungen hätten sich teilweise gegenseitig aufgehoben.

5.2 Nachhaltigkeit der Abnormal Returns

An Kapitalmärkten können Überreaktionen auf Informationen beobachtet werden (De Bondt und Thaler, 1985). Das bedeutet, dass die Investoren nach der Veröffentlichung einer Nachricht stärker den Kurs bewegen, als die Nachricht dies veranlassen sollte. Um zu prüfen, ob der Markt daraufhin mit einer Gegenbewegung auf die Überreaktion reagiert und damit die Überrenditen wieder ausgleicht, bietet sich die Überprüfung der Stabilität der abnormalen Renditen an.

Hierzu können so genannte Buy-and-Hold-Portfolios gebildet werden. Analog zu den kumulierten abnormalen Renditen CAR_i für einen Event i werden in diesem Fall die abnormalen Renditen über den Zeitraum zwischen t_1 und t_2 betrachtet. Die Renditen werden jedoch nicht aufaddiert, sondern miteinander multipliziert:

$$BHAR_i = \left[\prod_{t=t_1}^{t_2} (1 + r_t^i) - 1 \right] - \left[(1 + \alpha_i)^{t_2 - t_1 + 1} - 1 \right] - \beta_i \left[\prod_{t=t_1}^{t_2} (1 + r_{mt}) - 1 \right] \quad (15)$$

Die durchschnittliche Buy-and-Hold abnormale Rendite ergibt sich analog aus:

$$Average BHAR = \frac{1}{N} \sum_{i=1}^N BHAR_i \quad (16)$$

Tellis und Johnson (2007) wählen ein Zeitfenster über zehn Tage ab dem Event. Sie testen die so ermittelten Renditen und stellen fest, dass die mit den kumulierten durchschnittlichen abnormalen Renditen ermittelten Effekte auch über einen längeren Zeitraum – von in diesem Fall zehn Tagen – signifikant sind.

5.3 Langfristige Analysen

Insbesondere für den Fall, dass Events erst über einen längeren Zeitraum ihre Wirkung auf den Unternehmenswert entfalten, bietet sich die Verwendung des Calender-Time Portfolio-Ansatzes an (Sorescu, Shankar und Kushwaha, 2007). Werden beispielsweise Neuproduktvorankündigungen veröffentlicht, so sendet das Unternehmen mit der Nachricht ein Signal an den Markt, dass zukünftig mit zusätzlichen Cashflows zu rechnen ist. Zu dem Zeitpunkt der Nachricht ist jedoch die Unsicherheit dieser Cashflows noch außerordentlich hoch. Erst im Zeitablauf reduziert sich das Risiko für die Investoren, wenn beispielsweise Fortschritte in der Produktentwicklung berichtet werden.

Im Calender-Time Portfolio-Ansatz werden anders als in der kurzfristig ausgerichteten Untersuchung der abnormalen Renditen, Portfolios der Unternehmen mit Event gebildet. Deren Renditen werden auf der Basis von Monatsdaten mit den Renditen des Marktes abgeglichen. So können langfristige Effekte über mehrere Jahre untersucht werden. Einen Einstieg in den Calender-Time Portfolio-Ansatz bieten beispielsweise Sorescu, Shankar und Kushwaha (2007) und Diether, Malloy und Scherbina (2002).

6 Zusammenfassung und weiterführende Hinweise

In diesem Beitrag wurde die Event-Study-Methode vorgestellt. Zum Verständnis der Methode wurden zunächst mit der Darstellung der relevanten finanzwirtschaftlichen Hintergründe die Grundlagen gelegt. Für die Anwendung der Methode werden so genannte abnormale Renditen ermittelt. Hierbei handelt es sich um die prozentualen Veränderungen des Aktienkurses der betreffenden Firma. Die Schritte zur Berechnung der abnormalen Renditen und die Tests auf Signifikanz wurden detailliert dargestellt.

Die abnormalen Renditen können durch mehrere Einflüsse verzerrt werden. Zum einen können überlappende Events zu einem falschen Bild der Marktreaktionen führen und zum anderen können ungünstig gewählte Spezifikationen der Untersuchung dazu führen, dass Ergebnisse verzerrt sind oder instabil gegenüber selbst leichten Veränderungen des Designs werden. Um die Gefahren verzerrter Ergebnisse zu reduzieren, sollten mehrere Maßnahmen ausgeführt werden. Diese wurden im Einzelnen dargestellt. Von besonderer Bedeutung ist hierbei die

Frage, inwiefern sich die untersuchten Events mit anderen Kurs bewegenden Informationen überlagern. In einem Beispiel wurden die zentralen Aufgaben im Rahmen einer Event Study vorgestellt und diskutiert.

Abschließend sei gesagt, dass die Methode bislang nicht als Standard in der Statistik-Software enthalten ist. Zur Berechnung bietet sich beispielsweise Microsoft Excel an.

Literatur

- Agarwal, R. & Bayus, B.L. (2002).** The Market Evolution and Sales Takeoff of Product Innovations. *Management Science*, 48, 1024-1041.
- Agrawal, J. & Kamakura, W.A. (1995).** The Economic Worth of Celebrity Endorsers: An Event Study Analysis. *Journal of Marketing*, 59, 56-62.
- Böhmer, E. & Löffler, Y. (1999).** Kursrelevante Ereignisse bei Unternehmensübernahmen: Eine empirische Analyse des deutschen Kapitalmarktes. *Zeitschrift für betriebswirtschaftliche Forschung*, 51, 299-324.
- Brealey, R.A., Myers, S.C. & Allen, F. (2008).** *Principles of corporate finance*. 9. Auflage, Boston, Mass.
- Brown, S.J. & Warner, J.B. (1980).** Measuring security price performance. *Journal of Financial Economics*, 8, 205-258.
- Brown, S.J. & Warner, J.B. (1985).** Using daily stock returns: The case of event studies. *Journal of Financial Economics*, 14, 3-31.
- Campbell, J.Y., Lo, A.W.-C. & MacKinlay, A.C. (1997).** *The econometrics of financial markets*, Princeton, N.J.
- Chaney, P.K. und Devinney, T.M. (1992).** New Product Innovations and Stock Price Performance. *Journal of Business Finance*, 19, 677-685.
- Chaney, P.K., Devinney, T.M. & Winer, R.S. (1991).** The Impact of New Product Introductions on the Market Value of Firms. *Journal of Business*, 64, 573-610.
- Clement, M., Fischer, M. & Goerke, B. (2007).** Neuprodukteinführungen in der Filmindustrie: Wie reagieren Kapitalmarktinvestoren auf den Umsatzerfolg neuer Kinofilme? *Die Betriebswirtschaft*, 67, 418-444.
- De Bondt, W.F.M. & Thaler, R. (1985).** Does the Stock Market Overreact? *Journal of Finance*, 40, 793-805.
- Diether, K. B., Malloy, C.J. & Scherbina, A. (2002).** Differences of Opinion and the Cross-Section of Stock Returns. *Journal of Finance*, 57, 2113-2141.
- Elberse, A. & Eliashberg, J. (2003).** Demand and Supply Dynamics for Sequentially Released Products in International Markets: The Case of Motion Pictures. *Marketing Science*, 22, 329-354.
- Fama, E.F. (1965).** The Behavior of Stock-Market Prices. *Journal of Business*, 38, 34-105.
- Fama, E.F. (1970).** Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, 25, 383-417.
- Fama, E.F. (1991).** Efficient Capital Markets 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.

- Geyskens, I., Gielens, K. & Dekimpe, M.G. (2002).** The Market Valuation of Internet Channel Additions. *Journal of Marketing*, 66, 102-119.
- Koka, P.S., Jagpal, H.S. & Viswanath, P.V. (1997).** The Effect of New Product Announcements and Preannouncements on Stock Price. *Journal of Market Focused Management*, 2, 183-199.
- McWilliams, A. & Siegel, D. (1997).** Event Studies in Management Research: Theoretical and Empirical Issues. *Academy of Management Journal*, 40, 626-657.
- Sorescu, A., Shankar, V. & Kushwaha, T. (2007).** New Product Preannouncement and Shareholder Value: Don't Make Promises You Can't Keep. *Journal of Marketing Research*, 44, 468-489.
- Tellis, G.J. & Johnson, J. (2007).** The Value of Quality, *Marketing Science*, 26, 758-773.
- Worrell, D.L., Davidson III., W.N., Chandy, P.R. & Garrison, S.L. (1986).** Management Turnover Through Deaths of Key Executives: Effects on Investor Wealth. *Academy of Management Journal*, 29, 674-69.

B. Ausreißerwerte

Björn Goerke

Erschienen als:

Goerke, Björn (2009): Event Studies, in: Albers, Sönke, Daniel Klapper, Udo Konradt, Achim Walter und Joachim Wolf (Hrsg.): Methodik der empirischen Forschung (<http://www.bwl.uni-kiel.de/grad-kolleg/methodenbuch/>)

1 Einleitung

Extreme Werte in Datensätzen können die Ergebnisse substanziell beeinflussen. Das führt häufig zu einer Verzerrung der geschätzten Koeffizienten. Im Extremfall können sogar einzelne Beobachtungen dazu führen, dass signifikante Teststatistiken insignifikant werden und vice versa. Ebenso ist es möglich, dass Vorzeichen von geschätzten Koeffizienten wechseln (Mahajan, Sharma und Wind 1984; Martin und Roberts 2006; Mullen, Milne und Doney 1994). Aufgrund dieser Auswirkungen ist es angebracht, sich mit entsprechenden Werten eines Datensatzes auseinanderzusetzen.

Diese extremen Werte werden häufig als so genannte Ausreißer klassifiziert. Teilweise werden diese dann mit sehr einfachen Regeln identifiziert und eliminiert. In SPSS gibt es eine Option für die lineare Regression zur automatischen Elimination der Werte, die mehr als 3 Standardabweichungen vom Mittelwert abweichen. Dieses Vorgehen ist einfach umzusetzen und wird so auch in der Forschung berichtet (Aghion, Howitt und Mayer-Foulkes 2005). Es wird jedoch der Thematik nicht immer gerecht.

Bevor Ausreißer eliminiert werden, sollte versucht werden, festzustellen, wie es zu diesen extremen Werten gekommen ist. Sobald man die vermeintlichen Ausreißer versteht, kann die Entscheidung, wie mit ihnen umzugehen ist, besser getroffen werden. Und die Elimination ist dann eine von mehreren Möglichkeiten. Eine vorschnelle Elimination kann dazu führen, dass neue Erkenntnisse nicht aufgedeckt werden, wenn die extremen Werte beispielsweise aufgrund eines bislang nicht beobachteten Verhaltens entstanden sind. Dies würde einen Verlust für die Forschung darstellen. Ebenso ist es aber denkbar, dass fehlerhafte Beobachtungen in den Datensatz eingegangen sind, die – da nicht identifiziert und nicht korrigiert bzw. elimi-

niert – als neue Erkenntnisse ihren Weg in die wissenschaftlichen Journals finden (Hill und Rothaermel 2003).

Für die Entscheidung, wie man mit diesen extremen Werten umgeht, ist es also notwendig zu verstehen, woher sie stammen bzw. welche Ursachen sie haben und welchen Einfluss sie auf die Untersuchung und deren Ergebnisse nehmen. Dabei gibt es in diesem Thema kein eindeutig richtig oder falsch. Es bestehen weder verbindliche Richtlinien für den Umgang mit Ausreißern noch festgelegte Standardverfahren, um diese überhaupt erst zu identifizieren. Vielmehr ist es entscheidend, zu verstehen, ob die extremen Werte eher Fehler oder Ausnahmen darstellen, die man entweder korrigieren oder eliminieren sollte oder ob man sie schlichtweg nicht beachtet. Zusätzlich muss geprüft werden, ob es sich um Strukturen in den Daten handelt, die man mit der bislang gewählten Vorgehensweise noch nicht angemessen erfassen kann. Dies wäre insbesondere dann der Fall, wenn die Modellierung oder die eingesetzten Methoden die Datenstruktur nicht abbilden können (Cohen et al. 2003, S. 411-415).

Um die Entscheidung fundiert treffen zu können, werden in diesem Beitrag die gängigen Verfahren der Ausreißeranalyse in der linearen Regression anhand eines Beispiels vorgestellt und diskutiert. Mit Hilfe der Ergebnisse der Analysen kann dann die Entscheidung über das weitere Vorgehen getroffen werden. Neben der verbreiteten Elimination sind auch andere Lösungen möglich, die am Ende des Beitrags vorgestellt werden.

2 Quellen und Ursachen

Die Möglichkeiten, wie extreme Werte entstehen können, sind vielfältig (Mahajan, Sharma, Wind 1984). Eine wichtige Quelle für extreme Werte liegt in den verschiedenen Fehlermöglichkeiten empirischer Arbeit. Das beginnt bei einfachen Tippfehlern, die bei der Übertragung der Daten von Hand passieren können. Alleine die Verschiebung des Kommas um eine Stelle kann bereits zu extremen Werten führen. Ebenso entstehen Fehler, wenn beispielsweise durch fehlerhaft kalibrierte Messgeräte falsche Daten weitergeleitet werden oder beispielsweise Missverständnisse und Fehlinterpretationen entstehen (Podsakoff et al. 2003; Spector 2006). Letztendlich erfordert nahezu jeder Datensatz auch eine Aufbereitung in einem Tabellenkalkulationsprogramm. Auch hierbei können unabsichtlich Werte verfälscht werden.

Sollten Ausreißerwerte aufgrund von entsprechenden Fehlern entstehen, so sollte geprüft werden, inwiefern die Informationen noch korrigiert oder gegebenenfalls noch einmal erhoben werden können. Sollte das nicht möglich sein, weil die Daten beispielsweise anonym er-

hoben wurden, so würde hier tatsächlich eine Elimination der betroffenen Daten in Frage kommen.

Neben diesen Aspekten des gewissenhaften Umgangs mit dem Datensatz gibt es jedoch eine Reihe weiterer Aspekte, die zu extremen Werten führen können. Dies kann beispielsweise in Ausnahmen begründet sein (Weisberg 1980, S. 114 und S. 135). So nutzen Seaver und Triantis (1995) die Ausreißeridentifikation, um Unternehmen mit ungewöhnlichen Eigenschaften und außergewöhnlichen Ergebnissen aus einem Datensatz zu ermitteln und zu verstehen, was die Ursache für eine ungewöhnliche Performance in ihrer Studie ist.

Ein häufiges Problem stellen unbeobachtete Einflüsse dar, die zwar auf die abhängigen Variablen wirken, nicht jedoch im Datensatz erfasst sind. Denkbar wäre hier der Ausfall von Servern, über die Online-Shops betrieben werden. Möchte man beispielsweise feststellen, inwiefern der Versand von Newslettern sich auf die Absätze ausgewirkt hat, so können hier im Extremfall sogar negative Einflüsse geschätzt werden, sofern die Informationen über die Serverausfälle nicht in der Schätzung berücksichtigt werden können.

Schließlich können auch Modellfehlspezifikationen zu extremen Beobachtungen führen. Wenn der Zusammenhang zwischen einem Marketinginstrument und den Absätzen beispielsweise linear spezifiziert wird, allerdings abnehmende Grenzerträge vorliegen, so werden hohe Ausgaben für das Marketinginstrument im Modell zwangsläufig von den Beobachtungen abweichen. Auch so können extreme Werte entstehen, die Handlungsbedarf erzeugen.

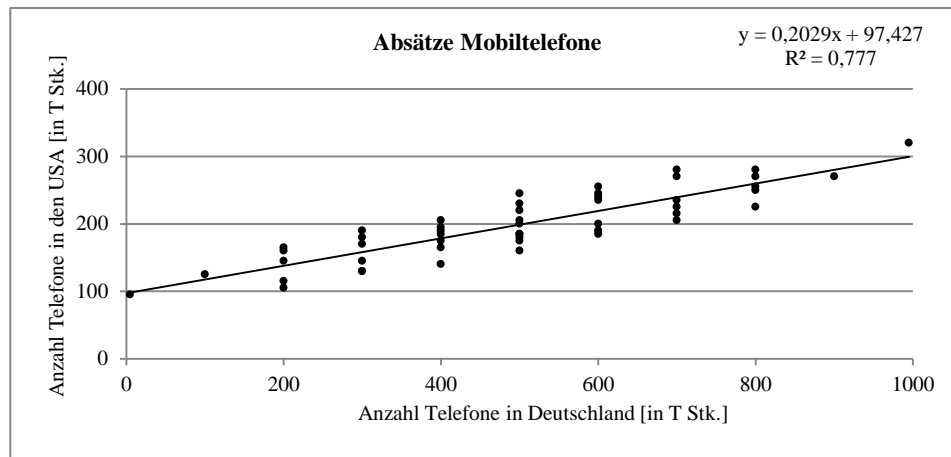
Neben dem Verständnis für die Herkunft der Ausreißer und der damit verbundenen Lösungsansätze ist es notwendig, die Auswirkungen der extremen Werte auf die Untersuchung beurteilen zu können. Hierfür werden im folgenden Kapitel die möglichen Verfahren vorgestellt.

3 Klassifizierung von Ausreißern

Ausreißerwerte können hinsichtlich ihrer Auswirkungen auf die Forschungsergebnisse in drei unterschiedliche Gruppen eingeordnet werden. Nicht alle Auswirkungen sind dabei für jedes Forschungsprojekt gleichermaßen von Bedeutung. Im Folgenden werden anhand eines Beispiels die unterschiedlichen Vorgehensweisen und die Auswirkungen vorgestellt. Grundsätzlich stehen für die Identifikation von Ausreißern viele unterschiedliche Verfahren zur Verfügung, die sich häufig nur in kleinen Details unterscheiden. Teilweise werden Verfahren mit unterschiedlichen Namen bezeichnet. Unterschiedliche Verfahren wiederum erhalten durch sehr ähnliche Namen, weshalb Vorsicht geboten ist. In diesem Beitrag werden etablierte

Verfahren vorgestellt, die in den gängigen Software-Paketen implementiert sind. Die anzuwendenden Verfahren werden im Folgenden anhand eines Beispiels diskutiert.

Abbildung 1: Beobachtungen Reaktionsfunktion ohne Ausreißer



Der Beispieldatensatz ist univariat aufgebaut. Dies ermöglicht eine einfache graphische Umsetzung, was für das Verständnis hilfreich ist. Die hier vorgestellten Verfahren stehen alle auch für den multivariaten Fall zur Verfügung. Die Erkenntnisse können dafür unverändert übernommen werden.

Das Beispiel stammt aus der Mobilfunkindustrie. Es liegen Daten eines Mobiltelefonherstellers über die Absatzzahlen von insgesamt 50 unterschiedlichen Modellen in Deutschland und in den USA vor. Diese Modelle wurden alle zunächst in Deutschland auf den Markt gebracht. Der Marktstart in den USA erfolgte jeweils mit deutlicher Verzögerung. Durchschnittlich konnte der Hersteller ungefähr 500.000 Stück pro Modell in Deutschland absetzen. In den USA ist die Marktposition des Herstellers schwächer. Dort lag der Absatz bei ungefähr 200.000 Stück pro Modell. Nicht alle Modelle konnten den gleichen Erfolg aufweisen. Einige Modelle konnten kaum verkauft werden, andere Modelle wiederum ermöglichten Absätze in Deutschland von bis zu 1.000.000 Stück (s. Abbildung 1).

Für die Vermarktung der Telefone in den USA ist die Einschätzung des zu erwartenden Markterfolgs von hoher Bedeutung, da in Abhängigkeit der Erwartungen bereits Telefone vorproduziert und Ressourcen für die unterschiedlichen Marketinginstrumente zugewiesen werden. Es soll nun anhand der vorliegenden Daten untersucht werden, ob ein Zusammenhang zwischen den Absätzen in Deutschland und in den USA besteht. Um den Zusammenhang zu untersuchen, wird ein univariates lineares Modell aufgestellt, das mit der linearen Regression geschätzt werden kann:

$$\text{Absätze USA} = \alpha + \beta \times \text{Absätze D} + u, \quad (1)$$

mit u als Fehlerterm der Schätzung und α und β als Koeffizienten des Modells.

In diesem Beispieldatensatz zeigt sich ein positiver Zusammenhang zwischen den Absätzen in den beiden Ländern. Der Koeffizient β nimmt einen Wert von 0,20 an. Das bedeutet, dass für 1.000 abgesetzte Telefone in Deutschland 200 Telefone in den USA abgesetzt werden. Da der Koeffizient α einen Wert von 97,43 annimmt, besitzt der Mobilfunkhersteller in den USA offenbar eine Kundenbasis, die anders als die Kunden in Deutschland auch dann die Telefone abnimmt, wenn ein Modell kein Markterfolg in Deutschland war. Andererseits bringen die Kunden in den USA offenbar nicht die gleiche Begeisterung für die in Deutschland erfolgreichen Telefone auf, was sich an dem niedrigen Steigungskoeffizienten zeigt. Das R^2 ist mit 0,78 ansprechend.

3.1 Leverage

Wenn einzelne Realisationen der unabhängigen Variablen eines Datensatzes extreme Werte annehmen, spricht man vom so genannten Leverage (Gao, Li und Wang 2005). In dem Beispiel des Mobiltelefonherstellers ist es denkbar, dass es um einzelne Modelle in Deutschland einen Hype gegeben hat, weil Testimonials von Stars besonders treffsicher waren. Sofern diese Stars nur in Deutschland diese Vorbildfunktion haben, wird sich der Effekt in den USA nicht wiederholen. Während also in Deutschland außergewöhnlich hohe Verkaufszahlen erzielt werden, können in den USA nur durchschnittliche Absätze realisiert werden. Damit ergibt sich im Datensatz eine Situation, dass Beobachtungen enthalten sind, die zwar hohe Werte bei den unabhängigen Variablen aufweisen, den Absätzen in Deutschland, aber nur durchschnittliche Werte bei den abhängigen Variable, den Absätzen in den USA.

3.1.1 Auswirkungen

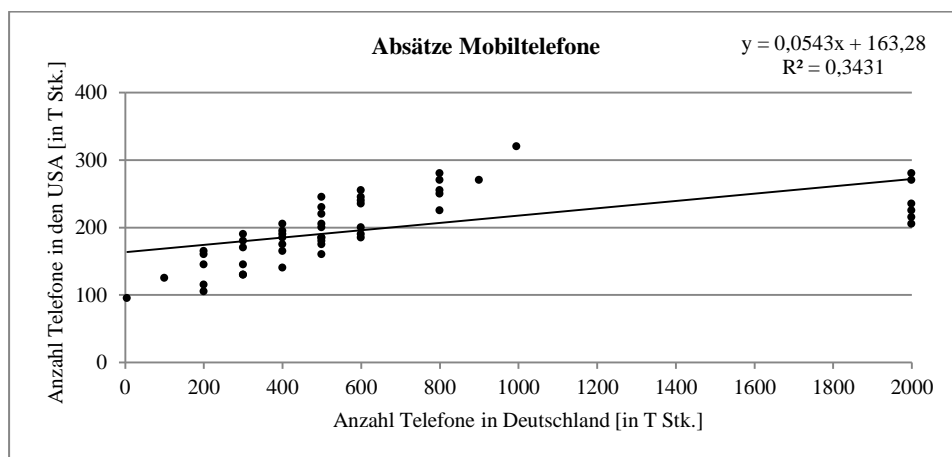
Es stellt sich die Frage, welche Auswirkungen diese extremen Werte auf die Schätzung haben. Abbildung 2 verdeutlicht die Situation. Insgesamt liegen in dem Datensatz sechs Modelle vor, bei denen in Deutschland ein Hype realisiert werden konnte. Das Unternehmen konnte jeweils ungefähr 2.000.000 Stück in Deutschland absetzen und damit ungefähr doppelt so viel wie mit dem erfolgreichsten Modell, das ohne Hype vermarktet wurde. Ohne den Hype-Effekt hätte das Unternehmen jeweils ungefähr 700.000 Stück von den Modellen absetzen können.

Betrachtet man die Ergebnisse der Schätzung mit den Leverage-Beobachtungen, so ergibt sich ein anderes Bild als im Fall ohne Ausreißer. Die Gerade durch die Punktwolke nimmt zwar weiterhin einen positiven Verlauf, aber die Steigung ist flacher. Der Einfluss der extremen Werte wird auch durch das β von 0,05 reflektiert, was einen schwächeren Zusammen-

hang der Absätze in Deutschland mit den Absätzen in den USA impliziert. Andererseits deutet der größere Achsenabschnitt darauf hin, dass die Absätze in den USA, die ohne Verbindung zu den Absätzen in Deutschland stehen, höher sind und bei ungefähr 163.000 Stück liegen. Das R^2 von 0,34 ist zwar deutlich geringer als das R^2 der Schätzung ohne Ausreißer, jedoch ist es immer noch befriedigend.

Da die Ergebnisse trotz der Ausreißer immer noch als plausibel eingestuft werden können, besteht das Risiko, dass man die Werte für die Handlungsimplicationen im Unternehmen nutzt, um zu prognostizieren, wie die Absätze in den USA in Abhängigkeit von den Absätzen in Deutschland laufen werden. Da der Zusammenhang so allerdings deutlich unterschätzt wird und der amerikanische Markt als relativ unabhängig vom deutschen Markt eingeschätzt werden würde, käme es zu Fehleinschätzungen. So würden eventuell Telefone, die in Deutschland sehr schlecht vermarktet werden konnten, in den USA trotzdem in den Markt gebracht, auch wenn es vorteilhaft wäre, diese Produkte dort überhaupt nicht einzuführen. Gleichzeitig würde man die Nachfrage nach erfolgreichen Mobiltelefonen unterschätzen und mit zu geringen Ressourcen in den Markt gehen, was wiederum dazu führt, dass das Potenzial nicht voll-

Abbildung 2: Beobachtungen Reaktionsfunktion mit Leverage-Werten



ständig abgeschöpft werden kann. Es besteht also das Risiko, die Wettbewerbsfähigkeit durch ein Übersehen dieser Ausreißer zu schwächen.

3.1.2 Identifikation

Die graphische Umsetzung der Situation mit Leverage-Werten in Abbildung 2 bietet einen ersten Hinweis, dass extreme Werte einen unerwünschten Einfluss auf die Ergebnisse nehmen. Neben dem Plot der Punktwolke stehen Statistiken zur Verfügung, deren Einsatz konkrete Hinweise über die Stärke des Einflusses einzelner Beobachtungen liefern. Um einzu-

schätzen, ob einzelne Werte der unabhängigen Variablen außergewöhnlich extrem sind, bietet sich die Berechnung der so genannten Leverage-Werte h_{ii} an (Cohen et al. 2003, S. 394):

$$h_{ii} = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum_{i=1}^n x_i^2} \quad (2)$$

wobei h_{ii} den Leverage für Beobachtung i , n die Anzahl der Beobachtungen, x_i den Wert für Beobachtung i und \bar{x} den Mittelwert der Beobachtungen darstellen. Der durchschnittliche Wert für h_{ii} ergibt sich aus $\overline{h_{ii}} = (k+1)/n$, wobei k die Anzahl der unabhängigen Variablen angibt. Diese Statistik sollte für jede einzelne Beobachtung ermittelt werden. In den Software-Paketen SAS und SYSTAT ist diese Statistik implementiert. Sie kann dort für jede einzelne Beobachtung automatisiert in einem Durchlauf berechnet werden.

SPSS hingegen verwendet den so genannten zentrierten Leverage h_{ii}^* , der vereinfacht ermittelt wird. Durch die Transformation

$$h_{ii}^* = \frac{(x_i - \bar{x})^2}{\sum_{i=1}^n x_i^2} \quad (3)$$

ergeben sich leicht abweichende absolute Werte. Eine andere Qualität der Information bietet diese Statistik jedoch nicht. Auch in SPSS kann man sich diese Statistik automatisiert für alle Beobachtungen ausgeben lassen. Eine weitere Statistik, die beispielsweise in SPSS angeboten wird, ist die so genannte Mahalanobis-Distanz, die wiederum als

$$Mahab = (n-1)h_{ii}^* \quad (4)$$

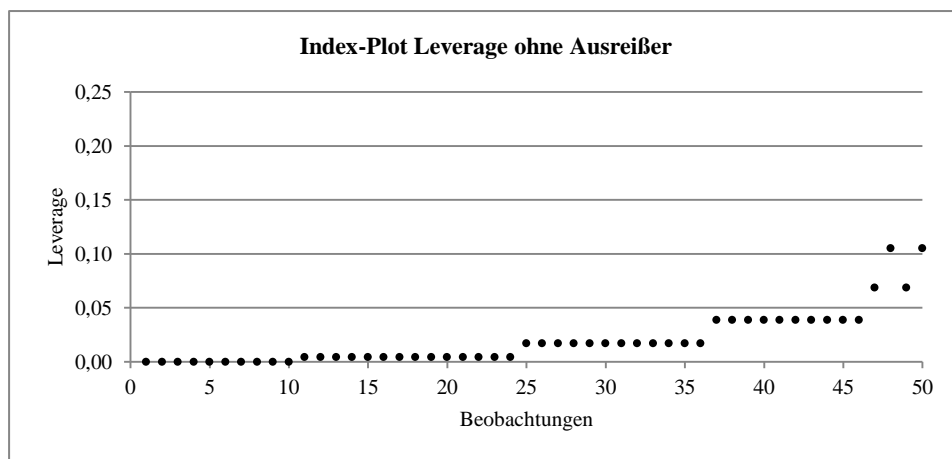
ermittelt werden kann. Wiederum ergibt sich keine zusätzliche Information.

Da die Begriffe jedoch teilweise ähnlich sind, die unterschiedlichen Berechnungen aber zu unterschiedlichen absoluten Werten führen, ist es für den Umgang mit den Statistiken entscheidend, jeweils die richtige Statistik des Software-Pakets auszuwählen.

Die Frage, ab welchem Leverage-Wert man davon ausgehen muss, dass eine einzelne Beobachtung zu großen und eben unerwünschten Einfluss nimmt, kann nicht eindeutig beantwortet werden. Für normalverteilte unabhängige Variablen geben Hoaglin und Welsch (1978) als Daumenregel $2(k+1)/n$ für große Datensätze und $3(k+1)/n$ für kleine Datensätze an. Sie lassen dabei offen, welche Anzahl von Beobachtungen als größer und welche Anzahl als kleiner Datensatz eingestuft werden sollte. Der größere Wert für größere Datensätze wird auch

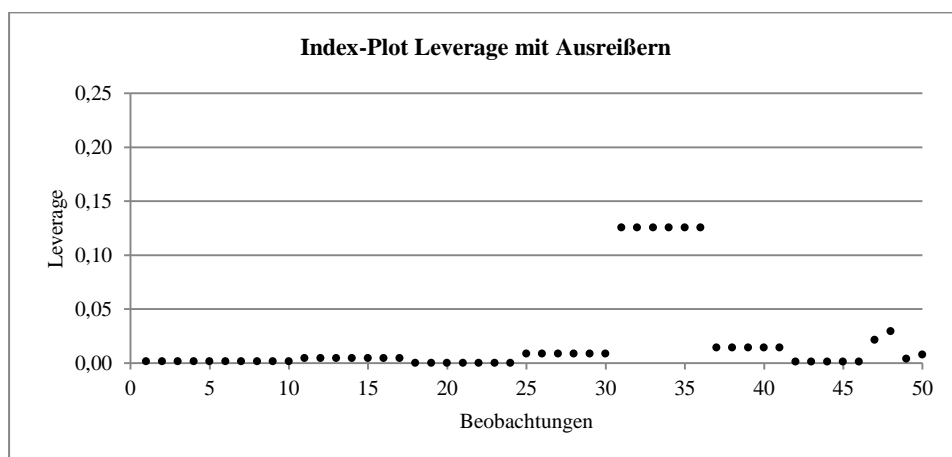
deswegen empfohlen, weil so weniger Werte überhaupt untersucht werden müssen. Diese Schwellen sind nicht als harte Kriterien zu verstehen, sondern als Hinweis, dass es sich um Ausreißer handeln könnte. Insbesondere sind jene Fälle von Bedeutung, deren Leverage-Wert eine Lücke zu den restlichen Leverage-Werten aufweisen. Die vorgeschlagenen Grenzen führen in dem Beispiel zu Werten von 0,08 bzw. 0,12 für h_{ii} . Die extremen Werte im Beispieldatensatz erreichen einen Wert von $h_{ii} = 0,13$ und liegen damit oberhalb der Grenzen.

Abbildung 3: Index-Plot der Leverage-Werte ohne Ausreißer



Für die Einschätzung einer entsprechenden Lücke bietet sich die Erstellung eines Index-Plots an (Cohen et al. 2003, S. 395; Gao, Li und Wang 2005). In so einem Index-Plot werden für

Abbildung 4: Index-Plot der Leverage-Werte mit Ausreißern

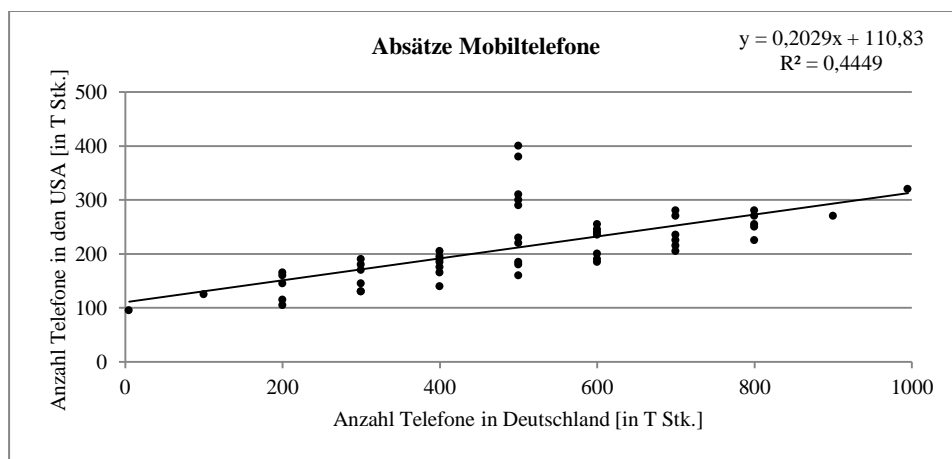


jeden einzelnen Fall die Leverage-Werte abgetragen. Die Abbildungen 3 und 4 veranschaulichen dieses Vorgehen. In Abbildung 3 sind die Leverage-Werte für jede einzelne Beobachtung ohne die Ausreißer abgetragen. Das Bild ist relativ homogen. Es entstehen keine großen Abstände zwischen den einzelnen Werten.

3.2 Discrepancy

Neben einer extremen Beobachtung der unabhängigen Variablen müssen auch extreme Werte in Bezug auf die beobachteten Realisationen der abhängigen Variablen berücksichtigt werden. Das bedeutet, dass extreme Werte für den Schätzfehler $e_i = Y_i - \hat{Y}_i$ genauerer Untersuchung bedürfen (Creech, Corzine und Huff-Corzine 1989; Sutradhar, Chu und Bari 2007). Sie sind ein Anzeichen für Ausreißer, die zur so genannten Discrepancy führen. Y_i und \hat{Y}_i sind die beobachteten bzw. die geschätzten Realisationen der Beobachtungen i der abhängigen Variablen.

Abbildung 5: Beobachtungen Reaktionsfunktion mit Discrepancy-Werten



Für das Entstehen entsprechender Werte im Beispiel des Mobiltelefonherstellers ist eine Situation denkbar, in der ein Wettbewerber in den USA Lieferschwierigkeiten mit Konkurrenzmodellen hat. Aufgrund fehlender Alternativen kaufen dann mehr Kunden die Telefone des betrachteten Herstellers. Dieser Umstand ist in dem Modell nicht berücksichtigt. In den Daten wiederum drückt sich dies durch ungewöhnlich erfolgreiche Absätze in den USA bei durchschnittlichen Absätzen in Deutschland aus.

3.2.1 Auswirkungen

Wiederum muss die Frage geklärt werden, welche Auswirkungen diese extremen Werte auf die Schätzung und auf die Handlungsimplicationen haben. Abbildung 5 verdeutlicht die Situation. Für insgesamt fünf Modelle werden außergewöhnlich viele Absätze registriert, obwohl in Deutschland noch nichts auf einen großen Verkaufserfolg in den USA hingewiesen hat.

In diesem Beispiel liegen die fraglichen Werte im Mittelpunkt der unabhängigen Variablen. Dadurch kommt es lediglich zu einer Verschiebung der Geraden, nicht zu einer gleichzeitigen Drehung. Entsprechend verhalten sich auch die geschätzten Koeffizienten. Der Steigungskoeffizient β bleibt unverändert. Der Achsenabschnitt α jedoch wird durch die extremen Werte auf knapp 111 vergrößert. Das R^2 bleibt mit 0,44 in einer befriedigenden Größenordnung. Der

Blick auf diese Parameter allein liefert wiederum keinen Hinweis auf den Einfluss der Ausreißer. Abhängig von der vorgesehenen Verwendung für die geschätzten Koeffizienten ergibt sich unterschiedlicher Handlungsbedarf. Interessiert nur der Zusammenhang zwischen den Absätzen in den beiden Märkten, so erhält man trotz dieser Ausreißer unverzerrte Ergebnisse. Eine weitere Beschäftigung mit diesen Werten ist dann nicht notwendig. Möchte man jedoch beispielsweise die Absatzzahlen für das nächste Quartal planen und daraufhin Geräte produzieren, dann ist es notwendig, diese Ausreißer zu untersuchen, da es so zu einer Unterschätzung des Bedarfs kommen würde.

3.2.2 Identifikation

SPSS bietet fünf unterschiedliche Methoden zur Identifikation von Discrepancy an. Besonders geeignet sind die so genannten Externally Studentized Residuals.¹ Dieses Verfahren wird auch von allen anderen gängigen Software-Paketen angeboten.

Bei den Externally Studentized Residuals handelt es sich um eine so genannte Eliminationsstatistik. Für jede einzelne Beobachtung wird die Regression erneut gerechnet – jeweils genau ohne die gerade betrachtete Beobachtung. Es wird also untersucht, was wäre, wenn die Beobachtung nicht Teil des Datensatzes wäre (Cook und Weisberg 1982, S. 20; Martin und Roberts 2006). Die Differenz in Bezug auf die abhängige Variable wird als

$$d_i = Y_i - \hat{Y}_{i(i)} \quad (5)$$

ausgedrückt, wobei Y_i für den beobachteten Wert steht und $\hat{Y}_{i(i)}$ für den geschätzten Wert nach der Elimination. Durch die Standardisierung dieser Differenz erhält man das Externally Studentized Residual für die Beobachtung i . Dabei gilt es zu beachten, dass für die Standardisierung die Standardabweichung des gelöschten Residuums für die Beobachtung i zu wählen ist:

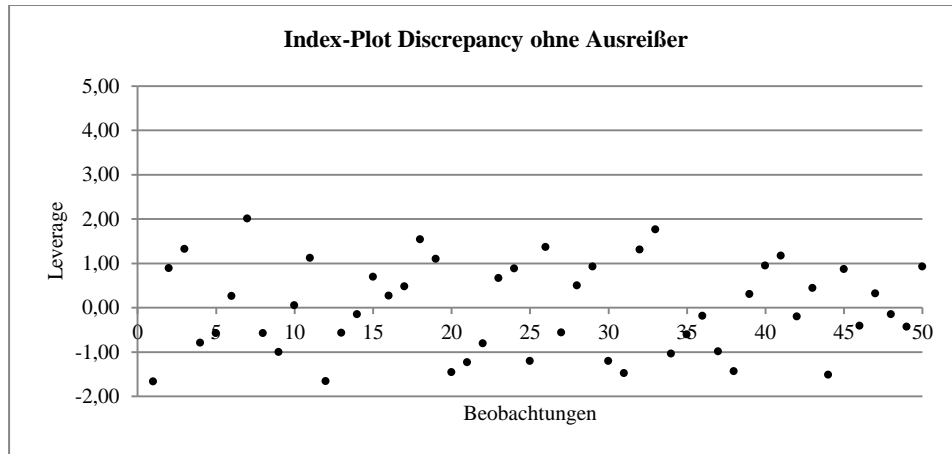
$$t_i = \frac{d_i}{\sigma_{d_i}} \quad (6)$$

Die Externally Studentized Residuals t_i sind dadurch t-verteilt, so dass zur Beurteilung der fraglichen Beobachtungen die kritischen Werte der aus den t-Tabellen genutzt werden können. Vereinfachend bietet sich jedoch die Möglichkeit an, Daumenregeln anzuwenden. Sobald eine Beobachtung die Grenze der Daumenregel überschreitet, wird der Wert als Ausreißer

¹ In der deutschen Version von SPSS findet man das Verfahren in der linearen Regression bei „Speichern ...“ und dann unter „Studentisiert, ausgeschlossen“. Die Ergebnisse werden verdichtet im SPSS-Output dargestellt und in dem Datensatz individuell für jede einzelne Beobachtung abgespeichert.

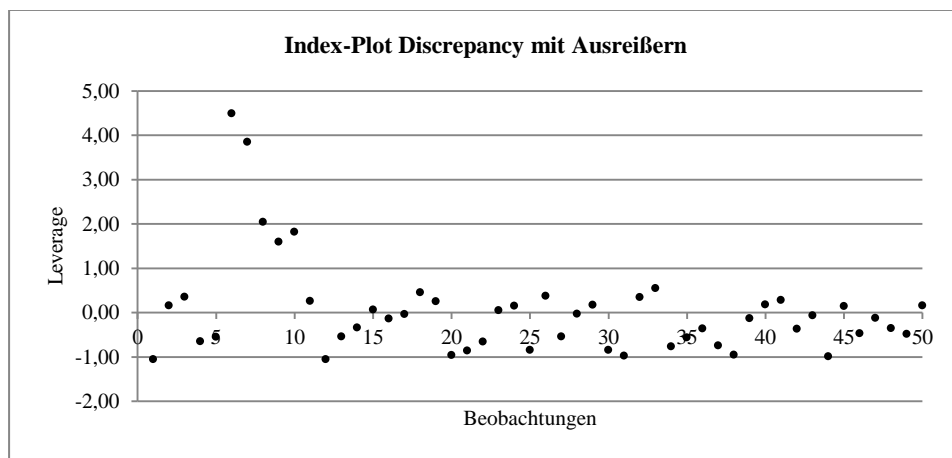
klassifiziert. Als Grenze werden Werte zwischen $\pm 2,00$ und $\pm 4,00$ für t_i in Betracht gezogen (Cohen et al. 2003, S. 401).

Abbildung 6: Index-Plot Externally Studentized Residuals ohne Ausreißer



Ähnlich wie bei der Einstufung der Leverage-Werte gibt es also keine verbindlichen harten Kriterien. Vielmehr sollen diese Statistiken einen Hinweis geben, welche Beobachtungen eine

Abbildung 7: Index-Plot Externally Studentized Residuals mit Ausreißern



potenzielle Gefahr darstellen. Für die fünf relevanten Beobachtungen in dem Beispiel ergeben sich für die t_i Werte zwischen 1,60 und 4,50. Je nach Auslegung der Kriterien fallen einige dieser Werte also in die Kategorie der Ausreißer oder nicht.

Neben der Berechnung der Statistiken bieten sich auch im Fall der Discrepancy-Ausreißer Index-Plots zur Beurteilung der Situation an (Langford und Lewis 1998). Abbildung 6 zeigt den Plot der Externally Studentized Residuals für den Datensatz ohne Ausreißer. Die einzelnen Punkte liegen alle nah beieinander. Abbildung 7 zeichnet ein anderes Bild. Es bilden sich zwei Gruppen außerhalb der allgemeinen Punktwolke. Insbesondere fällt die Lücke zu den beiden obersten Punkten auf. Auch die drei Punkte um

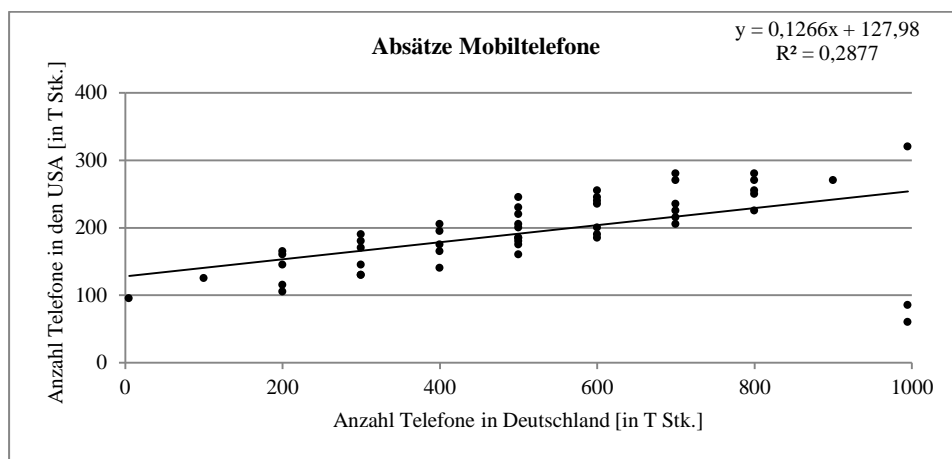
den Wert 2,00 sind auffällig. Eine weitere Beschäftigung mit diesen Beobachtungen scheint empfehlenswert.

Bei der Verwendung der unterschiedlichen Software-Pakete und der Durchsicht entsprechender Literatur ist zu beachten, dass es keine eindeutige Bezeichnung für die Externally Studentized Residuals gibt. SPSS, SAS und SYSTAT verwenden jeweils unterschiedliche Bezeichnungen. Insofern sollte im Zweifelsfall immer zunächst ein Blick auf die dahinter stehenden Formeln geworfen werden, um sicherzustellen, dass wirklich Externally Studentized Residuals verwendet wurden.

3.3 Influence

Als dritte Klassifizierung für Ausreißereigenschaften soll die so genannte Influence betrachtet werden. Mit dieser Klassifizierung wird eine Aussage darüber getroffen, inwiefern sowohl

Abbildung 8: Beobachtungen Reaktionsfunktion mit Influence-Werten



Leverage als auch Discrepancy vorliegen (Imon 2005). Bei der Betrachtung der Influence gilt es, zwei verschiedene Aspekte zu unterscheiden. Zum einen spricht man von globaler Influence, wenn die gesamte Schätzung beeinträchtigt wird. Zum anderen kann diese Aussage auch für jeden einzelnen zu schätzenden Koeffizienten getroffen werden. Dann liegt die so genannte spezifische Influence vor.

Betrachten wir in unserem Beispiel in Abbildung 8 folgende Möglichkeit, wie Influence entsteht. Der Mobiltelefonhersteller hat zwei in Deutschland sehr gut gelaufene Telefone wie gewohnt mit Verzögerung auf den US-Markt gebracht. In den USA schneiden diese Telefone jedoch nur unterdurchschnittlich lag. Das lag an einer Besonderheit des amerikanischen Marktes, wo zwischenzeitlich die Funktion Push-to-talk große Beliebtheit erfahren hat. Telefone ohne diese Funktion werden von einigen Kundensegmenten nicht mehr gekauft. Dieser Umstand ist jedoch im vorliegenden Datensatz nicht berücksichtigt worden.

3.3.1 Auswirkungen

Diese ungewöhnliche Konstellation bei den beiden Modellen wirkt sich auf die Schätzung aus. Die Steigung fällt deutlich. Das β liegt nur noch bei 12,66. Gleichzeitig kommt es zu einer Vergrößerung des Achsenabschnittes α auf 127,98. Für die Interpretation bedeutet dies, dass der Zusammenhang der Erfolge auf den beiden Märkten unterschätzt wird. Die Koeffizienten sind jedoch durchaus plausibel und das R^2 von 0,29 ist immer noch ausreichend, so dass hier die Gefahr besteht, dass die Ergebnisse trotz der Influence in den Daten verwendet und so Fehlentscheidungen getroffen werden.

3.3.2 Identifikation

Auch für die Identifikation der Influence stehen Verfahren zur Verfügung, die in den Software-Paketen implementiert sind. Dies gilt sowohl für die globale als auch für die spezifische Influence. Zur Identifikation der globalen Influence bietet sich das weit verbreitete Maß Cook's D an (Cook und Weisberg 1982, S. 10 ff.). Wie bei den Externally Studentized Residuals handelt es sich hier um eine Eliminationsstatistik. Das bedeutet, dass für jede einzelne Beobachtung ermittelt wird, was wäre, wenn diese Beobachtung nicht Teil des Datensatzes wäre. Die Statistik hat folgende Form:

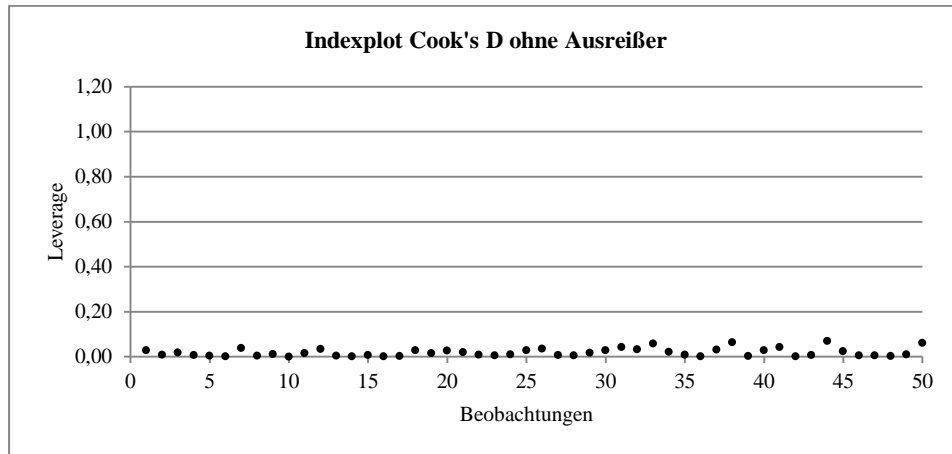
$$Cook's D_i = \frac{\sum_{i=1}^n (\hat{Y} - \hat{Y}_{(i)})^2}{(k+1)MS_{residual(i)}} \quad (7)$$

\hat{Y}_i steht dabei für den geschätzten Y-Wert für Beobachtung i für den Fall, dass die Schätzung mit dem vollständigen Datensatz vorgenommen wurde. $\hat{Y}_{(i)}$ steht für den geschätzten Y-Wert für Beobachtung i für den Fall, dass die Koeffizienten ohne Beobachtung i ermittelt wurden. n ist die Anzahl an Beobachtungen und k die Anzahl an Regressoren im Modell. Der Nenner standardisiert die Differenz im Zähler des Ausdrucks. $MS_{residual(i)}$ ist das durchschnittliche quadrierte Residuum unter Auslassung von Fall i. Für das Beispiel ergeben sich für die beiden fraglichen Beobachtungen Werte für Cook's D von 0,77 und 1,01. Pauschal werden Cook's D-Werte, die größer als 1,00 sind, als Ausreißer mit globaler Influence eingestuft. Genauer ist jedoch die Verwendung der kritischen Werte der F-Verteilung (mit k+1 und n-k-1 Freiheitsgraden) bei einem Konfidenzniveau von $\alpha = 0,5$ (Cohen et al. 2003, S. 404). Im Beispiel liegt der kritische Wert bei 0,43. Damit liegen die beiden identifizierten Werte außerhalb der kritischen Werte für Cook's D.

Ein weiteres bekanntes Maß zur Bestimmung der Influence ist DFFITS, das 1980 von Belsley, Kuh und Welsch eingeführt wurde (Langford und Lewis 1998). Es lässt sich in Cook's D überführen (Imon 2005). Insofern bietet es keine zusätzlichen Informationen zu

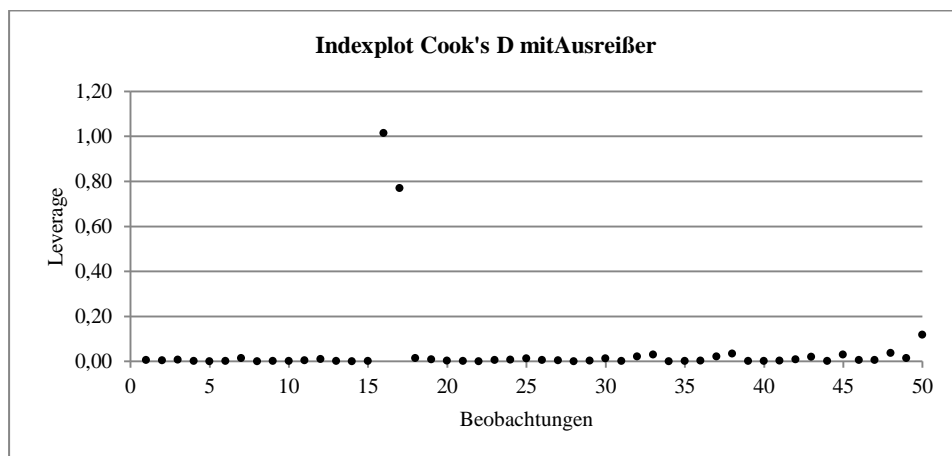
dem vorgestellten Maß. Beide Verfahren sind in den gängigen Software-Paketen implementiert.

Abbildung 9: Index-Plot Cook's D ohne Ausreißer



Zur Identifikation von Ausreißerwerten mit einem hohen Maß an Einfluss bietet sich zusätzlich wieder die Überprüfung eines Index-Plots an. Das Vorgehen entspricht dabei dem bereits beschriebenen Verfahren. Abbildung 9 zeigt den Index-Plot der 50 Cook's D-Werte für den Datensatz ohne die Ausreißer. Das Bild ist hier weitestgehend homogen.

Abbildung 10: Index-Plots Cook's D mit Ausreißern



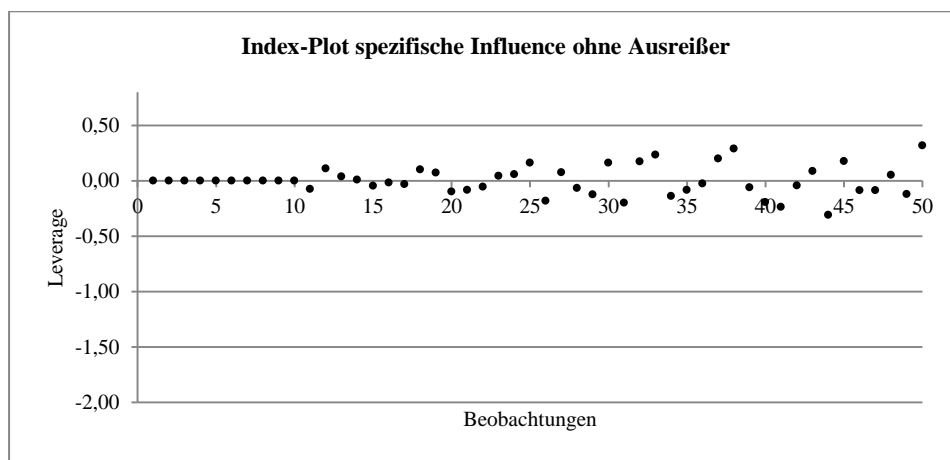
Einzelne Punkte weichen etwas ab, aber es entsteht dabei keine eindeutige Lücke zwischen diesen Punkten und der restlichen Punktwolke. Abbildung 10 wiederum zeigt die Situation für die Cook's D-Werte mit den Ausreißern. Deutlich zeigt sich, dass die beiden fraglichen Werte von der Punktwolke abweichen. Dies ist ein Hinweis dafür, diese Punkte genauer zu untersuchen.

Für die Feststellung, ob für die extremen Werte zusätzlich auch eine spezifische Influence relevant ist, stehen erneut Eliminationsstatistiken bereit (Chien und Tsou 2007). Die Statistik wird als DFBETAS bezeichnet und wie folgt berechnet:

$$DFBETAS_{ij} = \frac{\beta_j - \beta_{j(i)}}{\sigma_{\beta_{j(i)}}} \quad (8)$$

Der Nenner dient wiederum der Standardisierung. Hierzu wird die Standardabweichung des Schätzers für β_j nach der Elimination von Fall i verwendet. Die in SPSS implementierte Vor-

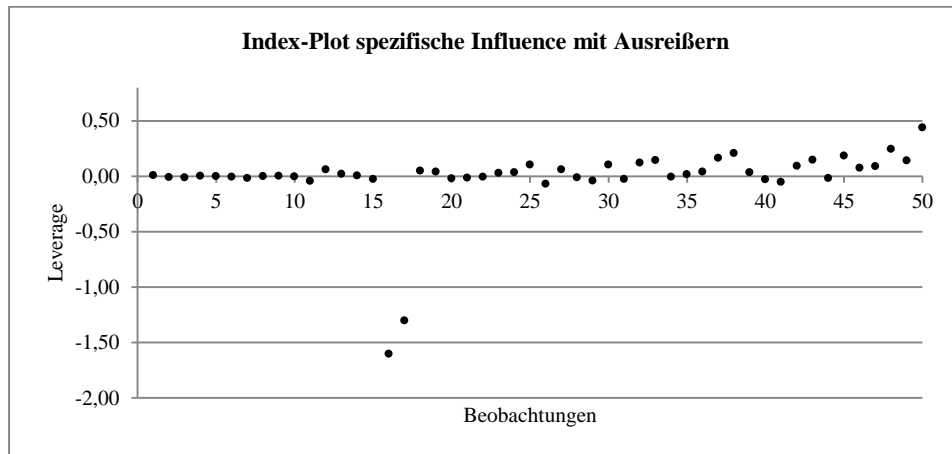
Abbildung 11: Index-Plot DFBETAS ohne Ausreißer



gehensweise sieht vor, dass diese Berechnung für jeden einzelnen Koeffizienten j und für jede einzelne Beobachtung i vorgenommen wird, so dass ein Überblick über jeden einzelnen Einfluss auf die Koeffizienten gegeben wird. Diese Berechnung erfolgt auch für den Achsenabschnitt. Im Beispiel ergeben sich für die beiden verdächtigen Punkte für den Achsenabschnitt von $DFBETAS_{16(0)} = 1,14$ und $DFBETAS_{17(0)} = 0,93$ und von $DFBETAS_{16(1)} = -1,62$ und $DFBETAS_{17(1)} = -1,31$ für den Steigungskoeffizienten. Auch für DFBETAS gilt die Daumenregel, dass Werte, die betragsmäßig größer als 1,00 liegen, als Ausreißer eingestuft werden können (Cohen et al. 2003, S. 405). Für den Achsenabschnitt ergibt sich hier somit ein Grenzfalle. Ist dieser von Bedeutung für die weiteren Handlungsimplicationen, so sollte überprüft werden, ob weitere Informationen darüber eingeholt werden können, weshalb die Werte so extrem sind. Für die beiden DFBETAS der Steigungskoeffizienten gilt dies ebenso, da beide Werte deutlich unter -1,00 liegen. Zur Identifikation von Ausreißern bietet sich wieder die Erstellung von Index-Plots an. Wiederum gilt, dass eine Lücke zwischen dem Wert für einen einzelnen Fall und den restlichen Werten auf einen Ausreißer hinweist. Abbildung 11 zeigt ein homogenes Bild der 50 Beobachtungen für die DFBETAS des Steigungskoeffizienten. Abbildung 12 wiederum weist eine deutliche Lücke zwischen den DFBETAS bei den ver-

dächtigen Punkte und der restlichen Punktwolke auf. Die restlichen Punkte liegen alle in einem engen Intervall. Die Situation der DFBETAS für den Achsenabschnitt ist vergleichbar.

Abbildung 12: Index-Plot DFBETAS mit Ausreißern



4 Behandlung von Ausreißern

Um sicherzustellen, dass die extremen Werte in einem Datensatz auch identifiziert werden, sollten die hier vorgestellten Verfahren standardmäßig Anwendung finden. Die entsprechenden Statistiken in SPSS oder ähnlichen Software-Paketen ausführen zu lassen, stellt einen geringen Aufwand dar. Sollten Grenzfälle auftreten, so sollte auch die Erstellung von Index-Plots in Erwägung gezogen werden, da diese durch die Visualisierung zusätzliche Sicherheit geben können.

Sobald extreme Werte als Ausreißer klassifiziert werden, muss festgestellt werden, wie es zu diesen extremen Werten gekommen ist. Insbesondere sollten Fehler in den Daten korrigiert werden, sofern dies möglich ist. Sofern Nacherhebungen notwendig wären, stößt man hier unter Umständen an die Grenze, wenn beispielsweise Anonymität zugesichert wurde und die Befragten aufgrund fehlender Kontaktdaten nicht wieder angesprochen werden können. Dann kann alternativ eine Elimination sinnvoll sein. Dies könnte auch im Fall des Mobiltelefonherstellers bei den Lieferschwierigkeiten des Wettbewerbers im Beispiel dieses Beitrags so sein. Wenn die Marketingabteilung die Zusammenhänge auswerten möchte, wird in der Datenbank nicht die Verfügbarkeit der Konkurrenzmodelle dokumentiert sein. Insofern ist es denkbar, dass die Verknüpfung dieser wichtigen Information für diese Modelle nicht erfolgen kann. Da es sich nur um wenige Modelle handelt, kann es sinnvoll sein, diese Beobachtungen zu vernachlässigen und ohne sie zu arbeiten. Da eine Nicht-Verfügbarkeit von Konkurrenzmodellen selten ist und von dem Mobiltelefonhersteller nicht eingeplant werden kann, wäre eine Model-

lierung ohne diese sinnvoll. Das Modell wäre für die Zwecke des Unternehmens trotzdem gut geeignet.

Grundsätzlich sollten vor einer Elimination der Ausreißer allerdings andere Optionen geprüft werden (Mahajan, Sharma und Wind 1984). Dies gilt insbesondere dann, wenn gleich mehrere extreme Beobachtungen auftreten und nicht nur einzelne Werte (Hill und Rothaermel 2003; Seaver und Triantis 1995). Hier besteht die Möglichkeit, dass es sich nicht um Ausreißer im klassischen Sinne handelt, die auf Fehlern oder Ausnahmen ohne Bedeutung für die eigentliche Fragestellung beruhen. Diverse andere Möglichkeiten könnten dann zutreffen. So ist denkbar, dass die gewählte Modellierung die Realität nicht mit ausreichender Güte abbildet, andere Methoden besser geeignet sind oder Transformationen der Daten notwendig werden. Im Folgenden werden die zusätzlichen Möglichkeiten vorgestellt.

4.1 Alternative Modellierung

Die Modellierung des Zusammenhangs zwischen den Verkäufen in Deutschland und denen in den USA ist bislang in diesem Beitrag linear vorgenommen worden. Die Annahme eines linearen Zusammenhangs ist in diesem Sachverhalt gerechtfertigt.

Viele Situationen erfordern jedoch eine andere Modellierung. Insbesondere wenn es um die Wirksamkeit von Maßnahmen geht, bei der Untersuchung von Marketinginstrumenten beispielsweise, sind lineare Zusammenhänge unrealistisch. Hier ist von abnehmenden Grenzerträgen auszugehen. Werden diese jedoch linear modelliert, so ist eine schlechte Modellanpassung zu erwarten. Dieses Problem kann beispielsweise durch eine Logarithmierung umgangen werden (Leeflang et al. 2000, S. 66-79). Durch die Stauchung der Datenpunkte verbessert sich die Modellanpassung dann wieder. Dadurch können insbesondere höhere oder sehr niedrige Werte der unabhängigen Variablen gut durch das Modell wiedergegeben werden. Ansonsten würden diese Punkte gegebenenfalls fälschlicherweise als Ausreißer klassifiziert werden. Das wiederum würde, wie in diesem Beitrag vorgestellt, zu einem Ausschluss der Beobachtungen führen oder zu Alternativerklärungen, die im schlimmsten Fall zu kontraproduktiven Implikationen führen könnten.

Eine weitere Modellierungsvariante besteht im Quadrieren einzelner Variablen. Sofern die Variable jeweils einmal einfach und einmal quadriert in die Schätzung eingeht, lassen sich umgekehrt u-förmige Zusammenhänge modellieren (Leeflang et al. 2000, S. 66-79). Dadurch wird es möglich, nicht nur abnehmende, sondern auch negative Grenzerträge abzubilden. Dies ist beispielsweise relevant, wenn so viele Werbeauslieferungen erfolgen, dass die Konsumenten-

ten nicht nur nicht mehr kaufen, sondern aus einer Ablehnung der Werbung heraus sogar auf Käufe verzichten. Sofern so ein Zusammenhang linear modelliert wird, ist mit Ausreißerwerten zu rechnen.

4.2 Alternative Verfahren

Neben alternativen Modellierungen sind auch andere Verfahren denkbar, um mit extremen Werten angemessen umzugehen. Eine etablierte Alternative für den Umgang mit Ausreißern ist die Least Absolute Deviations (LAD). Im Gegensatz zur linearen Regression, bei der der Fehlerterm quadriert wird, gehen bei der LAD die absoluten Werte des Fehlers in die Rechnung ein (Cohen et al. 2003, S. 417-419). Während die Methode so an Robustheit gegenüber der linearen Regression gewinnt, muss man bei der Anwendung der LAD mit instabilen bzw. multiplen Lösungen rechnen.

Ein Ansatz, der sowohl eine alternative Modellierung als auch eine Methode darstellt, ist die Spline Regression. Die Spline Regression ist dafür geeignet, eine Änderung des Steigungskoeffizienten einer Regressionsgeraden für definierte Grenzwerte explizit zu modellieren (Greene 2008). Damit können extreme Verhaltensweisen abgebildet werden. So modellieren van Doorn, Verhoef und Bijmolt (2007) den Zusammenhang zwischen persönlichen Einstellungen zu Bio-Produkten und dem tatsächlichen Kauf dieser Produkte mit Hilfe der Spline Regression, da sie die Hypothese aufstellen, dass der Konsum nicht linear mit der Affinität zu Bio-Produkten ansteigt, sondern erst ab einem Schwellenwert. An diesem Schwellenwert (spline knot) wird entsprechend die Änderung des Steigungskoeffizienten modelliert. Da der oder die Schwellenwerte nicht unbedingt vor der Untersuchung bekannt sind, wenden die Autoren ein iteratives Verfahren, basierend auf der Stepwise Regression, an, um die beste Anpassung der spline knots zu erreichen.

Eine weitere Alternative besteht in der Verwendung von Verfahren, die eine Segmentierung der Daten vornehmen (Wedel und Kamakura 2001). Anders als in der linearen Regression können so die Verhaltensweisen von unterschiedlichen Käuferschichten abgebildet werden, die auf Marketinginstrumente unterschiedlich reagieren. Gibt es beispielsweise in einem Datensatz ein großes Segment von durchschnittlichen Kunden, die auf Werbemaßnahmen eine homogene Reaktion zeigen, und gleichzeitig aber auch ein kleines Segment von Enthusiasten, die auf Werbung sehr viel stärker reagieren, so könnte man mit der linearen Regression Gefahr laufen, dieses kleine, sehr wertvolle Segment als Ausreißer einzustufen. Durch die Segmentierung wird es möglich, beide Segmente zu trennen und die Effekte jeweils segmentspezifisch zu ermitteln, so dass beide Segmente gezielt angesprochen werden können.

Die Möglichkeiten, mit alternativen Methoden die Problematik der Ausreißer zu reduzieren, sind mit den vorgestellten Vorgehensweisen noch nicht ausgeschöpft. Diverse weitere Möglichkeiten stehen zur Verfügung. Die LAD werden der Gruppe der Techniken der Robust Regressions zugeordnet. In dieser Gruppe stehen weitere Alternativen zur Verfügung. Hier gilt es, sorgfältig abzuwägen, welche Vor- und Nachteile die jeweiligen Methoden mit sich bringen und dann zu entscheiden. Eine grundsätzlich dominante Methode zur Behandlung von Ausreißern in der linearen Regression besteht also für einen sorgfältigen Umgang nicht.

5 Zusammenfassung

Zu Beginn wurden die möglichen Quellen und Ursachen für Ausreißerwerte vorgestellt. Voraussetzung für einen verantwortungsvollen Umgang mit den extremen Werten ist ihre eindeutige Identifikation. Dafür stehen diverse Verfahren zur Verfügung. Sind die Werte identifiziert, so muss fallbezogen entschieden werden, welche Maßnahmen getroffen werden müssen. Eine pauschale Vorgehensweise ist nicht empfehlenswert.

Dabei kommt zunächst der Identifikation eine große Bedeutung zu. Nicht identifizierte und dementsprechend nicht behandelte Ausreißerwerte können substanziellen Einfluss auf die Ergebnisse einer Forschungsarbeit nehmen. Im Extremfall ist eine Umkehr der Vorzeichen einzelner Koeffizienten möglich, was entsprechend zu Handlungsempfehlungen führt, die direkt kontraproduktiv im Sinne des zu erreichenden Ziels sind.

Die Identifikation kann mit jedem statistischen Standard-Software-Paket durchgeführt werden. Alle hier vorgestellten Klassifikationen für Ausreißerwerte werden beispielsweise durch SPSS abgedeckt. Gegebenenfalls gewünschte Index-Plots können mit einfachen Funktionen z.B. in Excel dargestellt werden. Der Aufwand, der durch eine sorgfältige Vorbereitung entsteht, kann später auftauchende Probleme und den daraus resultierenden zeitlichen Aufwand erheblich reduzieren.

Mit Hilfe der vorgestellten Verfahren ist es nicht nur möglich, die Ausreißerwerte zu identifizieren. Vielmehr kann auch eine erste Einschätzung hinsichtlich der Bedeutung für das jeweilige Forschungsprojekt gemacht werden. Durch diese Bewertung kann entschieden werden, ob sich ein größerer Aufwand für weitere Maßnahmen lohnt. Kommt es zu substanziellen Veränderungen der im Fokus stehenden Koeffizienten, so sollten angemessene Maßnahmen ergriffen werden.

Die weit verbreitete Elimination identifizierter Ausreißerwerte ist nicht ausreichend für ein verantwortungsvolles Vorgehen mit Ausreißern. Vielmehr muss man sich mit den Gründen

für diese Werte auseinandersetzen. Zunächst muss geklärt werden, ob die Werte aufgrund von Fehlern entstanden sind. Ist das nicht der Fall, müssen mögliche inhaltliche Gründe untersucht werden. Gegebenenfalls können dann einzelne Werte aufgrund inhaltlicher Überlegungen ausgeschlossen werden. Ein solches Vorgehen muss gut dokumentiert werden, damit nachvollziehbar ist, welche Werte aufgrund welcher Überlegungen ausgeschlossen wurden.

Sollte sich bei der inhaltlichen Aufarbeitung der Ausreißerwerte zeigen, dass mehrere Beobachtungen aufgrund ähnlicher Erklärungen entstanden sind, so sollte eine Umformulierung des Modells in Betracht gezogen werden. Ein zunächst angenommener linearer Zusammenhang kann sich beispielsweise als nichtlinear herausstellen. Gegebenenfalls sind auch zusätzliche methodische Erweiterungen notwendig, die die Abbildung von Segmenten ermöglichen.

Sollte all dies nicht dazu führen, dass die Probleme mit den Ausreißern gelöst werden können, so ist eine reine Elimination der Beobachtungen denkbar. Sie sollte jedoch das letzte Mittel bleiben und gut dokumentiert werden.

Literatur

- Aghion, P., Howitt, P. & Mayer-Foulkes, D. (2005).** The Effect of Financial Development on Convergence: Theory and Evidence. *The Quarterly Journal of Economics*, 120, 173-222.
- Belsley, D.A., Kuh, E. & Welsch, R.E. (1980).** *Regression diagnostics: identifying influential data and sources of collinearity*, New York.
- Chien, L.-C. & Tsou, T.-S. (2007).** Regression Diagnostic under Model Misspecification. *Journal of Applied Statistics*, 34, 563-575.
- Cohen, J., Cohen, P., West, S.G. & Aiken, L.S. (2003).** *Applied multiple regression/correlation analysis for the behavioral sciences*, Mahwah.
- Cook, R.D. & Weisberg, S. (1982).** *Residuals and influence in regression*, New York.
- Creech, J.C., Corzine, J. & Huff-Corzine, L. (1989).** Theory Testing and Lynching: Another Look at the Power Threat Hypothesis. *Social Forces*, 67, 626-630.
- Gao, S., Li, G. & Wang, D. (2005).** A New Approach for Detecting Multivariate Outliers. *Communications in Statistics: Theory & Methods*, 34, 1857-1865.
- Greene, W.H. (2008).** *Econometric analysis*, Upper Saddle River.
- Hill, C.W. & Rothaermel, F.T. (2003).** The Performance of Incumbent Firms in the Face of Radical Technological Innovation. *Academy of Management Review*, 28, 257-274.
- Hoaglin, D.C. & Welsch, R.E. (1978).** The Hat Matrix Regression and ANOVA. *The American Statistician*, 32, 17-22.
- Imon, A.H.M.R. (2005).** Identifying Multiple Influential Observations in Linear Regression. *Journal of Applied Statistics*, 32, 929-946.
- Langford, I.H. & Lewis, T. (1998).** Outliers in multilevel data, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 161, 121-160.
- Leeflang, P.S.H., Wittink, D.R., Wedel, M. & Naert, P.A. (2000).** *Building Models for Marketing Decisions*, Boston.

- Mahajan, V., Sharma, S. & Wind, Y. (1984).** Parameter Estimation in Marketing Models in the Presence of Influential Response Data: Robust Regression and Applications. *Journal of Marketing Research*, 21, 268-277.
- Martin, M.A. & Roberts, S. (2006).** An Evaluation of Bootstrap methods for Outlier Detection in Least Squares Regression. *Journal of Applied Statistics*, 33, 703-720.
- Mullen, M.R., Milne, G.R. & Doney, P.M. (1995).** An International Marketing Application of Outlier Analysis for Structural Equations: A Methodological Note. *Journal of International Marketing*, 3, 45-62.
- Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y. & Podsakoff, N.P. (2003).** Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology*, 88, 879-903.
- Seaver, B.L. & Triantis, K.P. (1995).** The Impact of Outliers and Leverage Points for technical Efficiency Measurement Using High Breakdown Procedures. *Management Science*, 41, 937-956.
- Spector, P.E. (2006).** Method Variance in Organizational Research – Truth or Urban Legend?, *Organizational Research Methods*, 9, 221-232.
- Sutradhar, B.C., Chu, D.P.T. & Bari, W. (2007).** Estimation effects on powers of two simple test statistics in identifying an outlier in linear models. *Journal of Statistical Computation and Simulation*, 77, 305-328.
- Van Doorn, J., Verhoef, P.C. & Bijmolt, T.H.A. (2007).** The importance of non-linear relationships between attitude and behaviour in policy research. *Journal of Consumer Policy*, 30, 75-90.
- Wedel, M. & Kamakura, W.A. (2001).** *Market segmentation: conceptual and methodological foundations*, Boston.
- Weisberg, S. (1980).** *Applied Linear Regression*, New York.

C.Long-term effects of subsidies on firm growth: introducing the concept of outcome additionality

Björn Goerke und Sönke Albers

Veröffentlicht als Working Paper:

Goerke, Björn und Sönke Albers (2016): Long-term effects of subsidies on firm growth: introducing the concept of outcome additionality, <http://hdl.handle.net/10419/142164>

Abstract

Public agencies provide subsidies for small and medium sized businesses (SMEs) to foster their development in terms of employment and sales. Although input and output additionality have been researched intensively little is known about the actual long-term effects of subsidies on SME growth. Relying on a unique dataset of actual SMEs we provide a means of evaluating whether subsidies lead to the expected positive long-term effects. We apply a specifically designed 3-stage-effects-model (3SEM) from the input of resources to the final outcome. The results imply that the effects of subsidies differ across types: While R&D grants unfold enduring positive effects other subsidies like corporate matchmaking might even harm companies.

1 Introduction

Since research and development (R&D) is considered the source of future business, governments strive to support private R&D (David et al., 2000; Wolff and Reinthaler, 2008). The logic is that R&D will “deliver” new products and increase the productivity that a company needs to grow and to withstand its competition (Ortega-Argiles et al., 2009; Jones, 2002; Aghion and Howitt, 1992) which eventually leads to increased welfare.

However, since investing in R&D means investing under uncertainty firms tend to spend their funds for R&D only hesitantly (Scherer and Harhoff, 2000; Ortega-Argiles et al., 2009; Segerstrom, 1991; Griliches, 1979, 1992; Caballero and Jaffe, 1993). This is due to the disproportionate effect of a possible negative outcome of the project. With regard to welfare any

project that is favorable in a risk neutral setting should be carried out as it adds to overall welfare (Arrow, 1962; Nelson, 1959; Aerts and Schmidt, 2008). Nevertheless, firms may turn down such a project due to risk-aversion. While the decision to turn down the project may be the correct decision for risk-averse firms, this causes a loss for welfare. The gap between favorability in terms of welfare and favorability for the decision making firm is especially wide for SMEs causing disproportionately high losses in welfare (for a comprehensive discussion of the reasons why SMEs are affected more, see Ortega-Argiles et al., 2009).

Governments have been trying to compensate this market failure for long. A common way is to offer grants, tax breaks or subsidies in general. Transferring part of the risk from the company to the taxpayer shall make companies change their minds about R&D investments. The objective of these subsidies is to increase the odds that a favorable R&D investment is carried out in order to add to welfare (Aerts and Schmidt, 2008).

When public funds are spent it is of utmost relevance to understand whether the return for society is positive (Ortega-Argiles et al., 2009), i.e. whether companies, e.g., create new jobs and pay higher taxes. Thus, a large body of research has evolved in the domain of the effectiveness of subsidies for SMEs (David et al., 2000; Klette et al., 2000; Brockhoff, 1983).

The effectiveness is commonly evaluated in one of the concepts of additionality as a measure of additional value created by infusing the subsidy. Numerous authors have evaluated the degree of additionality that is produced by subsidies (Wallsten, 2000; Aerts and Schmidt, 2008; Czarnitzki and Licht, 2006; Hujer and Radic, 2005; Clausen, 2009), i.e. the authors analyzed whether the subsidies led to additional funds spent on R&D (input additionality) or whether the subsidies led to a complete crowding-out of private funding of R&D which means that the subsidy completely replaces the otherwise privately provided funds for R&D. Unfortunately, the findings are ambiguous. The degree of additionality found depends on different factors like industry, scope of the subsidy or even the appropriateness of the funding (Gelabert et al., 2009; Görg and Strobl, 2007).

Yet, we only know little about the effects of the subsidies on the actual development at firm level in terms of *outcome*, e.g. growth in number of employees, sales, and - finally - profits. Though former studies show that there are *additionality* effects regarding *input* (i.e. R&D budget that is provided through subsidies does not fully replace the private funding but leads to an increase in total R&D budget) this does not necessarily lead to a positive development of the subsidized firms with respect to the above measures that reflect the *outcome* (Merito et

al., 2010; Piva, 2005). It could as well be the case that companies that received grants to foster R&D are encouraged to take inadequate risks. Then the companies could even lose money and destroy jobs instead of providing new jobs and earning additional income. So, in order to award subsidies efficiently we need to understand how the return structure of the subsidies on the *outcome* of the awardees, e.g. number of employees, sales, and profits evolves over time. Only by understanding what the long-term effects of the subsidies on these *outcome* measures are, policy makers and subsidy granting agencies can make sure that the subsidies are not spent inefficiently in a systematic manner.

The analysis of the effects of subsidies on the *outcome* contributes both to research in the field of subsidies as well as to the management of the awarding public agencies. The contribution to research is providing knowledge about the long-term effects of subsidies on firm growth. It is widely unknown what the effects of subsidies on figures like sales, profits, and the number of employees are in the long run. Do companies hire more employees after having received subsidies? And if so, do these companies maintain these jobs even after the subsidies have been discontinued?

While most studies in the field of subsidies center around *input additionality* which is defined as the additional input factor that is resulting from the subsidy and would not have been available without the subsidy (Clarysse et al., 2009) - only very little is known with regard to the actual growth of the receiving firm. Lerner (1999) and Wallsten (2000) shed some first light on the question but little has been contributed since. In this article we set forth a means of assessing the long-term effects and formalize a new concept for the effects on actual company development.

This question is highly relevant to the taxpayer because a better understanding of the long-term effects of subsidies leads to better management of this multi-billion dollar topic. Policy makers and the management of the awarding agencies benefit from the enhanced opportunities of structuring subsidy regulations and by selecting the awardees more profoundly. This will lead to better social returns of the subsidies.

We collected a unique dataset of annual report data from a set of companies that were subsidized. Due to an exemption from reporting requirements small and medium sized companies do not have to report their annual figures on a basis as detailed as it is necessary for our

study.² Therefore the companies in our sample had to be acquired one by one by contacting top management and convincing them of the importance of this study.

This article is organized as follows. In the following section we will present the current state of the literature. Then we move on to the methodological section where we present the data, the statistical method, and the model. Finally we present and discuss the results and their implications and present the limitations.

2 Introducing the concept of outcome additionality

For a company to grow investing in R&D is essential. Governments have been trying to support companies in their respective jurisdictions for long. The effects of those subsidies are called additionalities (Clarysse et al., 2009, David et al., 2000). Additionalities can be observed – or at least – investigated at different stages in the innovation process. There can be input-additionalities with regards to the resources that a company is willing to provide for R&D. Then there may be additionalities with regard to the output of the R&D projects and finally the companies might grow or stay the same (outcome). In the following, we categorize these different types of additionality by building on the work of Clarysse et al. (2009).

It begins with companies investing in their R&D portfolios. Companies need to provide input factors (e.g. expert staff, licenses etc.) in order to stimulate their own innovations and to come up with new products and services. The concept of *input additionality* aims to capture the effects of input factors when governments subsidize R&D projects. The main question is whether companies simply reduce their own R&D budget “Dollar by Dollar”, as Wallsten (2000) reports, which means a complete crowding-out of the private budget or whether subsidies add (at least to some degree) to the total R&D budget that companies provide for their R&D projects (Clarysse et al., 2009).

The second stage in the process is the *output* that can be achieved. Companies strive to develop new products and services. These are the obvious output figures. But there are other outputs that count as well. Companies might file new patents, enter into partnerships, and the like. These are examples of different outputs of innovation activities. Whenever subsidies add to these output figures this is called *output additionality* (Clarysse et al., 2009). This is without doubt an important concept as the companies’ growth dynamics might rely on strong out-

² While all companies are required to file their annual reports with the respective courts small and medium-sized companies only need to report a reduced set of numbers. E.g. companies do not need to report the number of employees and the sales figures can be reported in an aggregate way where the true sales figures cannot be derived anymore. Statistics on R&D budgets or patents and similar information are not required either.

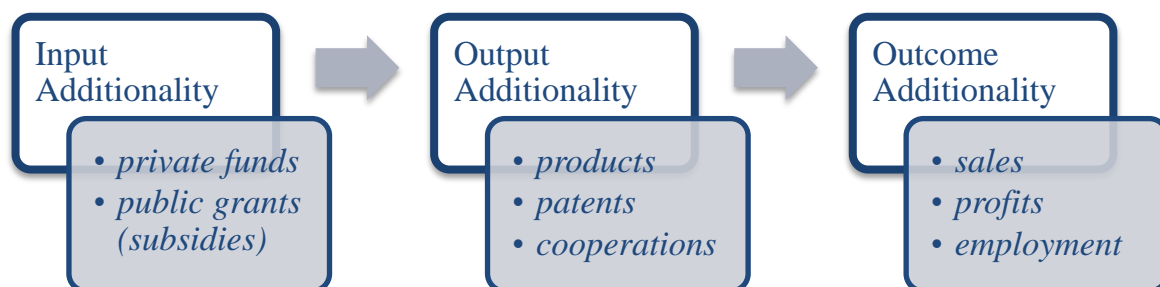
put figures. Still, output additionality does not cover the final and most relevant aspect: Does a company finally grow because it was awarded subsidies?

In the end companies as well as subsidy granting agencies and governments need to investigate the question whether subsidies actually add to the growth of a firm, i.e. do companies that receive subsidies grow in their sales and profit more than they would without the support? Public stakeholders also need to understand whether employment is positively affected as well. So, in order to create welfare-positive subsidy programs we need to understand whether and how much subsidies add to the core of these outcome figures. Accordingly we summarize this as *outcome additionality*. This is the third and final stage in the process of observing the effects of subsidies on the individual company level.

Figure 1 provides an overview of the three stages and the variables that are commonly measured. The further to the right the closer the concepts get to the effect on how the firms actually develop. The first two concepts lack the outcome of the subsidies. So, in case the subsidy creates substantial *input additionality* this does not necessarily lead to a positive development of the respective firm. It is the same with *output additionality*. In case a company invests heavily in new products because it was supported by subsidies this does not automatically lead to thriving sales figures or even employment numbers – it could well be the other way round when a company is being de-focused because of new projects and omits marketing the existing products.

Nevertheless, the concepts build on each another. *Outcome additionality* cannot be evaluated alone. We need to understand the whole process. Hence, we will set forth the current state of research in all three fields of additionality. Figure 1 shows the relationship of the different concepts of additionality. We call this the 3-stage-effects-model (3SEM).

Figure 1: Relationship of the different concepts of additionality



3 Literature Review

For a better understanding of the context that *outcome additionality* resides in we will first give a brief overview over the two concepts of input *and* output additionality. Both concepts will be reviewed regarding the findings in research. We will present the influences of the subsidies that were found and make use of those findings for a brief synthesis.

3.1 Input Additionality

Input additionality centers on the question whether the award of public grants leads to a crowding-out or even to a crowding-in of private R&D funds (Clarysse et al., 2009). A complete crowding-out of private funds by public support would mean that the firms simply replace their own funds by public funding and thus shift the risk from inside the company to the taxpayers. Crowding-in on the contrary means that companies invest even more in R&D than they would have without the support of public grants. As a measure of success of a subsidy program Georghiou (2002) suggested considering a project being a success in terms of public funding if the company invests at least the same amount of money that it is awarded as a subsidy.

The body of research in this field is so large that David et al. (2000) meta-analyzed the findings of the studies. The authors come to the conclusion that the studies deliver ambiguous results. In their sample of studies R&D subsidies could either add to R&D activity or they could just crowd-out private R&D funding. Blanes and Busom (2004) in their study of Spanish R&D subsidies review the literature with the same results.

Both research teams state that the results depend on different factors such as firm size and industry (David et al. 2000). Blanes and Busom (2004) discuss differing participation patterns in the different subsidy programs and differing goals of the agencies controlling the subsidies as a cause for the divergent findings.

In the following, we will give an overview over a sample of relevant studies in this field indicating whether the studies find crowding-out or crowding-in and the influencing factors they found if any (table 1). It seems as if there was no simple answer to the question whether public R&D financing adds to R&D activity or not.

Table 1: Literature in the field of input additionality

Author	Year	Journal	Crowding-out vs. crowding in
Wallsten	2000	RAND Journal of Economics	Complete crowding-out
Bönte	2004	Research Policy	Crowding-out in lower technology industries, positive effects in high-tech industries
Feldman and Kelley	2006	Research Policy	Partial crowding-out
Görg and Strobl	2007	Economica	Crowding-in for small subsidies
Aerts and Schmidt	2008	Research Policy	No crowding-out
Gonzales and Pazo	2008	Research Policy	No crowding-out
Özcelik and Taymaz	2008	Research Policy	Crowding-in and partial crowding-out in one sample, depending on company size

Therefore, the investigation of a possible crowding-out of private funds is highly relevant. The objectives of subsidy programs mainly center on providing new jobs and increasing sales, profits, productivity, or the like. If the R&D projects were largely carried out anyhow, i.e. with or without public support, this goal could not be met. Nevertheless, it is only a necessary condition for firm growth. It could still be that companies extend their R&D activities and do not develop better than they would have without (Merito et al., 2010; Piva, 2005). There can be different reasons, e.g. companies could be encouraged to take inadequate risks or companies might not be able to take advantage of the newly developed products since the subsidies only support R&D but not the similarly important sales resources.

3.2 Output Additionality

Output additionality takes research a step closer to the actual growth of the awardees. The goal of research in this domain is to analyze the effects of the awarded subsidies on more direct and immediate output figures. These are figures like new patents, new products and services, productivity, and so forth (Clarysse et al., 2009). Some studies go into more detail with regard to these output figures. E.g., Bérubé and Mohnen (2009) take a closer look at the degree of innovativeness that companies achieve in their new product development.

Research on output additionality draws less attention than input additionality. Nevertheless, there is research that establishes an understanding of the relevant processes in output additionality. Klette et al. (2000) worked out an overview in their meta-analysis of five studies that were in part focused on output additionality. The findings are predominantly positive. Table 2 gives an overview over the studies in the field of output additionality.

Table 2: Literature in the field of output additionality

Author	Year	Journal	Findings
Griliches and Regev	1995	Journal of Econometrics	Positive effect on productivity
Branstetter and Sakakibara	1998	Journal of Industrial Economics	Positive effect on patents filed
Klette and Møen	1999	Nordic Journal of Political Economy	Negative influence on productivity (assumed to be due to negative selection bias)
Hujer and Radic	2005	Scottish Journal of Political Economy	Positive effect on the number of new products
Czarnitzky and Licht	2006	Economics of Transition	Positive effect on patents filed
Berubé and Mohnen	2009	Canadian Journal of Economics	Positive effect on new products

These findings are plausible. Since there is evidence that subsidies do not crowd out private funds completely there must be additional R&D resources in the game. And as long as there are companies that spend the money on their R&D we should find additional output. Although these findings are desirable from the public point of view there is still no evidence that these positive conclusions result in better company growth (Hujer and Radic, 2005; Falk, 2007). Companies could well be able to develop a whole set of new products but if they are not able to market them efficiently these companies might even take financial harm from the R&D that they were encouraged to undertake. Therefore, we move on to studying the actual growth of the outcome of the companies that received public funding (Czarnitzky and Licht, 2006).

3.3 Outcome Additionality

The final step in evaluating the effects of R&D subsidies is to analyze the actual development of the companies after having been awarded. The core outcomes for assessing firm development are the number of employees and sales. Further on we will address the additionality associated with these metrics as *outcome additionality*.

Astonishingly, the presence of outcome additionality in research is still in its infancy. To the best of our knowledge there is no meta-analysis yet covering this topic but only two articles discussed below. Please note that we will emphasize the long-term effects achieved over time that the subsidies have on outcome. In contrast to *input additionality*, where there is a *point in time* when the decision on the external funding is made, outcome additionality *evolves over time* so that the long-run effects are of importance. With respect to *outcome* there is no “natural” ending to the investigation of the effects on, e.g. sales (Falk, 2007). As long as it is statistically meaningful we will investigate how the subsidies and the subsequent inputs and outputs carry-over to the final outcome.

There are two articles that deal with the well-known Small Business Innovation Research program (SBIR). The SBIR was tailored for small businesses in the USA. In the course of this program high-tech SMEs in the USA were provided with more than US-\$ 7 billion in grants for R&D between 1983 and 1997.

Lerner (1999) researched the effects of the SBIR. He analyzed the effects of the funds granted on sales and on the number of employees of the awardees over a ten year period starting 1985. The author constructed a sample of awarded and matching non-awarded firms. Lerner (1999) had the 1985 and the 1995 sales and employment figures of the companies to compare. He tested these data for significant growth of the awarded firms over the non-awardees. The awarded firms grew significantly better. On average the awardees reported an increase in sales of US-\$ 4.0 Mill. compared to an increase of US \$ 1.1 Mill. of the non-awardees. With regard to employment the figures were positive as well. The mean increase in employment was 26 employees compared to the increase of the non-awardees of 6 additional jobs.

Lerner (1999) reports a highly skewed distribution with the top awardees contributing disproportionately high to these positive findings. Lerner (1999) also reports variations in the findings due to different influences like region, industry, the number of SBIR awards received, or venture capitalist activity in the area. This more detailed analysis suggests that the average increase of the outcome mainly stems from companies in certain areas with a lot of activity by venture capital firms.

Wallsten (2000) researched the SBIR as well and came to very different conclusions. He criticized the endogeneity bias in Lerner's study. According to Wallsten (2000) it may well be that companies do not grow because of the grants they received but that they – the other way round – only received the grants because these companies were the faster growing ones. Therefore, endogeneity needed to be controlled for explicitly.

The way Wallsten (2000) constructed the sample for his study allowed for addressing the endogeneity bias. The author makes use of a system of equations that is being simultaneously estimated making use of 3SLS. One equation is designed to capture the influence of firm-size on the probability of receiving an award. The influence is positive and significant. But the estimates of the remaining equations (Wallsten covers two different phases of the SBIR program) do not indicate a significant influence of the awards on employment.

Due to specific requirements for this procedure Wallsten (2000) could only make use of data supplied by NASA and the US Department of Defense (DoD). Since there was a wide range

of agencies awarding the SBIR grants this is a substantial restriction as NASA and DoD act as awarding agencies and as potential customers for the newly developed product at the same time which was also acknowledged by Wallsten (2000) in the same article. Though, other agencies may act in these two roles as well, NASA and DoD are acting in very special markets. Wallsten explains why he thinks that the sample is representative, anyhow. Nevertheless, this severe restriction in the sampling process remains.

The findings of Wallsten (2000) are contrary to those of Lerner (1999). According to Wallsten (2000) the impact of the grants on the KPI development of the awardees is insignificant. Additionally, Wallsten (2000) finds that larger and more patent-intensive firms are more likely to win the awards. Altogether, Wallsten concludes that the grants awarded under the regulations of the SBIR programs were crowded-out “Dollar by Dollar”.

In another case, Merito et al. (2010) investigate the effects of R&D grants on several firm-level indicators of manufacturing firms in Italy. The authors point out that even if the private funds are not crowded out a rise in sales or employment does not have to be the result. Therefore, they compare groups of awarded companies to groups of non-awarded companies with regard to the development of the outcome. They construct a sample for short-term effects (2000 - 2002) and a sample for medium-term effects (2000 - 2004). While there are no overall significant effects on the development of the awardees, for the subset of SMEs the authors find a positive effect on employment in the medium term.

The inconclusive results point out the need to learn about subsidy effects in the long run (Falk 2007) but also stress the challenges that come with a long-term study of R&D grants on firm-level indicators. Lerner (1999) and Wallsten (2000) have started this important stream in research. Both scholars demonstrated that it is a feasible task to tackle the long-term effects but that it is also necessary to control for effects that merely stem from company characteristics that may have an influence on the probability of receiving awards. The ambiguity of their findings suggests that there is a wide range of challenges that need to be researched in order to come to a generalized long-term evaluation in terms of outcome additionality. Merito et al. (2010) add to understanding the effects but still their findings are ambiguous. By analyzing different time frames they allow for the first time to distinguish between different intervals of company development.

In the following we will present an approach that deals with the different challenges of analyzing the long-term effects of the subsidy. We start by elaborating the theoretical framework for the study.

4 Hypotheses

We make use of the resource-based view of the firm introduced by Wernerfelt (1984) and Barney (1991). According to the resource-based view a firm consists of a bundle of manifold resources. They usually comprise funds, machinery and knowledge (e.g. patents or exceptionally skilled staff) but they could also consist of brand value, distribution channels, or the like. Altogether the resources enable the firm to deliver its products and services, to grow, and to withstand its competition. In a situation of a lack of resources a company will not be able to grow and might even be forced to exit the market.

According to the resource-based view a company that acquires additional resources permanently lifts itself to a higher level of company development and possibly to an improved market position. This means, e.g., that a company that has acquired an improved market position by hiring an excellent salesperson will continue to benefit from this shift as long as the salesperson stays within the company – or until it is otherwise eliminated from the list of competitive advantages. This could occur, e.g., by direct competitors improving their own sales competences as well. Subsidies add to the bundle of resources as well. When a company receives a grant in order to carry out an R&D project this is an infusion of a new resource that the company can make use of.

Another aspect evolves along the question whether it is merely the resources that hinder companies from developing (Merito et al. 2010). Especially for non-SMEs the assumption is that companies can rather easily access capital markets that will provide companies with new funds in case the company can prove the upcoming projects to be favorable. SMEs may not have staff dedicated to managing the interactions with capital markets. Rather these companies rely on senior managers that handle finance as only one aspect of their management tasks. Following these arguments the resource-based view is especially relevant for SMEs.

As was shown by Merito et al. (2010) SMEs will develop better with regard to firm growth in case that they receive additional resources from the outside. Since the resource-based view does not distinguish between the different types of resources we can form one single hypothesis that can be applied to any subsidy awarded to SMEs:

Hypothesis 1a. *The growth of an SME with regard to the outcome is positively associated with subsidies awarded to the SME.*

According to the resource-based view the effects of the resources that were added to the bundle of resources will not diminish over time. This even accounts for hiring a highly skilled R&D manager who has knowledge the value of which could decay over time. On the other hand shifting a company to a new level of productivity could also lead to disproportionate growth due to overcoming barriers of growth. In order to test whether the assumption of an enduring and constant effect of the additional resources holds we state the following second hypothesis.

Hypothesis 1b. *The change in outcome is enduring and constant.*

5 Model

For testing the hypotheses we designed a sequential equation approach where *input additionality* is explained by subsidies, *output additionality* is explained by *input additionality* and subsidies and finally *outcome additionality* by *input and output additionality* together with subsidies. Thus, we cover the influences the complex structure of effects has on *outcome additionality*. The framework is displayed in figure 2.

In the first stage, input additionality, we employ a model that reflects the influences of the subsidies on the R&D budget that is being provided by the companies. We introduce firm size measured as number of employees to control for endogeneity that might result from the pure size of the company as Wallsten (2000) found in his study of the SBIR.

The second stage, output additionality, is structured in a similar way. We make use of the number of patents filed as output measure.³ The independent variables consist of the subsidies and controls again but they are extended by the R&D budget provided by the companies which was the dependent variable in the first stage. This way we connect the two intertwined stages.

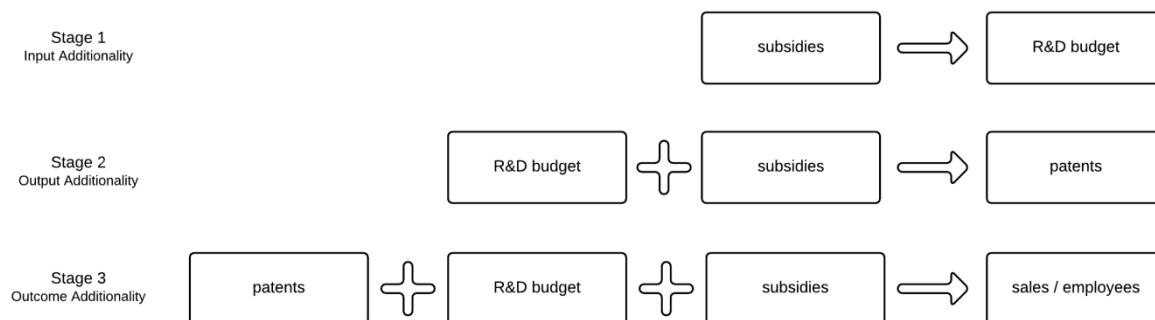
Finally, the third stage, the stage of outcome additionality, is developed in the same way. The independent variables consist of the subsidies and controls but also of the dependent variables of the two preceding stages. Three key measures to assess whether the aim of growing was

³ There is a wide array of different measures for output additionality with new patents being in line with former research. We opted for patents, first, as filing a new patent is clearly defined and unambiguously measurable whereas new products, e.g., will always be less precise to measure as some respondents might interpret a new feature of an existing product a new product while others may not. Thus, patents reduce noise in the data. Second, patents are often filed substantially earlier than the release of the products that are based on the respective patents. This reduces noise in the data as well.

accomplished are sales, number of employees, and profits.⁴ While profits can easily be manipulated by the management (e.g. profits can be deferred to later periods or to other entities), the sales figures and the number of employees are rigorously defined and reported.

This set of models allows for a comprehensive investigation of the effects of subsidies in all the stages. Furthermore, this also allows assessing the impact of preceding stages on the following stages.

Figure 2: Stylized overview of the 3SEM frameworks



6 Data

The agency providing the data for this study, WTSH, is a so-called “One-Stop-Agency”.⁵ It is the largest public agency covering the region with a wide portfolio of subsidy instruments. Within this portfolio 4 of these instruments are commonly used. These 4 instruments are described in more detail below and will serve as independent variables in this study. The hypotheses stated before will be tested on all of these instruments.

6.1 The subsidy instruments

6.1.1 R&D grants

SMEs receive a substantial amount of funds for single R&D projects. The grants do not have to be repaid. The awardees only need to document that they spent the grants according to the project proposal. The purpose is to come up with new products and services.

6.1.2 Patent consultancy

Companies are offered a large array of patent-related support. As patenting is a difficult and lengthy process many SMEs do not take advantage of the chance to safeguard their intellectual property rights. WTSH offers extensive support to the SMEs in the area with regards to

⁴ These measures are common targets of subsidy programs. Programs specialized in single industries or technologies may have other target measures.

⁵ The Wirtschaftsförderung und Technologietransfer Schleswig-Holstein GmbH (short: WTSH) is located in Kiel, Schleswig-Holstein, Germany. It is the general subsidy awarding agency in the federal state of Schleswig-Holstein. Please refer to <http://www.wtsh.de> for further information.

patenting. Companies can take part in information workshops, receive personal assistance in developing an IP strategy, have patent databases researched, and make use of support filing a new patent. Beware that a strategy for intellectual property protection may be not to patent the knowledge but to exploit it as fast as possible (time-to-market).

6.1.3 Corporate matchmaking

WTSH also offers a service of matching two or more companies that seek to find partners in developing new products. This matchmaking is not aimed at one company developing and producing the products and the other one selling them. WTSH engages in finding technological partners that, e.g., enable a company to develop a new technological approach by making use of the competencies of a partner.

6.1.4 Trade fairs

WTSH offers taking part in international trade fairs at a price considerably below market level. Additional to that companies receive further support on location, e.g. they are invited to networking activities. Besides offering a discount for trade fairs this instrument is especially designed to increase a firm's probability to engage in export business.

6.2 The Dependent Variables

In each of the three stages we observe a specific dependent variable that could be affected by the subsidies. In the third stage we estimate the effects of the two dependent variables of the earlier stages on the last stage.

Stage 1 – input additionality

Alongside the large body of research in the field of input additionality (e.g. David et al., 2000) we investigate the R&D budget that a company invests in new product development in a given year. The variable is operationalized as numeric in thousands of Euro. Companies were asked to state the budgets they actually used for new product development. As these figures are not reported in the annual Profit & Loss accounts the respective managers were required to estimate the figures.

Stage 2 – output additionality

Output additionality is commonly measured as direct results of R&D. This could be, e.g., licenses, patents, new product, or the like. As new patents filed are unambiguously measurable this is a favorable output measure to research on. Accordingly, in our study we use new patents filed in the respective year as the figure we build the model upon.

Stage 3 – outcome additionality

Following Lerner (1999) we focus on the outcome of the number of employees and the sales to assess the companies' growth. These outcome figures also have certain advantages: These figures will be reliable since they can be extracted from the companies' financial reports. With profits this is different. While this figure is part of the financial reports as well, companies can easily manipulate their profit and loss calculations for tax or other reasons, e.g. defer income to other periods. This would add noise to the data. Additionally, the companies under investigation do not have to report any of the figures we asked them for. The test of asking for the profits was vastly negative. Only very few companies provided this piece of information. Therefore the two outcome figures that we investigate are sales and employment. In the following data section we will discuss how these variables were coded.

In stages 1 and 2 that data on the number of employees also serves as a control variable for capturing possible influences that stem from firm size.

6.3 Dataset

The data in this study were collected by WTSH. WTSH strives for being in contact with every single SME in the region that invests in R&D. Thus, they cover the whole region in a way only a state-funded agency can operate.

Each company in our sample is regarded as an SME according to the subsidies rules of the European Union, i.e. fewer than 250 employees and annual sales below 50 million Euros or total assets below 43 million Euros.⁶ Since these companies have relaxed reporting regulations it is not possible to obtain appropriate data of companies of this size from databases. In an unprecedented effort WTSH collected this unique dataset by asking the companies' general managers to take part in this study and share their otherwise undisclosed financial information for this study.

The data comprises the years from 1996 until 2005. We have annual data on the 2 outcome figures of sales and number of employees for 33 companies. The sales data and the information on the R&D budget were deflated to 1996 prices, so that inflation effects are being controlled for. Inflation data was obtained from Destatis (German statistical office).

⁶ The term small and medium sized business is used differently throughout the literature in this field of research. We are using the official European definition. Many scholars in this field investigated reasonably larger companies – albeit, they are called SMEs as well. As the size of the company was often found to be an influence factor on the probability of receiving subsidies this is different in our sample as there are strictly no larger companies involved.

While the data on the dependent variables was provided by the companies themselves the data on the subsidy instruments were extracted from the customer database of WTSH. While the instruments of Trade fairs, Corporate matchmaking, and Patent consultancy were coded as a dummy variable (1 if subsidy awarded, 0 otherwise), R&D grants were coded in Euros and deflated to 1996 prices as well.

Since we obtained annual data covering a time span of 10 years we set up a panel dataset. Not all companies in our sample existed throughout the entire period, so our panel is unbalanced. The information on the subsidies received was assigned to the first year the subsidy was granted to allow for an accurate assessment of the long-term effects of the instruments.

The average annual sales figure of the companies in our dataset during the 10 year period is 5.08 million Euros. This sales volume was generated by an average number of employees of 37.4 per company. None of the companies in the study belongs to the services sector. The subsidies were granted for the development and promotion of products. Table 3 shows the descriptive statistics of the dataset.

Table 3: Descriptive statistics and correlations

	Dependent variables				Subsidy instruments			
	R&D budget	patents	sales	Employees	Patent consultancy	Corporate matchmaking	Trade fairs	R&D grants
Mean	245.118	4.030	5,081.595	37.415	.181	.242	.140	37.022
Std. Dev.	326.893	12.686	6,769.454	46.621	.386	.428	.369	123.283
Correlations								
	R&D budget	patents	sales	Employees	Patent consultancy	Corporate matchmaking	Trade fairs	R&D grants
R&D budget	1.000							
Patents	.218	1.000						
Sales	.299	.670	1.000					
Employees	.488	.746	.849	1.000				
Patent consultancy	.265	-.055	.148	.048	1.000			
Corporate matchmaking	.390	.260	.052	.203	.143	1.000		
Trade fairs	.119	-.028	-.046	-.043	.212	.237	1.000	
R&D grants	.238	.036	.107	.157	.178	.051	.101	1.000

7 Methodology

Building on previous research we constructed a framework that allows for a more precise and richer analysis. Testing for differences between two points in time limit the results to the specific time span of the respective study. The comparison of, e.g., sales figures at the time the subsidy was granted and a certain number of years later, does not allow for understanding whether this was the maximum of the effect. The annual observations of our panel dataset allow for a different, more information rich analysis. Specifically, employing the stock model approach we can test whether the effects of the subsidies are enduring or if there is a decay after the subsidy might have had an effect right after it was awarded. Following we describe the methodological approaches we employed in our study.

Panel Regression

First of all as we observe the companies repeatedly on an annual basis the dataset is structured in the form of a panel. Either Random Effects or Fixed Effects panel regressions can be employed to compute the effects inherent in the data. We carried out the Hausman specification test in order to decide for the method. The Hausman specification test indicated using the Random Effects panel regression ($\text{Prob} > \text{Chi}^2 = 0.0000$).

Grid Search

The second hypothesis states an enduring and constant effect of the subsidies. This means that the effects of the subsidies are not only observed in the period when they are granted but in the subsequent periods as well. In order to capture whether the effects are actually enduring and constant, diminish, or only last for the period when they are granted we employ the carry-over concept. Specifically we make use of the stock model by Nerlove and Arrow (1962).

Equation (1) reflects the Random Effects model. *model* depicts whether it is the equation for input, output or outcome additionality (either sales or employees), *i* depicts the subsidy instrument with *t* indicating the period. Instead of the common independent variables of the subsidy instruments the equation contains stock variables $s_t^{i,model}$. On the one hand the stock variables contain the information of the current period *t* but – on the other hand – they also contain the information on how much of the effect of previous periods has been carried over.

$$y_t^{model} = \alpha^{model} + \beta^{i,model} s_t^{i,model} + \varepsilon_t^{model} \quad (1)$$

Equation (2) depicts the stock model that is built by introducing the carry-over coefficient. The carry-over coefficient $d^{i,model}$ for every subsidy instrument *i* carries the information

which proportion of the effect of the previous periods is carried over. i denotes the respective subsidy instrument while t denotes the successive year of the subsidy starting at 0 in the first year. In case the carry-over coefficient $d^{i,model}$ is 0 no stock of the previous periods is build and the stock variable $s_t^{i,model}$ equals the original independent variable x_t^{model} . In case $d^{i,model}$ is 1 the whole value of the previous periods is added to the current stage. For any value between 0 and 1 the carry-over effect will lead to adding fractions of the values of the previous stages. The equation is as follows:

$$s_t^{i,model} = x_t^{model} + d^{i,model} s_{t-1}^{i,model} + \dots + d^{i,model} s_{t-n}^{i,model}, \quad (2)$$

where $s_t^{i,model}$ is the stock of instrument i in period t with n being the maximum amount of periods taken into account from previous years.

The empirical coefficients are derived in an iterative process known as Grid Search (Greene 2003, p. 566). We work with idiosyncratic carryovers per subsidy instrument and additionality measure. The boundaries of the variation of the coefficients are 0 (no carry-over of the effect at all with the effects taking place in the respective period only) and 1 (enduring and constant effect). In every iteration the model is estimated anew. Finally, the model with the best fit in terms of R^2 is chosen. This process determines the carry-over coefficients for the subsidies and the coefficients of the subsidies models jointly.

Control Function Approach

Our primary aim is to estimate the effects of the subsidies on growth in sales and employment. It is straightforward to include the subsidies as independent variables. In addition, we assume that the results of the two preceding stages of additionality will have influence as well. Thus, the additionality variables are endogenous. In order to control for endogeneity, we make use of the Control Function Approach (Petrin and Train, 2009). The Control Function Approach demands that the residuals of the regression explaining the endogenous variable are inserted in addition to the endogenous variable into the equation under investigation.

8 Results

The 3SEM framework of the additionalities results in a multi-step presentation. First we will focus on the degree that the subsidies affect input and output additionality. Afterwards we will show the estimates of the effects on sales and the number of employees.

The influence of the subsidies on the R&D budget that is being invested by the companies is being estimated. The results are presented in table 4.

Table 4: Estimates of stage 1 and 2 (input and out additionality)

Variable	Stage 1 R&D budget		Stage 2 Patents	
	Beta (Standard Error)	z-value (p-value)	Beta (Standard Error)	z-value (p-value)
R&D grants [in 1.000 €]	0.183 (0.074)	2.48 (0.013)	- 0.011 (0.004)	- 2.82 (0.005)
Patent consultancy [dummy 0/1]	123.496 (20.033)	6.16 (0.000)	- 3.932 (1.518)	- 2.59 (0.010)
Corporate matchmaking [dummy 0/1]	24.931 (10.803)	2.31 (0.021)	6.746 (1.512)	4.46 (0.000)
Trade fairs [dummy 0/1]	-47.819 (23.483)	- 2.04 (0.042)	- 1.444 (1.606)	- 0.90 (0.368)
Number of employees [nominal value]	3.960 (0.617)	6.42 (0.000)	0.148 (0.030)	4.99 (0.000)
R&D budget	-	-	- 0.003 (0.003)	- 0.97 (0.331)
Residuals of the Control Function Approach	-	-	-0.012 (0.007)	-1.74 0.082
constant	-4.494 (53.151)	-0.08 (0.933)	- 1.372 (1.742)	- 0.79 (0.431)
R ² overall	0.39		0.61	
R ² within	0.54		0.33	
R ² between	0.38		0.59	
Wald Chi ² (Prob > Chi ²)	199.43 (0.000)		113.13 (0.000)	

The model itself is significant (Wald = 190.43, $P > \text{Chi}^2 = 0.0000$) and it shows adequate quality with an overall R^2 of 0.39 and R^2 within and R^2 between in similar ranges. All coefficients of the model have significant estimates. While the effect of the subsidy of Trade fairs is negative the other variables show positive estimates. The variable Number of employees which was introduced to control for size of the companies is significant and positive indicating that the budget that is provided for R&D projects is positively correlated with the size of the company itself.

The second model shows a different picture (table 4). The aim of this model is to capture the influence of the subsidies on the number of patents that the companies file. In order to increase the explanatory power the R&D budget of the companies is introduced into this equation. As this variable is the dependent variable of the preceding stage we need to assume that

the error terms are not independent of each other. The expected endogeneity is then controlled for by adding the residuals of the first stage to this second stage equation (Petrin and Train, 2009).

The model itself is significant (Wald = 113.13, $P > \chi^2 = 0.0000$). R^2 of 0.61 is higher than in the first stage. Three of the four subsidy coefficients have significant estimates. While Trade fairs do not seem to affect the number of patents filed, R&D grants, Patent consultancy, and Corporate matchmaking are significant. However, the negative signs of R&D grants and Patent consultancy are not in line with expectations. Again the control variable Number of employees is positive and significant indicating that the number of patents filed is correlated with the size of the company as well. The residuals of the first stage are also significant. This implies that the R&D budget is endogenous while the residuals correct for this endogeneity so that the coefficient of R&D represents the corrected influence. The dependent variable of stage 1 itself, R&D budget, has no significant influence.

The third stage is the one we focus on. The results of both equations are presented in table 5. There are two models capturing the effects on the two outcome variables under investigation. On the one hand there is the model of the number of employees and on the other hand we have the model of the sales figures. For both models we also report the carry-over coefficients as described in the methodology section. The aim is to estimate the long-term effects of the subsidies.

The employees model of stage 3 has high explanatory power with an R^2 of 0.77 (Wald = 579.23, $P > \chi^2 = 0.000$). The model was enriched by the dependent variables of the preceding stages. Both variables – R&D budget and patents filed – carry a significant coefficient. So do the variables of R&D grants and Corporate matchmaking. Patent consultancy as well as Trade fairs are insignificant. Both carry-over coefficients of R&D grants and Corporate matchmaking are 1.00, indicating an enduring effect that, once it was installed, carries on to the following periods without decay. While R&D grants shows the expected positive sign Corporate matchmaking has a negative effect in growth on the number of employees. The residuals that were introduced according to the Control Function Approach are positive and significant. Again, this shows that the two equations are endogenous and that controlling for this endogeneity is necessary.

Table 5: Estimates of the third stage – employees and sales model (outcome additionality)

	Stage 3: employees			Stage 3: sales		
Variable	Carry-over coefficient d	Beta (Standard Error)	z-value (p-value)	Carry-over coefficient d	Beta (Standard Error)	z-value (p-value)
R&D grants [in 1.000 €]	1.00	0.027 (0.009)	2.95 (0.003)	1.00	- 1.291 (1.357)	- 0.95 (0.342)
Patent consultancy [dummy 0/1]	0.76	1.406 (3.026)	0.46 (0.642)	0.94	1,459.005 (319.301)	4.57 (0.000)
Corporate matchmaking [dummy 0/1]	1.00	-27.969 (5.531)	- 5.06 (0.000)	0.79	- 696.082 (396.611)	- 2.42 (0.015)
Trade fairs [dummy 0/1]	0.39	- 0.233 (4.375)	- 0.05 (0.957)	0.75	- 653.855 (396.611)	- 1.65 (0.099)
R&D budget [nominal value]		0.048 (0.007)	6.43 (0.000)		5.451 (1.140)	4.78 (0.000)
Patents [nominal value]		3.476 (0.227)	15.30 (0.000)		96.039 (25.753)	3.73 (0.000)
Residuals of the Control Function Approach		- 3.280 (0.431)	- 7.61 (0.000)		141.917 (129.742)	1.09 0.274
constant		16.466 (2.643)	6.23 (0.000)		3,799.028 (887.617)	4.28 (0.000)
R ² overall	0.77			0.34		
R ² within	0.22			0.40		
R ² between	0.95			0.34		
Wald Chi ² (Prob > Chi ²)	579.23 (0.000)			107.03 (0.000)		

Finally, the sales model of stage 3 shows adequate explanatory power as well. R^2 is 0.34 (Wald = 107.03, $P > \text{Chi}^2 = 0.000$). Again, two of the four subsidies carry significant coefficient estimates, namely Patent consultancy and Corporate matchmaking. While Patent consultancy shows the expected positive effect, Corporate matchmaking has a negative effect like in the employees' model of stage 3. The carry-over coefficient of Patent consultancy of 0.94 is close to one. This indicates a slowly diminishing effect over time. The carry-over coefficient of 0.79 for Corporate matchmaking indicates a faster diminishing process than Patent consultancy but, e.g., there still remain about 30% of the effect after 5 years. Like in the employees model both dependent variables of the preceding stages show significant and positive effects. Other than in the employees' model and the second stage model the residuals of the Control Function Approach are not significant. These two equations do not seem to be endogenous.

The hypotheses are partially supported. Table 6 presents the results regarding the hypotheses. Hypothesis 1a states a positive relationship between the subsidy and the outcome. In the employees model this is true for R&D grants while the findings of the remaining three subsidies are not in line with this hypothesis. Hypothesis 1b states that the effect shall be enduring. This means that the carry-over coefficient needs to be 1.00. With regards to the R&D grants of the employees' model this is the case. So, R&D grants support both hypotheses concerning the growth in employment.

Table 6: Overview of results for hypotheses

Overview of results for hypotheses					
		R&D grants	Patent consultancy	Corporate matchmaking	Trade fairs
employees	Hypothesis 1a Positive influence of the subsidies	yes	not significant	no, negative	not significant
	Hypothesis 1b Positive influence is enduring and constant	yes	-	-	-
sales	Hypothesis 1a Positive influence of the subsidies	not significant	yes	no, negative	not significant
	Hypothesis 1b Positive influence is enduring and constant	-	no, to some extent	-	-

In the sales model we find one variable supporting Hypothesis 1a of the positive effects on the outcome. This is the subsidy of Patent consultancy. Hypothesis 1b of endurance is not met perfectly but the high carry-over coefficient of 0.94 indicates a long-lasting effect of the subsidy. The three remaining subsidies are either insignificant or they even show a significant negative sign contradicting the first hypothesis.

All in all the results in table 6 show some support for the hypotheses. These results will be discussed in the following section.

9 Discussion

The finding of the positive influence of the R&D grants in stage 1 implies that these R&D grants do not result in a complete crowding-out. In fact, since the value is positive this means that the overall budget grows. A coefficient of 0.183 means that for a 1,000 € that a company is awarded the overall budget grows by 183 €. So, a substantial amount of the award replaces

the private funding by the company. What we observe is a partial crowding-out of the funds which is in line with the majority of the studies carried out in the field of input additionality.

The positive influence of Patent consultancy and Corporate matchmaking are plausible as both subsidies reduce the risk of an R&D project. This in return increases the expected returns and improves the favorability of the projects. The negative influence of Trade fairs seems counter-intuitive at first. But since taking part in a trade fair means investing cash-flow in this marketing instrument this results in a reduced free budget to be allocated. While we find a negative sign this does not automatically translate into a negative decision.

The second equation on output additionality captures the effects of the subsidies granted on the patents filed by the companies. Interestingly, the coefficient of Patent consultancy is negative which seems counter-intuitive. Patent consultancy of course comprises the consulting on how to file a patent for a new product. But a major task of Patent consultancy is working on an appropriate strategy on how to protect the intellectual property of a company. Especially for SMEs patenting technology can jeopardize the whole investment in R&D. In order to file a patent companies have to describe their technology in detail. This description is free to be investigated by any other company. So, first, other companies could just copy the patent. Though the technology is then protected by a patent it is within the responsibility of the patent holding company to defend its rights. This can be time-consuming and can cause large expenses. Second, other companies can learn how to solve a technological problem and build their own solution using this knowledge making sure that the technological gap to the patent is just big enough not to violate the existing rights. Thus, many SMEs refrain from patenting their intellectual property especially after they have learned about other strategies of protection of intellectual property. This often comprises a very high level of confidentiality and aggressive time-to-market.

The findings of a negative influence of R&D grants on the number of patents filed follow the same reasoning. When a company applies for an R&D grant it must explain in great detail how it is going to protect the intellectual property that is being created in the course of the R&D project. Because of this companies need to consider the threats explicitly which they do not have to do carrying out an R&D project by themselves. Going into detail the companies might learn about the threats associated for the first time. The positive influence of Corporate matchmaking on patents filed is interesting. Carrying out R&D projects in a joint venture with another company or even other companies increases economic power which in turn increases the probability of a successful defense of a patent violation. Developing new products with

another organization also means that confidentiality as a strategy might not work as it might not be possible to control the dissemination of information as is it needed.

The results of the third stage are the ones in our focus. Hypothesis H1a states that there is a positive effect of the subsidies on the growth of the outcome figures of the companies. The second hypothesis H1b states that this effect be enduring and constant which means that once the company has received this support the positive effect will not only occur in the year the subsidy was received but that this effect will be observed in the following years as well.

The results of the subsidy of R&D grants support the employment hypothesis H1a as well as hypothesis H1b. The positive effect of R&D grants carries over with a factor of $d = 1.00$. This means that the effect does not diminish over time. A new job that gets created because of the subsidy will be sustained. It is a different situation in the sales model of stage 3. The effect of the subsidy on sales is insignificant. This may be an indicator that companies mostly invest the subsidies in R&D staff as an investment in the sales team which should increase the sales capacity and should lead to an increase in sales which we do not observe. Since both dependent variables of the stages 1 and 2 - the R&D budget and the number of patents filed - have a significant and positive influence on the growth in employment as well as in sales these findings further support the results of the employment model while for the sales model this can be interpreted as a hint that there is a positive compound effect of the R&D grants received on sales. Since the carry-over coefficient d of the instruments are individually estimated for every equation this effect cannot just simply be computed by inserting the values of the preceding equations in the respective equations of stage 3.

Patent consultancy supports hypothesis H1a in the sales equation while it is insignificant in the employees' model. The carry-over coefficient d of 0.94 in the sales model is high and close to 1.00. While this finding does not support hypothesis H1b completely this is a partial acceptance of the hypothesis. Even after, e.g., 3 years 83.1% of the effect still remains. The high coefficient of 1,459,005 € of additional sales is in line with the results that Patent consultancy often has to offer for SMEs. While it might not be favorable to actually file a patent there are two common strategies that can be combined. One is confidentiality and the other one is aggressive time-to-market. The latter option aims at monetizing the intellectual property as fast as possible. In order to do so many companies may make use of distributions channels outside their own organization as growing the own sales team is time consuming and comes with an increased risk.

The direct effect of Corporate matchmaking on employment as well as sales is negative. Both coefficient estimators are significant and the values of the coefficient are relatively high implying that Corporate matchmaking will harm a company on its growth track. Hypothesis H1a is not being supported for both equations. Hypothesis H1b is supported with regards to the employment equation which - taking into account that the effect is negative - would be a warning signal for any company seeking support in Corporate matchmaking. With regards to employment the negative effect is enduring. Findings for H1b in the sales equation suggest a long-lasting albeit not an enduring effect with a carry-over coefficient of $d = 0.79$. This result partially supports hypothesis H1b. It says that after, e.g., 3 three years the negative effect of taking part in Corporate matchmaking still is at a level of 49.3%. Taking into account the effects of Corporate matchmaking on the preceding stages changes the conclusion on this subsidy. As Corporate matchmaking has significant and positive effects on both preceding stages the compound effect is positive as well. It is not only that the effects in the preceding stages are positive they are also considerably high in numerical terms. So, carefully compounded the effects of Corporate matchmaking indicate a slightly negative effect on employment and a slightly positive effect of sales.

The subsidy of Trade fairs is insignificant in both equations. Hence, the hypotheses are not supported with respect to this subsidy. While this seems counter-intuitive at first glance this result gets plausible when put into context with the findings of the two preceding stages with respect to Trade fairs. The effects of Trade fairs on R&D budget were negative. The investment in a trade fair does not only consume free funds, it also focuses the organization towards this event lowering the output in other parts of the company. Nevertheless, the compound effect of the negative influence on R&D budget in first stage implies a negative effect on both employment and sales. Anecdotal evidence from managers of SMEs in the region shed a light on this counter-intuitive finding. Most managers only expect positive returns from a specific trade fair after taking part a few times. The first two or three fairs will result in negative returns. Since WTSH focuses on supporting the first time of taking part in a specific trade fair the negative effects become plausible.

10 Implications

The implications are mainly driven by the findings on the different degrees of endurance of the subsidy instruments that we analyzed in this study. The results have implications on different levels. To start we take a closer look at what the results imply for research.

Research implications. As raised in the literature review of this paper the main body of research in the field of additionality focuses on input additionality. In this paper we apply a broader approach and investigate the three stages of input, output, and outcome additionality simultaneously. The findings clearly show that this broader approach is necessary in order to understand how the effects on input additionality actually proceed through the subsequent stages. The theoretical approach that input additionality will finally translate into company growth may fall short due to processes that are not factored in common theory.

The second implication is that of the long-term effects of subsidies. While input additionality focuses on a single decision made by the companies' management all other processes take time and the effects need to build up or the effects diminish. Here the resource based view falls short stating that the effects of newly added resources be enduring. The results of our study clearly show that this needs to be considered in more detail. While R&D grants have an enduring effect on the number of employees this does not show for the other subsidies. Research will need to investigate how subsidies that were once granted wear off and how this wear off can be slowed down in order to help policy makers improve the subsidy granting programs.

Finally, the findings of the negative effects of Trade fairs can be plausibly explained to some degree. Anyway, being able to explain the effect does not improve the result with regards to decision making. Talking to SMEs in the region we came to the conclusion that the decision to start using the marketing instrument of Trade fairs is something that will only pay off in the long run. Anecdotal evidence is that it takes a few repetitions of taking part in the Trade fairs before the contacts made on the fairs do actually turn into business as most business people conservatively want observe the progress a new player makes. WTSH especially supports the first time a company takes part in a trade fair. In order to investigate the effects of the Trade fairs on company growth one needs to design a study that is especially capable of capturing effects that may only occur 3, 4 or even more years later than the subsidy was granted.

Policy implications. The implications for policy makers and the management of the subsidy granting agencies are manifold. The positive effects of R&D grants and Patent consultancy are straightforward. The companies receiving this kind of support grow better than they would do without. This is in line with the aim of adding to welfare through subsidies. While the granting agencies could just carry on giving away the subsidies policy makers need to assess whether the outcome that follows the subsidies actually justifies the investment as there is not

only the subsidy itself that needs to be earned in terms of welfare but also process cost that are incurred just by granting subsidies.

With regards to the more complex results for the subsidies of Trade fairs and Corporate matchmaking this is not so straightforward. While the plain results of stage 3 suggest being very careful with awarding these subsidies further investigation of the effects seems reasonable. Trade fairs may pay off considerably later than the other subsidy instruments. While investing in Trade fairs may still be favorable the awarding agencies could increase the value of the subsidies by accelerating the process of repayment with regards to welfare. The agencies could provide a more comprehensive framework of marketing activities so that taking part in a trade fair is just one means in a set of options. Another option would result in giving away these subsidies more carefully. The agencies could check whether the applying companies are all set for the long run. If this is not the case agencies could simply turn down applications in order to save their own resources and also to protect the companies from engaging into activities they have trouble overseeing the consequences. This same change in mindset could be appropriate for Corporate matchmaking as well.

11 Limitations

As this study is an early one in the field of the research of outcome additionality we had to deal with a number of limitations. The differences in the results of the different instruments show that the results are not uniform across subsidy instruments. It is important to distinguish between the instruments in the market of subsidies and to understand the processes that are associated with the different instruments. So, in order to understand the whole field of subsidies further research is to be directed at additional subsidy instruments.

WTSH is the agency for one of the federal states in Germany. Though regulation in the EU provides some level of comparability across regions further research needs to collect data from other regions as well. Lerner (1999) found different effects across the zip code areas in the US. Further studies should investigate the influences the different regions and their unique patterns of the economy, capital resources, and education have.

Finally, matching methods are common in this field of research in order to cope with the (self)-selection bias that is to be expected in the awarding process. While the matching algorithms can provide some accuracy in case that there are matching observations we do not come to the conclusion that this is true for researching a set of broadly awarded, general subsidy instruments. There are numerous agencies in the field providing similar instruments to

the same companies. In order to arrive at reliable results in a matching process one would have to get the data of all agencies that provide the same sort of instruments. The publicly available database of subsidy instruments “Förderdatenbank” (www.foerderdatenbank.de) returns 117 results for companies seeking R&D grants in the federal state of Schleswig-Holstein.⁷ This is simply not a feasible number of programs to collect data from.

Self-selection bias

According to WTSH, the agency aims at contacting between 1.000 and 2.000 companies per year intensively. The agency actively approaches technology-oriented firms in its region. This usually leads to some sort of public support. Over a timeframe of 10 years which is what we cover in this study only a very small number of companies has not been supported by WTSH or other subsidy granting agencies. Due to the confounding effects that are to be expected by other subsidies and that cannot be controlled for it is merely not feasible to carry-out a reliable, non-biased matching approach which is widely used in subsidy research for which we would have had to form a sample of non-treated firms. Therefore we decided to use a panel-based approach which allows isolating the effects of every single subsidy that can be controlled for.

The advantage of the panel is that every company in the sample is treated with at least one of the subsidy instruments. Simultaneously, this company is not being treated with the other instruments. So, for these instruments we implicitly have a non-treated subsample. This allows assessing the difference between companies that were treated and those that were not treated. As described above it is not a realistic venture to try and construct a sample that was not treated by any instrument at all.

References

- Aerts, K. & Schmidt, T. (2008).** Two for the price of one? Additionality effects of R&D subsidies: A comparison between Flanders and Germany. *Research Policy*, 37, 806-822.
- Aghion, P. & Howitt, P. (1992).** A Model of Growth through Creative Destruction. *Econometrica*, 60 (2), 323-351.
- Arrow, K.J. (1962).** Economic welfare and the allocation of resources of invention. In: Nelson, R.R. (Ed.), *The Rate and Direction of Inventive Activity: Economic and Social Factors*. Princeton, 361-392.
- Barney, J. (1991).** Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17 (1), 99-120.

⁷ The result was retrieved on Jan 24, 2016.

- Bérubé, C. & Mohnen, P. (2009).** Are firms that receive R&D subsidies more innovative? *Canadian Journal of Economics*, 42 (1), 206-225.
- Blanes, J.V. & Busom, I. (2004).** Who participates in R&D subsidy programs? The case of Spanish manufacturing firms. *Research Policy*, 33, 1459-1476.
- Bönte, W. (2004).** Spillovers from publicly financed business R&D: some empirical evidence from Germany. *Research Policy*, 1635-1655.
- Branstetter, L. & Sakakibara, M. (1998).** Japanese Research Consortia: A Microeconomic Analysis of Industrial Policy. *Journal of Industrial Economics*, 46 (2), 207-233.
- Brockhoff, K. (1983).** The measurement of goal attainment of governmental R&D support. *Research Policy*, 12, 171-182.
- Caballero, R.J. & Jaffe, A.B. (1993).** How High are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth. In: Blanchard, O., Fischer, S. (Ed.), *NBER Macroeconomics Annual 1993*, Volume 8. MIT Press, 15-86.
- Clarysse, B., Wright, M. & Mustar, P. (2009).** Behavioural additionality of R&D subsidies: A learning perspective. *Research policy*, 38, 1517-1533.
- Clausen, T.H. (2009).** Do subsidies have positive impacts on R&D and innovation activities at the firm level? *Structural Change and Economic Dynamics*, 20, 239-253.
- Czarnitzki, D. & Licht, G. (2006).** Additionality of public R&D grants in a transition economy. *Economics of Transition*, 14 (1), 101-131.
- David, P.A., Hall, B.H. & Toole, A.A. (2000).** Is public R&D a complement or substitute for private R&D? A review of the econometric evidence. *Research Policy*, 29, 497-529.
- Falk, R. (2007).** Measuring the effects of public support schemes on firms' innovation activities – Survey evidence from Austria. *Research Policy*, 36, 665-679.
- Feldman, M.P. & Kelley, M.R. (2006).** The ex ante assessment of knowledge spillovers: Government R&D policy, economic incentives and private firm behavior. *Research Policy*, 35, 1509-1521.
- Gelabert, L., Fosfuri, A. & Tribó, J.A. (2009).** Does the Effect of Public Support for R&D Depend on the Degree of Appropriability? *Journal of Industrial Economics*, 57 (4), 736-767.
- Georghiou, L. (2002).** Impact and additionality of innovation policy. Innovation policy and sustainable development: can innovation incentives make a difference. 57-65.
- González, X. & Pazó, C. (2008).** Do public subsidies stimulate private R&D spending? *Research Policy*, 37, 371-389.
- Görg, H. & Strobl, E. (2007).** The Effect of R&D Subsidies on Private R&D. *Economica*, 74, 215-234.
- Greene, W.H. (2003).** *Econometric Analysis*. 566.
- Griliches, Z. (1979).** Issues in assessing the contribution of research and development to productivity growth. *Bell Journal of Economics*, 10, 92-116.
- Griliches, Z. (1992).** The search for R&D spillovers. *Scandinavian Journal of Economics*, 94, 29-47.
- Griliches, Z. & Regev, H. (1995).** Firm productivity in Israeli industry 1979-1988. *Journal of Econometrics*, 65 (1), 175-203.
- Hujer, R. & Radic, D. (2005).** Evaluating the Impacts of Subsidies on Innovation Activities in Germany. *Scottish Journal of Political Economy*, 52 (4), 565-586.

- Jones, C.I. (2002).** Sources of U.S. Economic Growth in a World of Ideas. *American Economic Review*, 92 (1), 220-239.
- Klette, T.J. & Møen, J. (1999).** From Growth Theory to Technology Policy – Coordination Problems in Theory and Practice. *Nordic Journal of Political Economy*, 53-74.
- Klette, T.J., Møen, J. & Griliches, Z. (2000).** Do subsidies to commercial R&D reduce market failures? Microeconomic evaluation studies. *Research Policy*, 29, 471-495.
- Lerner, J. (1999).** The Government as Venture Capitalist: The Long-Run Impact of the SBIR Program. *Journal of Business*, 72 (3), 285-318.
- Merito, M., Giannangeli, S. & Bonaccorsi, A. (2010).** Do incentives to industrial R&D enhance research productivity and firm growth? Evidence from the Italian case. *International Journal of Technology Management*, 49 (1), 25-48.
- Nelson, R.R. (1959).** The Simple Economics of Basic Scientific Research. *Journal of Political Economy*, 67 (3), 297-306.
- Nerlove, M. & Arrow, K.J. (1962).** Optimal Advertising Policy under Dynamic Conditions. *Economica*, 29, 129-142.
- Ortega-Argilés, R., Vivarelli, M. & Voigt, P. (2009).** R&D in SMEs: a paradox? *Small Business Economics*, 33 (3), 3-11.
- Özçelik, E. & Taymaz, E. (2008).** R&D support programs in developing countries: The Turkish experience. *Research Policy*, 37, 258-275.
- Piva, M. & Vivarelli, M. (2005).** Innovation and Employment: Evidence from Italian Microdata. *Journal of Economics*, 86 (1), 65-83.
- Petrin, A. & Train, K. (2009).** A Control Function Approach to Endogeneity in Consumer Choice Models. *Journal of Marketing Research*, 47 (1), 3-13.
- Scherer, F.M. & Harhoff, D. (2000).** Technology policy for a world of skew-distributed outcomes. *Research Policy*, 29, 559-566.
- Segerstrom, P.S. (1991).** Innovation, Imitation, and Economic Growth. *Journal of Political Economy*, 99 (4), 807-827.
- Wallsten, S.J. (2000).** The effects of government-industry R&D programs on private R&D: the case of the Small Business Innovation Research program. *RAND Journal of Economics*, 31 (1), 82-100.
- Wernerfelt, B. (1984).** A Resource-Based View of the Firm. *Strategic Management Journal*, 5, 171-180.
- Wolff, G.B. & Reinthaler, V. (2008).** The effectiveness of subsidies revisited: Accounting for wage and employment effects in business R&D. *Research Policy*, 37, 1403-1412.

D.Neuprodukteinführungen in der Filmindustrie: Wie reagieren Kapitalmarktinvestoren auf den Umsatzerfolg neuer Kinofilme?

Michel Clement, Marc Fischer und Björn Goerke

Erschienen als:

Clement, Michel, Marc Fischer und Björn Goerke (2007): Neuprodukteinführungen in der Filmindustrie: Wie reagieren Kapitalmarktinvestoren auf den Umsatzerfolg neuer Kinofilme?, Die Betriebswirtschaft, 64. Jg., Heft 4, 418-444.

Abstract

In der Kinovermarktung durchläuft ein Film einen Lebenszyklus, der in der Regel nur wenige Wochen andauert und erheblichen Einfluss auf nachfolgende Verwertungsstufen hat. Dieser Aufsatz analysiert im Rahmen einer Ereignisstudie, welche frühen Ereignisse im Kinolebenszyklus (z.B. die Veröffentlichung der Umsatzzahlen des ersten Kinowochenendes) die Investoren zu einer Neubewertung ihrer Investition in Medienunternehmen bewegen und so den Unternehmenswert verändern. Hierbei werden die a priori Erwartungen der Investoren an den Umsatzerfolg der Innovation mit berücksichtigt. Die Ergebnisse implizieren, dass Investoren unmittelbar nach dem Start des Filmes ihre Erwartungen anpassen und darauf reagieren. Zudem zeigt sich ein Einfluss der Umsatzzahlen des zweiten Wochenendes auf die Marktkapitalisierung von Medienunternehmen.

E.The Effect of Unexpected Chart Positions on the Firm Value of Music Labels. An Event Study of Album Success.

Nima Mehrafshan, Björn Goerke und Michel Clement

Veröffentlicht als Working Paper:

Mehrafshan, Nima, Björn Goerke und Michel Clement (2016): The Effect of Unexpected Chart Positions on the Firm Value of Music Labels. An Event Study of Album Success., <http://hdl.handle.net/10419/142161>

Abstract

We conduct an event study to (1) analyze whether investors revise their expectations about a music album's success when new chart information is published and (2) estimate how these revised expectations affect the value of a music label. We find that expectations about the success of an album are formed with respect to the performance of the promotional singles and that failure to meet these expectations leads to negative stock returns. However, unexpectedly high chart positions do not lead to significantly higher valuations of labels. The initial album success is anticipated at a very early stage when single charts are released one week prior to the release of the initial album charts.

1 Introduction

In many product categories – ranging from perfumes to computer software – firms offer product samples to support the upcoming market entry of innovations. The samples provide consumers with a product experience that potentially triggers purchases and creates favorable word of mouth for the subsequently released main product (Heiman et al. 2001). Such product sampling strategies are well known in the music industry. Music labels release singles as pre-album samples to promote high-margin albums that are released a few weeks later (Yanbin et al. 2011).

Considering the relevance of innovations for firm value (Sorescu and Spanjol 2008), we argue that the market success of a product sample (e.g., the single) may serve as a valuable information source for investors regarding the upcoming release of the main product (i.e., the full album). Because information about new products is often rare, presentations of samples or prototypes attract the attention of capital markets (e.g., Apple's Macworld presentations). Capital market theory suggests that investors update their expectations regarding new products when relevant new information becomes available. If a new piece of information changes the expectations of investors about the future cash flows of a company, they immediately engage in stock transactions correcting the valuation of the firm and rendering the stock market "efficient" (Fama et al. 1969). An illustrative example from the music industry was the 16% plunge in the share price of EMI after its management announced the delay of two new album releases by Coldplay and Gorillaz to the next fiscal year, because this delay presumably signaled management issues with the bands (Goodway 2005). We argue that the success of a single may form music label investors' expectations with respect to the performance of the album. Especially, *unexpected* sales deviations of sequentially released products (e.g., albums) may cause investors to alter their expectations and therefore reevaluate a record label's stocks. For example, our data show that the surprising second chart rank of Mariah Carey's album "The Emancipation Of Mimi" after her rather disappointing single 'It's Like That' (peaking at chart rank 16) was followed by an *abnormal positive stock return* of her label Universal-Vivendi of approximately 2%.

The objective of this study is to gain insight regarding the relevance of product samples in the product introduction process of music and when and how investors update their expectations with regard to a new product to be introduced. We conduct an event study to analyze how the unexpected billboard chart success or failure of a new album influences firm value. Grounding on a unique sample of music albums we model investors' expectations of the chart success of a new album and compare predicted with observed chart success. We find that expectations about the success of an album are formed with respect to the chart performance of the promotional singles and that a failure to meet these expectations leads to negative stock returns. However, unexpectedly high chart positions do not lead to significantly higher valuations of labels. Our findings further reveal that the initial album success is anticipated by investors at a very early stage using single charts one week prior to the release of the initial album charts.

With this study, we contribute to the field of media economics in different ways. Although the effects of new product (pre-) announcements on firm value have been studied previously (e.g., Hendricks and Singhal 1997; Sorescu et al. 2007), little is known about the effects of product samples on firm value when the samples are introduced to the market prior to the main product. By employing the event study methodology within a time period covering the most crucial phase of the introduction process surrounding the album release, we are able to determine when investors react to new information and thus update their evaluation of the stock and how the reactions of investors move stock prices depending on market expectations. For managers both inside and outside of the music industry, this knowledge is valuable because the strategic management of expectations can lead to higher stock prices and, more importantly, may prevent stock price drops if expectations are not met.

We continue with an overview of related literature and provide our hypotheses. We then discuss our methodological approach, our data and the results of the event study. The paper closes with conclusions and implications of our empirical study.

2 Product Introduction Events

2.1 Product Introductions in the Music Industry

The evaluation of future market success is generally a difficult but crucial task for producers and investors. Consumer feedback regarding product samples may serve as an important indicator of the demand for new experience goods (Chellappa and Shivendu 2005; Hirschman and Holbrook 1982). Most music labels follow a standardized product introduction process to reduce consumers' uncertainty with respect to the quality of the product. Promotional singles are distributed several weeks prior to the release of an album; these singles are partly even free of charge through (Internet) radio airplay or music videos to create media awareness and to enable consumers to evaluate the songs.⁸ Thus, due to the hedonic nature of the product, labels regularly release singles to support the sales of the subsequent album.

The music industry has several advantages for studying the influence of product samples on firm value using event studies. First, information regarding product success is transparent and publicly available in the music industry via chart rankings that are published in short intervals (e.g., by *Billboard* magazine, Bradlow and Fader 2001). Second, there is a large number of product releases each week, and each release has the character of a venture with a high risk of

⁸ Although the single is regaining importance as a product itself due to growing single-download sales, albums are still generating the majority of music sales (Elberse 2010). In 2012, albums worth US\$ 3.86 bn, and singles worth US\$ 1.63 bn were sold in the US (both including downloads, Source: RIAA).

not recovering its initial investment (production and especially marketing) and a slight chance of becoming a big hit and thus covering the losses of many flops (Krasilovsky and Shemel 2007). Third, a few large stock exchange listed companies (major labels), some of which are parts of technology (e.g., Sony Music) or media (e.g., UMG/Vivendi) conglomerates, dominate a large number of small and medium-sized firms (independent labels) that partially distribute their products through the infrastructure of the major labels or other service companies. According to Nielsen Soundscan, the major labels account for about 67% of the value of the U.S. music industry (or even 88% if calculated by distribution ownership; (Billboard 2013). Forth, although the music industry has experienced massive structural changes, the basic mechanisms of product introductions in the music industry, including large marketing investments in the weeks before an album release using promotional singles, have not been affected. Most notably, these changes include the efficient digitalization of music records, which enabled music to be distributed through the internet, but also facilitated illegal reproduction replacing a supposedly substantial proportion of global music sales (Rob and Waldfogel, 2006). As a reaction, music labels introduced so called ‘360 deals’, which include participation in live performances and merchandise, to benefit from a wider range of musicians income streams, partly induced by their marketing investments (Leeds 2007). On the other hand, technological advances reduced the cost of music production, marketing and distribution, leading some artists to abandon major record deals (Graham 2009).

However, rather than reversing economic relationships, these changes have intensified prevalent patterns, such as the superstar phenomenon and the use of (free) samples (Hamlen 1994; Bhattacharjee et al. 2007; Elberse 2008 and 2010). Finally, similar to the movie industry, in which cinematic box office success determines the expected revenues from DVD sales and television licensing (Hennig-Thurau et al. 2007), the chart success of new songs indicates the total monetary success in downstream markets (e.g., the sale of recorded media, concert tickets, merchandising, and music licensing). Thus, album sales are not only an important driver of record label profits but they also serve as an indicator for future revenues. Album releases mark highly important milestones in a musician’s life cycle.

2.2 Prior Research on Product Introduction Events

The actual market entry is a highly sensitive stage in any product’s life cycle. Hendricks and Singhal (1997) analyze the detrimental effects of delaying the release of a product, which illustrates how negative information during this crucial phase may affect firm value. However, there is little prior research on the direct effects of new information regarding the actual process of product *introductions* (as opposed to product introduction announcements) on the fi-

nancial value of firms, especially concerning product samples as primary information sources. Table 1 provides an overview of the most relevant literature on product introductions and firm value. If not stated otherwise, we extracted the average abnormal return (AAR) on the event day (-0,+0) and listed significant moderators with an indication of their direction.

Most prior studies have investigated the announcements or pre-announcements of new products as opposed to the success of introductions of the actual products or product samples. Chaney et al. (1991) were the first to analyze the stock market reactions to new product announcements, and showed that returns to new products are more pronounced for technology firms and for original products as opposed to mere remakes of existing products. They also find that stock returns are larger when detailed information is provided in an announcement.

Sorescu et al. (2007) extend these findings by differentiating short and long term effects, showing that capital markets react more sustainably if information provided in new product preannouncements are reliable and if capital markets are continuously being updated about the progress of the product introduction. Their research suggests that there are several events (preannouncements and follow ups) in the introductory process of new products that lead to reactions by investors. In this context, product samples may be a very reliable source of information because investors and consumers can to some extent examine and test a product prior to its market introduction. However, experimental studies have shown that product samples, as opposed to other consumer promotions, may have negative long-term effects on sales and lead to the cannibalization of the main products (Bawa and Shoemaker 2004). It is thus unclear how the success of product samples influences stock market expectations with respect to the overall firm's performance.⁹

Sood and Tellis (2009) analyzed different types of events in the initiation, development and commercialization phases of innovation processes and found that the total average stock returns for an innovation project are 13 times the average returns for a single innovation event (\$49 m). By contrast, Kelm et al. (1995) find that returns for events in commercialization phases are lower than for events in other phases. This result may have arisen because market expectations toward a project have already been formed by the time the product is launched; thus, it is crucial to incorporate expectations into the analysis to ensure accurate valuations.

⁹ We assume in this paper that a music single serves as a product sample of the music album. However, due to increased unbundling in the digital world (Elberse 2010) it can be argued that singles and albums are not necessarily complements. While this trend is likely to continue, we argue that higher single sales will also lead to publicity for the corresponding album resulting in additional sales of either the full album or unbundled parts of the album.

Joshi and Hanssens (2009) applied an expectation model to address this problem. They analyzed new product introductions in the movie industry and found that unexpected box office successes or failures (difference between expectations and actual revenue and absolute profit) explain abnormal returns after movie releases. We extend their research by focusing on the analysis of (music) product samples and examining how the success of samples are used as

Table 1: Event Studies on New Products and Innovations

Authors	Event	Effect on Firm Value	
		AAR (0,0)	Moderators
Chaney et al. 1991	Product introductions	+ .22%	Technology intensity in the industry (+), degree of innovation (+)
Kelm et al. 1995	Product developments	+ .96% ¹	<i>Pre-announcement</i> : Technology strength (+)
			<i>Announcement</i> : R&D intensity of the industry (-), concentration within industry (+), firm size (u-shaped)
Hendricks and Singhal 1997	Delays of product introductions	-3.25% ²	Competition in industry (-), diversification (+), provision of estimate of delay (+)
Kalaignanam et al. 2007	Product development alliances	not reported (+)	Development alliances ³ : Extension of alliance (+), cooperation experience (+), innovativeness of partner (+)
Sorescu et al. 2007	Announcements of new products	n. s. ¹	<i>Short term</i> ⁴ : Information depth of pre-announcement (+), reliability of information (+); <i>Long term</i> ⁴ : Follow-up information (+), reliability of information (+)
Joshi and Hanssens 2009	Theatrical movie release	+ .42% ⁶	Expectation shock ⁵ (+), movie profit (+), ad intensity (+)
Fosfuri and Giaratana 2009	Rival's new product announcements	not reported (-)	None
Sood and Tellis 2009	8 innovation event type announcements	+ .40%	Firm size
Ransbotham and Mitra 2010	Technology acquisition announcements	-1.26% ⁷	Target age (-), recent patents (+), privately held company (+)

¹ CAR (-1,0)

² The negative AAR results from the announcement of a product's introduction delay

³ Effects are asymmetrically distributed depending on the size of the partner.

⁴ Short term: Daily return; Long term: One-year return

⁵ Difference between estimated and actual open week gross

⁶ Difference between above and below average advertising CAR (-2,+2), i.e. the cumulated AARs starting two days prior to and ending two days after the event

⁷ Difference of AAR between above and below median age target companies

highly relevant informational resources for estimating the success of later released products (i.e., albums). Further, Joshi and Hannsens (2009) focus on the motion picture industry in which expectations are formed in a different way, as there are no full samples provided to consumers (except movie trailers). Joshi and Hannsens (2009) do not include the “quality” or demand of trailers in their expectation model due to data limitations in the motion picture industry. However, in the music industry labels typically release the “best” single prior to album release for free on radio in order to advertise the new album. Thus, the music industry is a well suited industry for studying the influence of product samples on firm value.

3 Hypotheses

A number of sources provide public information regarding the success of music singles and albums. Shortly after the release of a song as a single, the first charts reflect the actual demand of consumers for the product. This occurs several weeks prior to the release of the album. Airplay, DJ, and sales charts are published and provide information about the success of new releases. We argue that, at least for typical pop records, the promotional singles act as product samples for consumers and can thus be used as indicators of the future success of the main products (i.e., the albums). Thus, single charts prior to album release already contain considerably reliable information about probable chart ranks and can influence the expectations of investors. However, only *new* information will lead to adjustments of investors’ evaluations. Investors form expectations with respect to a new release based on market data. Therefore, they will expect a new single of a super star to enter the charts at a high position indicating a bestselling forthcoming album. Consequently, the capital market will only react if the success of the new album is either *unexpectedly* high or low. If the expectations are simply met, capital market theory predicts no changes due to the absence of new information.

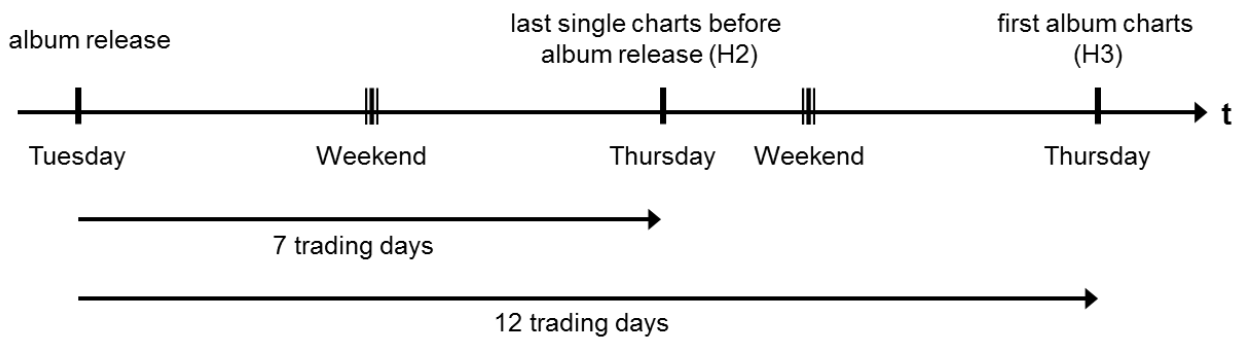
Unexpected rankings: Previous research grounding on prospect theory has shown that investors are usually more sensitive towards negative news due to loss aversion. This sensitivity is sometimes referred to as the negativity bias (Akhtar 2012; Kahneman and Tversky 1979; Sood and Tellis 2009; Tellis and Johnson 2007). Therefore, we expect stronger capital market reactions for flops than for hits. We use “hits” and “flops” to refer to unexpectedly high-/low-ranking records. According to this definition, a flop may be a profitable title, but may not meet the (estimated) expectations. Analogously, a hit may be a title that did not rank very highly in the charts, but has exceeded its (low) expectations (see method section). Referring to the theoretical framework of prospect theory, we assume that stocks of labels facing unex-

pectedly low-ranked albums will suffer more than stocks associated with unexpectedly high-ranked albums will gain, leading to hypotheses 1:

H1: The magnitude of negative cumulative abnormal returns will be higher for unexpectedly low-ranked albums compared to the positive cumulative abnormal returns of unexpectedly high-ranked albums.

Timing: We formulate two hypotheses regarding the timing of new chart information. In the U.S., singles and albums are typically released on Tuesdays, and the corresponding first week charts are published on Thursdays of the week after the following week (i.e., 16 days later).¹⁰ Figure 1 shows the context of album release, i.e. relevant single chart, and album chart publications.

Fig. 1: Timeline of relevant charts releases



Higher initial chart rankings (i.e., strong demand of early adopters) correspond to higher levels of acceptance of the song by followers (Strobl and Tucker 2000). Furthermore, high chart positions reflect higher demand which leads to herding behavior of consumers (Salganik, Dodds, and Watts 2006). Therefore, labels attempt to optimize the initial chart ranks of their artists using advertising and promotion measures (e.g., by encouraging airplay). Labels thus rely on daily trend chart information provided by market research services (e.g., Nielsen SoundScan), which approximates the official chart positions. Official chart movements are also available to investors of music labels and can be used to estimate future cash flows from an album and related products. Album success may also affect the artist's brand value and thereby the cash flow of their future releases. Thus, if singles act as product samples and the album is a milestone in the process of artist brand value building, the ranking of a single may change the expectations about the performance of the album, which will be reflected in share prices in an efficient market. Hence, a rational investor will use the most recent information about a product sample (i.e., single charts) to gauge the future success of the main product.

¹⁰ The week after the release week is referred to as the *first chart week*.

We expect that an album's sales expectations are updated one week before the release of the album charts, when the *last single chart ranks* (prior to the first album charts) publicly indicate the potential success of the album. Hence, we hypothesize:

H2: Positive (negative) abnormal returns will be observed on the Thursday on which the last single charts are made available prior to the album release for unexpectedly high- (low-) ranked albums.

The expectations for an album increase with the performance of a single release. Of course, there can still be surprising discrepancies between single and album chart performance that will cause investors to re-evaluate the stock. If these discrepancies are sufficiently high to change the expectations of the investors regarding the financial contribution of the album, the re-evaluation should instantly be reflected in the share price. Therefore, we also expect to observe abnormal returns for the *first album charts* (week 3) if the chart rank of the album does not correspond to the investors' expectations that are formed based on the single chart performance and other album characteristics.

H3a: Positive (negative) abnormal returns will be observed on the Thursday on which the first album charts are made available for unexpectedly high- (low-) ranked albums.

However, because the majority of the expectation revisions have already been made after the release of the single charts (week 2), the ARs surrounding the album chart release will be smaller.

H3b: The magnitude of these abnormal returns will be smaller than those of the single charts released a week earlier (H2).

4 Expectation model

4.1 Relevance of Expectations in Event Studies

According to the theory of efficient capital markets, new information will lead to an immediate reaction in stock prices, reflecting changes in market expectations (Fama et al. 1969). We use the well-established event study methodology to obtain firm-specific returns (abnormal returns or ARs) related to events (new chart information). Because the expected positive and negative returns would simply mutually be cancelled out when calculating average abnormal returns (AARs), we divided our sample into a group of albums that outperformed (hits) and a group that underperformed (flops) in comparison with the *expectations* of the market. According to the theory of efficient capital markets, these expectations are based on all relevant mar-

ket information available to investors, including the past successes of the artists, the marketing abilities of the labels and, most importantly, the performance of product samples (i.e., single releases). We model these expectations ex post using an econometric model of album chart ranks. Predictions of this model approximate market expectations of chart ranks based on the included variables. Positive or negative differences between predictions and actual chart positions indicate the extent to which an actual chart rank is a positive or negative surprise with respect to prior expectations.

The event study methodology quantifies the effects these surprises have on the firm value of the stock listed labels, by examining the statistical significance of daily abnormal stock returns averaged over as many occurrences of the respective event (e.g., chart releases) as can be obtained. Other influences not correlated with the event, which may affect abnormal returns, represent ‘noise’ in the statistical sense and are mutually cancelled out. To further reduce this noise, it is common practice to exclude observations where so called ‘confounding events’ occur, i.e., new information unrelated to the event under study, but likely to affect abnormal returns. It is thereby possible to include stocks of conglomerates (such as Sony Music/Sony BMG, Universal Music Group/Vivendi) in the analysis that generate only a fraction of their cash flows with music.

4.2 Data

In order to estimate the expectation model, we combine chart data with individual information regarding 853 albums in the sample collected from multiple sources (RIAA [riaa.com], allmusic.com, grammy.com). We use Billboard Top 200 album chart ranks, Billboard Hot 100 single chart ranks from January 2004 to February 2006, which are derived from the physical (i.e., CDs, Vinyl records, etc.) and as of February 2005 also digital (e.g., iTunes and Napster downloads) album and single sales, respectively.

To compare the ARs of unexpectedly under- and outperforming albums, we estimate a model to reconstruct the expectations of the capital markets regarding potential chart success prior to the release of the album chart. The dependent variable is the first Billboard Top 200 album chart rank after the release, which we have coded as an ordinal variable with five chart categories ($Charts_i^{album}$, where i denotes the album, see table 2) to account for the non-linear relationship between sales and chart ranks (Chevalier and Goolsbee 2003). Higher chart ranks represent disproportionately higher sales per release than lower ranks due to the specific supply and demand characteristics of artistic markets; this situation is generally referred to as the superstar phenomenon (Adler 1985; Giles 2006; Hamlen 1994; Rosen 1981). This situation

may be problematic when sales charts are used directly as a measure of financial success in this study setting. By relying on these charts, we (and likewise the capital markets) have limited information on the exact levels of sales and only know their order. Therefore, we cannot derive the functional relationship between chart ranks and sales. Thus, we account for the underlying skewed distribution of record sales by categorizing the chart positions. This categorization also accounts for possible jumps or breaches between different chart slots. For example, climbing one chart rank and reaching a higher slot (e.g., rank 11 to 10) may have a greater effect than simply a one-rank increase (e.g., rank 8 to 7) because the release becomes available in media and retail stores that cover or promote the higher slot (e.g., radio top 10 countdowns), and this promotion and coverage may lead to higher product awareness and availability (Connolly et al. 2006). Hence, we formed our chart categories to represent common chart rankings in the music industry (e.g., top 5, top 10, top 20, top 40 and top 200 for album charts and, similarly, the top 10, top 40, and top 100 for singles¹¹). This scale also solves the problem of how to treat unranked titles: Releases that are not ranked form the (lowest) base categories of the ordinal charts variables.

Table 2: Coding of chart variables

Variable	Highest album chart rank			Highest single chart rank		
	$Charts_i^{album}$			$Charts_{ij}^{single}$		
Categories	Top 200 album	k	n_k	Hot 100 Singles	j	n_j
	Ranks 1-5	5	187	Ranks 1-10	3	35
	Ranks 6-10	4	80			
	Ranks 11-20	3	125	Ranks 11-40	2	77
	Ranks 21-40	2	150			
	Ranks 41-200	1	150	Ranks 41-100	1	143
	No Rank	base	161	No Rank	base	598
		Σ	853			Σ 853

To predict album chart ranks, we use single chart information and control for a range of artist- and album-related measures (see table 3 for descriptive statistics and the correlation matrix of the data): The independent variables $Charts_{ij}^{single}$ take the value 1 for the respective peak chart rank category j of the highest-ranking single in the Billboard Hot 100 Singles from the con-

¹¹ This configuration is modeled along the typical chart categories used by Billboard. It resembles the thresholds the analysts are confronted with when assessing chart success of either singles or albums.

sidered album l prior to the release of the album chart (table 2).¹² The variable *Amazonstars* (mean=4.1, s. d.=.57) represents the mean of the customer ratings of an album on Amazon.com as a measure of perceived product quality. We also included a squared *Amazonstars* variable (mean=16.7, s. d.=4.1) to account for non-linear quality effects.¹³ $\ln(\text{Prior_albums})$ (mean=1.4, s. d.=1.3) is the natural log of the number of albums an artist formation has released prior to the release of the respective album. We include two variables that measure prior success and star power as the performance of an established band might be of more interest than of a band that barely entered the Billboard Top 200 once: $\ln(\text{Gold_platinum})$ (mean=-14.0, s. d.=14.8) is the natural log of the number of gold and platinum records a formation has been awarded prior to an album release (platinum records receive a double weight because they represent sales of 1 m records as opposed to 500 k for gold records).¹⁴ *Grammys* (mean=.7, s. d.=2.0) is the number of Grammy awards a formation has received prior to the release of an album. *Pop* (31%), *Rock* (28%), and *Rap* (15%) are dummy variables reflecting the genre of an album; these variables take the value 1 for the respective genre (pop, rock/heavy metal, and rap/hip-hop) and 0 otherwise (26%). We control for the distribution power of major labels by including dummy variables for the major labels. *Warner* (18%), *EMI* (11%), and *Conglomerate* (55%) are dummy variables taking the value 1 if an album has been released by Warner Music/WMG, EMI Music, or by one of the two major labels that were owned by conglomerates during the study (Sony Music/Sony BMG, Universal Music Group/Vivendi).¹⁵ The base category are independent labels (16%). Finally, *Foreign_country* (13%) takes the value 1 if an artist or formation is not from an English-speaking country (U.S., Canada, U.K., Australia) and 0 otherwise.

¹² Since in the sample only 23 albums had two singles and only two albums had three singles released prior to the album release, we only included the best performing single peak rank as an explanatory variable.

¹³ We note that the relation of chart positions and the ranking of an album at Amazon might be subject to endogeneity. However, the primary model objective is a prediction task (and not obtaining consistent parameter estimates). Hence, we follow Ebbes, Papies and van Heerde (2011) who recommend refraining from correcting for endogeneity with IV estimation when prediction is the main modeling objective.

¹⁴ We used the natural logs of *Prior_albums* and *Gold_platinum* to account for the diminishing returns that we found to be prevalent in the data. With an increasing number of albums (awards), the amount to which each additional album (award) contributes to the probability of higher or lower chart ranks is assumed to decrease.

¹⁵ We have combined Sony Music and UMG into one variable, because according to a Chow test, coefficients on separate dummy variables have shown to be indifferent from one another in the model ($p=.77$). This combination is plausible because albums from these conglomerates seem to have similar characteristics with respect to chart success, whereas the more specialized and less integrated major labels Warner and EMI have distinctive properties.

Table 3: Descriptive statistics and correlation matrix of expectation model data

Variables	N	Mean	S. d.	Correlations												
				1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1. Top 200 album	692	34.025	46.338	1												
2. Hot 100 singles	255	48.267	29.149	.217	1											
3. Amazonstars	853	4.067	.550	.119	.264	1										
4. Prior_albums	853	8.461	9.682	-.063	.006	-.014	1									
5. Gold	853	3.566	5.605	-.086	.051	.000	.361	1								
6. Platinum	853	2.882	6.549	-.048	.012	.002	.255	.812	1							
7. Grammys	853	.686	1.697	-.117	-.003	-.022	.373	.353	.303	1						
8. Pop	853	.309	.463	-.001	-.158	.012	.070	.161	.052	.137	1					
9. Rock	853	.285	.357	.051	.140	.039	.050	-.029	-.002	-.021	-.423	1				
10. Rap	853	.149	.460	-.109	-.100	-.233	-.132	-.098	-.068	-.079	-.280	-.264	1			
11. Warner	853	.179	.375	-.039	-.008	.028	-.046	-.051	-.036	.029	-.075	.125	-.024	1		
12. EMI	853	.107	.298	.087	.124	-.034	.079	.108	.030	.042	.122	-.041	-.059	-.162	1	
13. Conglomerate	853	.549	.466	-.106	-.089	-.029	-.005	.075	.099	.030	.042	-.117	.088	-.515	-.381	1
14. Foreign_country	853	.131	.228	.127	-.046	.082	-.019	-.095	-.106	-.105	.055	.008	-.153	-.019	.068	-.136

4.3 Model estimation

To estimate the expectation model, we use ordinal regression with a complementary log-log link function from the class of generalized linear models (GLM) because the largest category of the dependent variable is the highest rank category in our data:

$$\begin{aligned} \eta_{lk} = \log(-\log(1 - \gamma_{lk})) = \\ \theta_k - [\sum_{j=1}^3 \beta_j (Charts_{ij}^{single}) + \beta_4 (Amazonstars_i) + \beta_5 (Amazonstars_i^2) + \beta_6 \ln(Prior_albums_i) + \\ + \beta_7 \ln(Gold_platinum_i) + \beta_8 (Grammys_i) + \beta_9 (Pop_i) + \beta_{10} (Rock_i) + \beta_{11} (Rap_i) + \beta_{12} (Warner_i) + \\ + \beta_{13} (EMI_i) + \beta_{14} (Conglomerate_i) + \beta_{15} (Foreign_country_i)] \end{aligned} \quad (1)$$

with $\gamma_{lk} = \text{Prob}(Charts_i^{album} \leq k | \mathbf{x}_i)$ for $k = 1, \dots, 5$

Table 4 shows the estimation results.

According to the LR test, the joint model is highly significant, and the pseudo R^2 values suggest a reasonably good model fit. Multicollinearity among independent variables does not seem to be a severe problem in modeling the expectations. The correlation between *Amazonstars* and *Amazonstars*² is naturally high ($r=.975$, $VIF=21.4$ and 21.1). However, since both coefficients are significant, collinearity is not detrimental. Apart from that, the highest correlations are between $\ln(Gold_platinum)$ and $\ln(Prior_albums)$ ($r=.441$) and between $\ln(Gold_platinum)$ and *Grammys* ($r=.306$). However, apart from the quality variables (*Amazonstars* and *Amazonstars*²) there are no VIFs higher than 1.60. Finally, we conducted an out-of-sample hold-out prediction test to verify that the model serves as a reasonable approximation of market expectations. We re-estimated the model using a sub-sample of 700 randomly selected observations. Using this model, we predicted the chart categories for the 153 observations of the hold-out sample. The model coefficients did not significantly differ from our final model, which utilizes all available observations. The out-of-sample hit rate of the predicted categories was 42.5%.

Table 4: Ordinal regression results of the dependent variable $Charts^{album}$ (complementary log-log link function)

Independent variables	Parameter estimate	Standard error	Wald coefficient
$Charts_{Top41-100}^{single}$.758***	.122	38.382
$Charts_{Top11-40}^{single}$	1.129***	.185	37.218
$Charts_{Top1-10}^{single}$	1.374***	.304	20.408
<i>Amazonstars</i>	.936**	.303	9.564
<i>Amazonstars-squared</i>	-.182***	.043	18.142
$\ln(Prior_albums)$	-.122**	.039	9.916
$\ln(Gold_platinum)$.018***	.003	31.314
<i>Grammys</i>	.104***	.027	15.210
<i>Pop</i>	.128	.108	1.411
<i>Rock</i>	.145	.106	1.858
<i>Rap</i>	.429**	.149	8.258
<i>Warner</i>	.556***	.133	17.393
<i>EMI</i>	.236	.150	2.470
<i>Conglomerate</i>	.920***	.114	65.276
<i>Foreign_country</i>	-.977***	.120	66.337
** p < .01, *** p < .001 (n=853)			
LRT Chi-square (15)			562.4 (p<.001)
Pearson Chi-square (3,825)			4,402.0 (p<.001)
Cox and Snell / McFadden R ²			.483 / .187

For a better understanding of the influence the single chart rank has on the expectations with regards to the subsequent album chart rank, we conduct an analysis of the marginal effects of the single chart rank. We use the expectation model to predict the probabilities for each outcome of $Charts^{album}$ for each level of $Charts^{single}$ (holding all other variables at their means). The following results in table 5 hold for an average album in the dataset.

Table 5: Predicted probabilities for each category of $Charts^{album}$ by levels of $Charts^{single}$

			$Charts^{single}$			
k			No Rank	Ranks 41-100	Ranks 11-40	Ranks 1-10
$Charts^{album}$	5	Ranks 1-5	.141	.329	.435	.503
	4	Ranks 6-10	.085	.106	.103	.098
	3	Ranks 11-20	.148	.148	.134	.122
	2	Ranks 21-40	.190	.154	.130	.114
	1	Ranks 41-200	.211	.142	.111	.093
	base	No Rank	.226	.122	.087	.070

Calculating the album chart slot expectation for each level of $Charts^{single}$ through:

$$E(Charts_l^{album}) = \sum_{k=0}^5 [\hat{\gamma}_{lk} \cdot Charts_l^{album}] \quad (2)$$

results in table 6:

Table 6: Marginal effects of $Charts^{single}$ on $Charts^{album}$

	$Charts^{single}$			
	No Rank	Ranks 41-100	Ranks 11-40	Ranks 1-10
$E(Charts_l^{album})$	2.078	2.962	3.359	3.594
Marginal effect	-	.884	.397	.235

An average single peaking in ‘ranks 41-100’ compared to ‘no Top 100 ranking’ increases the expected album chart slot from 21-40 ($E(Charts_l^{album})=2.1$) to 11-20 ($E(Charts_l^{album})=3.0$). Going from single chart ranks 41-100 to ranks 11-40 does not result in an increase to a higher album chart slot. Moving from single chart ranks 11-40 to Top 10 increases the expected album chart slot from 11-20 to 6-10.

4.4 Discussion

All coefficients have the expected signs. The coefficients of the dummies of the single charts increase with each category, but the increases diminish. Thus, the probability for a strong album chart rank is higher for albums with chart-listed singles. The probability increases if the single chart rank is higher (categories 2 and 3), but the increase is less than the probability increase associated with being listed in the single charts, at all. We find an inverted U-shaped non-linear relationship for *Amazonstars*, suggesting higher rankings for albums which have neither very poor nor exceptionally good average quality ratings. This finding might be due to possibly higher variance in the reviews. Sun (2012) reports that niche products that some consumers love and others hate are often associated with a high variance of ratings. Based on the informational

content of reviews, a higher variance of reviews *Amazonstars* may in turn lead to a higher subsequent demand, because it spurs curiosity and may lead to customer and media controversy.

The star power and quality-related variables $\ln(\text{Gold_platinum})$ and *Grammys* capture the effect of prior successes and have a strong positive effect. The negative sign of $\ln(\text{Prior_albums})$ may indicate that (holding prior success constant) younger (more trendy) musicians have higher average chart rankings than older bands that may already have completed the growth stage of their ‘artist life cycles’.¹⁶ Moreover, domestic rap albums from major labels have the highest probability for high album chart entries. With the exception of EMI, we find a significant impact of the major label’s dummy variables *Warner* and *Conglomerate* (Sony and Universal Music) on the album’s chart success. Finally, we find a significant disadvantage with respect to the album’s chart success if the album is not from an English-speaking country.

The model results are used to calculate the deviations of the expected from the observed chart position and to assign unexpected hits and unexpected flops to the respective group for the event study. If the prediction error is greater than zero ($\hat{\varepsilon}_l > 0$), which indicates that the true or realized chart category is higher than expected, the album is allocated to the group of unexpected hits, and vice versa for unexpected flops. We define the prediction error as follows:

$$\hat{\varepsilon}_l = \underbrace{\text{Charts}_l^{\text{album}}}_{\text{actual category}} - \underbrace{\sum_{k=0}^5 [\hat{\gamma}_{lk} \cdot \text{Charts}_l^{\text{album}}]}_{\text{category expectation}}. \quad (3)$$

5 Event Study

5.1 Event study methodology

In most event studies, the market model, which is a simple regression of the return of a firm on a market index return, is used to predict abnormal returns (McWilliams and Siegel 1997):

$$AR_{it} = \hat{\varepsilon}_{it} = R_{it} - \hat{\alpha}_i - \hat{\beta}_i r_{Mt} \quad (4)$$

- R_{it} is the actual return of stock i on trading day t
- $\hat{\alpha}_i$ is the autonomous return of stock i
- $\hat{\beta}_i$ is the systematic (market) risk factor of stock i
- r_{Mt} is the return of the market portfolio M on trading day t .

¹⁶ In our dataset, the 395 albums of bands that have released more than 4 albums (median) have a mean initial chart category of 2.3 compared with 2.6 for the 458 albums that have released 4 or fewer albums ($p < .001$).

To draw inferences regarding the hypotheses, the abnormal returns (AR) must be aggregated over all firms in the sample to test their statistical significance. The simplest method is to consider average abnormal returns (AAR) and conduct ordinary cross-sectional t-tests.

$$AAR_t = \frac{1}{I} \sum_{i=1}^I AR_{it} \quad (5)$$

The t-statistics can be obtained following Boehmer et al. (1991):

$$T_1 = \frac{AAR_t}{\sqrt{\frac{s^2}{I-1}}} \quad (6)$$

$$\text{with } s^2 = \frac{1}{I} \sum_{i=1}^I (AR_{it} - AAR_t)^2$$

To unveil the effects that manifest over several trading days, the AARs are cumulated over the event window (L_2) to obtain cumulated average abnormal returns (CAR).

$$CAR_{t_2, t_3} = \sum_{t=t_2}^{t_3} AAR_t \quad (7)$$

The respective t-distributed test statistic is as follows (Hendricks and Singhal 1997):

$$T_2 = \frac{CAR_{t_2, t_3}}{\sqrt{\sum_{t=t_2}^{t_3} S_{AAR_t}^2}} \quad (8)$$

To control for heteroscedastic ARs on the event day and to prevent highly volatile stocks from dominating the tests, standardized abnormal returns (SAR) are calculated following the method by Dodd and Warner (1983):

$$SAR_{it} = \frac{AR_{it}}{SD_{it}} \quad (9)$$

SD_{it} is the estimated forecast standard deviation of the abnormal returns.

$$SD_{it} = \sqrt{\left(s_{it}^2 \left(1 + \frac{1}{L_1} \cdot \frac{(r_{Mt} - \bar{r}_{Mt})^2}{\sum_{\tau=t_0}^{t_1} (r_{M\tau} - \bar{r}_{Mt_1})^2} \right) \right)} \quad (10)$$

with
$$s_{it}^2 = \frac{1}{L_1} \sum_{\tau=t_0}^{t_1} (AR_{i\tau})^2$$

The accumulation of the SARs to obtain standardized average abnormal returns (SAAR) is analogous to equation (4):

$$SAAR_t = \frac{1}{I} \sum_{i=1}^I SAR_{it} \quad (11)$$

The corresponding test statistic is formed with the now constant standard deviation $[(L_1 - 2) / (L_1 - 4)]^{0.5}$ of the SAR:

$$Z = \frac{SAAR_t}{[(L_1 - 2) / (L_1 - 4)]^{0.5}} \cdot \sqrt{I} \quad (12)$$

This test statistic follows a t-distribution, as well (Dodd and Warner 1983).

Both significance tests require that the ARs are not correlated. The test statistic for the SAARs underlies the implicit assumption that the event-induced variance is insignificant.

5.2 Data

We use the event study to control for economy- or industry-wide price movements to obtain firm-specific returns. To calibrate the AR model, we choose a prediction window of 250 trading days (L_1) ending 20 trading days before the event (the day of the initial album chart release). By defining the event window (L_2) from seven trading days before the event to three trading days after the event, we are able to capture all effects that occur from the last pre-album single chart release to the album chart release. We chose the DJ Stoxx Media 1800 as the market index and collected share prices and index return data from *Datastream*.

Since stock returns can only be retrieved for stock exchange listed companies only those 712 albums from major labels could be considered in the event study. Because record labels release multiple albums on one day, the sign of the errors from the expectations model may differ across the albums released by one company on one Tuesday. In these cases and in case the event windows overlap, the abnormal returns of the respective music labels have been eliminated from the sample, which decreases the sample size. Moreover, confounding events have been eliminated from the sample for each individual trading day. The following types of confounding events have

been searched for: (1) financial results (e.g., company reports, earnings warnings), (2) co-operations and significant contracts, (3) changes in management, (4) changes in analyst recommendations, (5) mergers and acquisitions activity, (6) other (e.g., purchase of licenses, other contracts, investigations against the company, court decisions, restructuring announcements), (7) financing measures (e.g., issuance of shares or bonds), and (8) outliers (absolute returns that are larger than 3 times their standard deviation).¹⁷ Therefore, the sample sizes vary for each day, ranging between 145 and 159 albums.

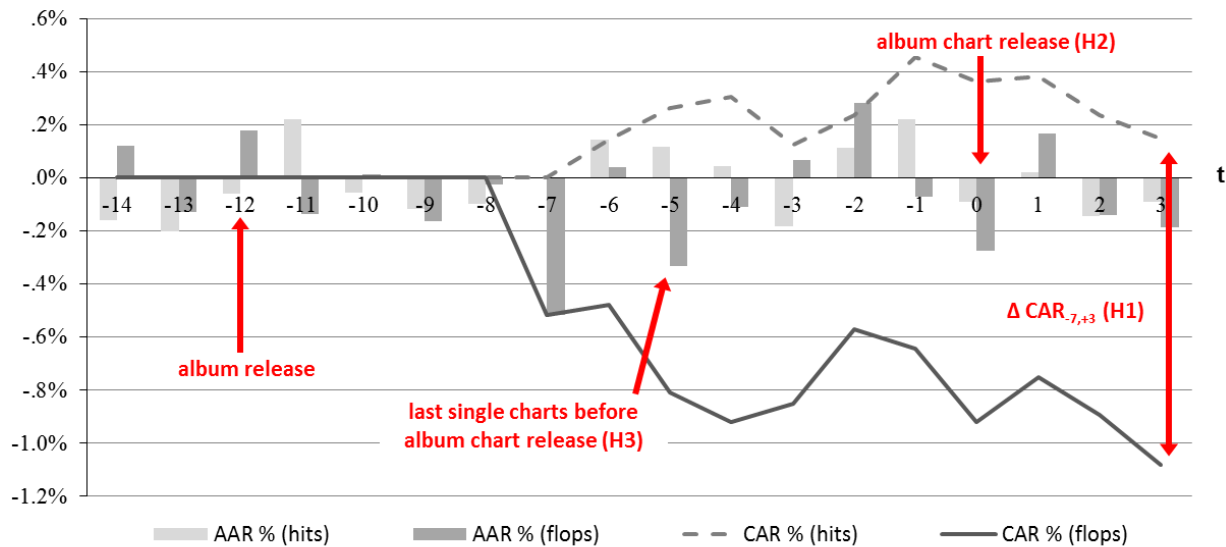
5.3 Results

Table 4 shows the event study results within the two groups of the unexpected hits and flops for the period beginning two trading weeks prior to the album chart release and ending three trading days after the release of the album chart.

Figure 2 presents a visual representation of the results. We observe that the cumulative average abnormal returns (CARs) within the flops group are (with the exception of trading day -2) significantly negative beginning seven trading days before the album chart release. Within the event window the CARs (-7; +3) for the flops group accrue to 1.08% ($p < .05$). The CARs are significant (at least at the 10% significance level) throughout this period, with the exception of the Tuesday of the first chart week, where a positive AAR of .28% ($p < .10$) may represent a slight reversal effect. On the contrary, during the event period the CARs of the hits group are all insignificantly different from zero. The differences (Δ CAR) between the CARs of the hits and flops groups are statistically significant over the whole event window ($p < .10$), with two exceptions: the Tuesday of the first chart week (-2) and the Monday after the album release (2). The t-values of the respective tests are lower due to the high standard errors of the insignificant CARs of the hits group.¹⁸ These differences peak on the Thursday of the third week when the album charts are released (0).

¹⁷ Sources used for confounding events identification: IR news on the corporate websites of major labels and press releases from a broad range of publications via the press database LexisNexis; for U.S.-listed companies: releases of SEC forms 10-K (annual report), 10-Q (quarterly report), and 8-K (current report filing) from the EDGAR database (www.sec.gov/edgar). Album announcements themselves were not eliminated as they are not identifiable. Information on upcoming albums usually diffuse into the market over time (as opposed to dedicated product announcements in the technology sector). Any piece of information adds marginally to the set of information for the investors. This also applies to single releases.

¹⁸ We used t-tests assuming unequal variances and sample sizes for the difference tests. We used the Welch–Satterthwaite equation for calculating the degrees of freedom.

Figure 2: AARs and CARs over the event period (t =trading days)

Thus our data supports H1 of asymmetric reactions of investors toward positive and negative information, as suggested by prospect theory and previously observed for movies (Kahneman and Tversky 1979; Joshi and Hanssens 2008), because we observe significant negative effects for unexpectedly low chart ranks, whereas unexpectedly high chart ranks are not rewarded with significant positive AARs.

During the week of the album release (week 1) there is only one significantly positive AAR of .23% ($p < .10$) in the hits group on the Wednesday one day after the album release (-11). Other than that there are no significant AARs in the hits group. We find a significant negative AAR of -.33% ($p < .05$) in the group of underperforming albums one week after the album has been released (week 2) on the day of the single charts release (-5) providing support for the “negative side” of H2. Interestingly, however, the largest one day AAR of -.52% ($p < .001$) can be observed two days earlier (-7). As a robustness test we also calculated standardized average abnormal returns (SAAR) that account for heteroscedastic ARs and conducted non-parametric sign tests (as indicated by the N+/N- columns in table 7), which substantially confirmed the results obtained by ordinary tests of AARs. This result is further evidence for the “negative side” of H2 and indicates that negative ARs will be observed at (or near) the Thursday on which the last single charts are made available prior to the album release, though only for unexpectedly low ranked albums. Interestingly, the market seems to partly anticipate the new information provided by the single charts, which suggests insider information that may be based on trend charts provided by market research companies.

In week 3, on the day of the album chart release (0), the average abnormal return for the flops group is $-.28\%$ ($p < .05$), providing support for the hypothesis that unexpected album charts provide new information leading to abnormal returns (H3a). However, the cumulative abnormal returns (the sum of the AARs over the respective days) for the single chart release week ($CAR(-7, -3) = -.85$, $S.E. = .10$) is significantly larger in magnitude than those of the album chart week ($CAR(-2, +2) = -.04$, $S.E. = .09$, $diff. = -.81$, $p < .10$); this result is in accordance with the hypothesis that most of the information regarding the success of the album is already included in the single charts in week 2 (H3b).¹⁹

Thus, the results suggest that there are two main points in time at which relevant information becomes available. In week 2, single charts indicate the success of product samples and provide the basis for estimating the success of albums (H3a). The effect of the album charts is much lower; this result suggests that a product sample has provided most of the information needed for updating expectations on product introduction success (H3b).

¹⁹ As a standard procedure we also conducted an event window sensitivity analysis. Therefore, we test three alternative event windows with $-1;+1$, $-2;+2$, and $-3;+3$ trading days around the event (0). The CAARs for all these event windows are insignificant. Table 8 shows the results in detail. These findings further support the hypotheses.

Further, we regressed the $CAR_i(-7,3)$ on the residuals from the expectation model in order to test whether we find a linear relationship. However, the result is an insignificant model ($F=0.03$) and an insignificant coefficient of the expectation error ($t=-0.18$).

Table 7: Event study results

		Unexpectedly high chart ranks						Unexpectedly low chart ranks						Difference	
t	Week day	AAR %	S. E.	CAR %	S. E.	N	N+	AAR %	S. E.	CAR %	S. E.	N	N-	ΔCAR%	S. E.
Week 1: Release of the album (Tuesday)															
-14	Friday	-.153	(.210)			43	18	.117	(.197)			60	32		
-13	Monday	-.189	(.222)			45	18	-.135	(.151)			59	35*		
-12	Tuesday	-.050	(.178)			45	23	.172	(.175)			56	23		
-11	Wednesday	.234*	(.178)			43	26*	-.154	(.200)			55	33*		
-10	Thursday	-.054	(.210)			43	21	.014	(.192)			55	26		
-9	Friday	-.108	(.189)			44	18	-.166	(.163)			57	31		
Week 2: Release of last single charts before album charts (Thursday)															
-8	Monday	-.098	(.166)			41	20	-.025	(.166)			60	30		
-7	Tuesday	.000	(.233)	.000	(.233)	46	24	-.516***	(.186)	-.516***	(.186)	58	36**	.516**	(.298)
-6	Wednesday	.145	(.212)	.145	(.317)	45	23	.038	(.196)	-.478**	(.269)	59	29	.622*	(.416)
-5	Thursday	.116	(.177)	.261	(.376)	41	21	-.332**	(.180)	-.809***	(.321)	60	32	1.071**	(.495)
-4	Friday	.043	(.237)	.304	(.430)	45	23	-.109	(.193)	-.919***	(.377)	59	32	1.223**	(.572)
-3	Monday	-.181	(.162)	.123	(.464)	44	17	.067	(.182)	-.852**	(.422)	58	30	.975*	(.627)
Week 3: Release of first album charts (Thursday)															
-2	Tuesday	.111	(.220)	.234	(.519)	43	22	.282*	(.191)	-.570	(.450)	62	31	.804	(.687)
-1	Wednesday	.220	(.247)	.454	(.598)	39	19	-.073	(.176)	-.643*	(.494)	59	28	1.097*	(.776)
0	Thursday	-.090	(.215)	.364	(.608)	43	24	-.276**	(.158)	-.919**	(.527)	57	37**	1.283*	(.805)
1	Friday	.019	(.193)	.383	(.625)	45	23	.168	(.189)	-.751*	(.547)	60	31	1.134*	(.831)
2	Monday	-.145	(.212)	.238	(.654)	46	18	-.141	(.152)	-.892*	(.591)	55	32	1.129	(.881)
3	Tuesday	-.092	(.182)	.146	(.693)	44	21	-.188	(.155)	-1.080**	(.596)	58	30	1.226*	(.914)

* p < .10, ** p < .05, *** p < .01, one-sided tests

t = Number of trading days counted from album release

6 Conclusion

We extend prior research on how new information regarding product introductions affects firm value by focusing on product samples as primary information sources. In the music industry, singles serve as a major promotional tool to initiate the demand for later released music albums. We have reconstructed capital market expectations regarding the chart success of the albums using an expectation model that incorporated the performance of promotional singles and analyzed the effects of deviations from these expectations on the stock valuation of the labels. Our hypotheses were largely supported by the data. In particular, we found that labels that introduce outperforming albums gain considerable firm value over companies that introduce underperforming albums. This effect is significant for a period of eleven trading days (with one exception). As hypothesized, we find that the magnitude of the abnormal returns is greater for negative expectation revisions than for positive expectation revisions (H1). We only found significant abnormal returns during the chart release weeks (weeks 2 and 3) for the group of disappointing albums; this result clearly supports this hypothesis.

With respect to the timing of the effects, we find that labels introducing underperforming albums significantly lose firm value near the days of the chart release. Most of these negative abnormal returns are observed in the week before the album charts are released when the single charts are made public (H2). The loss of firm value on the day of the release of the album charts is smaller (H3).

These findings help us to understand the process of firm value generation during the introduction phase of new products. Our results provide specific insight regarding the role of product samples, which clearly formed the basis for expectations with regard to the success of the main product. Firm value will be affected only when the main product substantially fails to meet the expectations that were formed in the market when the success of the product sample was evaluated. In accordance with prospect theory, investors in the capital market tend to react more strongly to negative news, i.e., when the main product (the album) misses the expectations built by the product sample (the single), as opposed to main products that exceed expectations.

This study has important implications for managing the introduction process of new products. Companies may be tempted to raise expectations before a product is introduced to the market. If the product sample is successful, product awareness is raised and creates publicity. However, there is a severe risk that a company may fail to meet these expectations. Our event study

provided evidence that product introductions that fail to confirm expectations may lead to a considerable decrease in firm value. Therefore, new products should meet the expectations that were formed on the basis of promotional tools. Product samples can be used to decrease perceived consumption risks and preference uncertainty but should be chosen as sustainable representations of the main products.

Music managers should be aware of the fact that their actions are under the scrutiny of the capital market and that they influence market expectations of album success by manipulating chart success of respective promotional singles. Our results show that disappointments of expectations with album success, which were raised by single success may decrease firm value or at least create stock price volatility due to the alternating raises and falls of capital market expectations. For value orientated managers two strategies to avoid stock price volatility and firm value destruction come to mind. First, to avoid disappointments with album success, managers should concentrate marketing investments on albums that had the most successful singles prior to album release. And second, they should avoid investing too much of their marketing budget in promotional singles that are extracted from albums, which they do not expect to match the expectations raised by the promoted singles.

However, these considerations have to be made, of course, in light of expected cash flows from single and album sales, as well as effects on the brands of the artists. If a large fraction of the prospected future profits of an artist are expected to be generated by single record sales, management of album success expectations should not be the main focus. If, on the other hand, an artist is expected to make much of their money with album sales and other subsequent products the label participates in (e. g., royalites, live performances, merchandise), managing capital market expectations may be vital. Our results prompt further research questions. As noted above, product trials are particularly important for experience goods, such as music, to enable consumers to “experience” and evaluate products before making a purchase. Moreover, information regarding the demand for sample products is readily available in the music industry via single chart rankings. Thus, further studies in the field of media economics should ascertain whether the results are generalizable to other experience or non-experience goods and identify additional indicators of product sample demand, such as downloads of movie trailers or software trials.

References

- Adler, M. (1985).** Stardom and Talent. *American Economic Review*, 75 (1), 208-212.
- Akhtar, S., Faff, R., Oliver, B. & Subrahmanyam, A. (2012),** Stock Salience and the Asymmetric Market Effect of Consumer Sentiment News. *Journal of Banking & Finance*, 36, 3290-3301.
- Bawa, K., & Shoemaker, R. (2004).** The Effects of Free Sample Promotions on Incremental Brand Sales. *Marketing Science* 23 (3), 345-63.
- Bhattacharjee, S., Gopal, R.D., Lertwachara, K., Marsden, J.R. & Telang, R. (2007).** The Effect of Digital Sharing Technologies on Music Markets: A Survival Analysis of Albums on Ranking Charts. *Management Science*, 53 (9), 1359-74.
- Christman, E. (2013).** Universal Music Still Market Top Dog In 2012,” Billboard.com, <http://www.billboard.com/biz/articles/news/1510504/universal-music-still-market-top-dog-in-2012> (retrieved 5-June-2013).
- Boehmer, E., Masumeci, J. & Poulsen, A.B. (1991).** Event-study methodology under conditions of event-induced variance. *Journal of Financial Economics*, 30 (2), 253-72.
- Bradlow, E.T. & Fader, P.S. (2001).** A Bayesian Lifetime Model for the “Hot 100” Billboard Songs. *Journal of the American Statistical Association*, 96 (454), 368-81.
- Chaney, P.K., Devinney, T.M. & Winer, R.S. (1991).** The Impact of New Product Introductions on the Market Value of the Firm. *Journal of Business*, 64 (4), 573-610.
- Chellappa, R.K. & Shivendu, S. (2005).** Managing piracy: Pricing and sampling strategies for digital experience goods in vertically segmented markets. *Information Systems Research*, 16 (4), 400-17.
- Chevalier, J. & Goolsbee, A. (2003).** Measuring Prices and Price Competition Online: Amazon.com and BarnesandNoble.com” *Quatitative Marketing and Economics*, 1 (2), 203-222.
- Connolly, M. & Krueger, A.B. (2006).** Chapter 20. Rockonomics: The Economics of Popular Music, In V. A. Ginsburg, & D. Throsby (Eds.), *Handbook on the Economics of Art and Culture* (pp. 667-719), Vol. 1, Oxford: Elsevier.
- Dodd, P. & Warner, J.B. (1983).** On corporate governance : A study of proxy contests. *Journal of Financial Economics*, 11 (1-4), 401-38.
- Ebbes, P., Papies, D. & van Heerde, H.J. (2011).** The Sense and Non-sense of Hold-Out Sample Validation in the Presence of Endogeneity. *Marketing Science*, 30 (6), 1115-1122.
- Elberse, A. (2008).** Should You Invest in the Long Tail? *Harvard Business Review*, 86 (7/8), 88-96.
- Elberse, A. (2010).** Bye-Bye Bundles: The Unbundling of Music in Digital Channels. *Journal of Marketing*, 74 (3), 107-23.
- Fama, E.F., Fisher, L., Jensen, M.C. & Roll, R. (1969).** The Adjustment of Stock Prices to New Information. *International Economic Review*, 10 (1), 1-21.
- Fosfuri, A. & Giarratana, M.S. (2009).** Masters of War: Rivals' Product Innovation and New Advertising in Mature Product Markets. *Management Science*, 55 (2), 181-91.
- Giles, D.E. (2006).** Superstardom in the US popular music industry revisited. *Economics Letters*, 92 (1), 68-74.
- Goodway, N. (2005).** Coldplay album setback sends EMI into a spin. In *The Evening Standard*, London.

- Graham, J. (2009).** Musicians ditch studios for tech such as GiO for Macs,” USA TODAY, http://usatoday30.usatoday.com/tech/news/2009-10-13-apogee-gio-music_N.htm (retrieved 29-March-2014).
- Hamlen, W.A., Jr (1994).** Variety and Superstardom in Popular Music. *Economic inquiry*, 32 (3), 395-407.
- Heiman, A., McWilliams, B. Shen, Z.H. & Zilberman, D. (2001).** Learning and forgetting: Modeling optimal product sampling over time. *Management Science*, 47 (4), 532-46.
- Hendricks, K.B. & Singhal, V.R. (1997).** Delays in new product introductions and the market value of the firm: The consequences of being late. *Management Science*, 43 (4), 422-36.
- Hennig-Thurau, T., Henning, V., Sattler, H., Eggers, F. & Houston, M.B. (2007).** The Last Picture Show? Timing and Order of Movie Distribution Channels. *Journal of Marketing*, 71 (4), 63-83.
- Hirschman, E.C., & Holbrook, M.B. (1982).** Hedonic Consumption: Emerging Concepts, Methods and Propositions. *Journal of Marketing*, 46 (3), 92-101.
- Joshi, A.M. & Hanssens, D.M. (2009),** Movie Advertising and the Stock Market Valuation of Studios: A Case of “Great Expectations”? *Marketing Science*, 28 (2), 239-50.
- Kahneman, D. & Tversky, A. (1979).** Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47 (2), 263-91.
- Kalaighnam, K., Shankar, V. & Varadarajan, R. (2007).** Asymmetric New Product Development Alliances Win-Win or Win-Lose Partnerships. *Marketing Science*, 53 (3), 357-74.
- Kelm, K.M., Narayanan, V.K. & Pinches, G.E. (1995).** Shareholder Value Creation During R&D and Commercialization Stages. *Academy of Management Journal*, 38 (3), 770-86.
- Krasilovsky, M.W. & Shemel, S. (2007).** *This business of music: the definitive guide to the music industry*. 10 ed., New York: Billboard Books.
- Leeds, J. (2007).** The New Deal: Band as Brand,” The New York Times, <http://www.nytimes.com/2007/11/11/arts/music/11leed.html?pagewanted=all&r=0> (retrieved 29-March-2014).
- McWilliams, A. & Siegel, D. (1997).** Event studies in management research: Theoretical and empirical issues. *Academy of Management Journal*, 40 (3), 626-57.
- Ransbotham, S. & Mitra, S. (2010).** Target Age and the Acquisition of Innovation in High-Technology Industries. *Management Science*, 56 (11), 2076-93.
- Rob, R. & Waldfogel, J. (2006).** Piracy on the high C’s: Music downloading, sales displacement, and social welfare in a sample of college students. *Journal of Law and Economics*, 49, 29-62.
- Rosen, S. (1981).** The Economics of Superstars. *American Economic Review*, 71 (5), 845-58.
- Salganik, M.J., Dodds, P.S. & Watts, D.J. (2006).** Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science*, 311, 854-856.
- Sood, A. & Tellis, G.J. (2009).** Do Innovations Really Pay Off? Total Stock Market Returns to Innovation. *Marketing Science*, 28 (3), 442-56.
- Sorescu, A.B. & Spanjol, J. (2008).** Innovation's Effect on Firm Value and Risk: Insights from Consumer Packaged Goods. *Journal of Marketing*, 72 (2), 114-32.
- Sorescu, A., Shankar, V. & Kushwaha, T. (2007).** New Product Preannouncements and Shareholder Value: Don't Make Promises You Can't Keep. *Journal of Marketing Research*, 44 (3), 468-89.

- Sun, M. (2012).** How Does the Variance of Product Ratings Matter? *Management Science*, 58 (4), 696–707.
- Strobl, E.A. & Tucker, C. (2000).** The Dynamics of Chart Success in the U.K. Pre-Recorded Popular Music Industry. *Journal of Cultural Economics*, 24 (2), 113-34.
- Tellis, G.J. & Johnson, J. (2007).** The Value of Quality. *Marketing Science*, 26 (6), 758–73.
- Yanbin, T. & Min, L. (2006).** An Experimental and Analytical Study of on-Line Digital Music Sampling Strategies. *International Journal of Electronic Commerce*, 10, 39-7

Erklärung zum selbständigen Verfassen der Arbeit

Ich erkläre hiermit, dass ich meine Doktorarbeit „Essays zur Förderung und Bewertung von Innovationen“ selbstständig und ohne fremde Hilfe angefertigt habe und dass ich als Koautor maßgeblich zu den weiteren Fachartikeln beigetragen habe. Alle von anderen Autoren wörtlich übernommenen Stellen, wie auch die sich an die Gedanken anderer Autoren eng anlehenden Ausführungen der aufgeführten Beiträge wurden besonders gekennzeichnet und die Quellen nach den mir angegebenen Richtlinien zitiert.

Ort, Datum

Björn Goerke