

Kumulative Dissertation zur Erlangung des akademischen Grades  
eines Doktors der Wirtschaftswissenschaften  
Doktor rerum politicarum (Dr. rer. pol.)  
an der Wirtschaftswissenschaftlichen Fakultät  
der Katholischen Universität Eichstätt-Ingolstadt

## Forecasting Price Decline Events for Consumer Goods

vorgelegt von  
Benjamin Buchwitz

Referent  
Prof. Dr. Ulrich Küsters

Korreferent  
Prof. Dr. Thomas Setzer



Ingolstadt im September 2019

## Überblick

Die vorliegende kumulative Dissertation wurde am Lehrstuhl für Statistik und Quantitative Methoden an der Wirtschaftswissenschaftlichen Fakultät der Katholischen Universität Eichstätt-Ingolstadt verfasst und beschäftigt sich mit der prototypischen Entwicklung, Anwendung und Konfiguration von Prognosesystemen zur Vorhersage von Preissenkungsereignissen. Im Rahmen der Promotion wurden die folgenden fünf Forschungsbeiträge angefertigt:

1. Buchwitz, Benjamin; Küsters, Ulrich (2018): **A Time Series based Monitoring Methodology to Optimize Purchase Timing Decisions**, Working Paper, SSRN ID 3179987, Catholic University of Eichstätt-Ingolstadt.
2. Buchwitz, Benjamin; Küsters, Ulrich (2018): **Should I buy my new iPhone now? Predictive Event Forecasting for Zero-Inflated Consumer Goods Prices**, Proceedings of the 39<sup>th</sup> International Conference on Information Systems 2018, San Francisco, California (Ranking: VHB-JQ2: A, VHB-JQ3: A, CORE2018: A\*).
3. Buchwitz, Benjamin (2019): **Evaluation and Optimal Calibration of Purchase Time Recommendation Services**, Proceedings of the 52<sup>nd</sup> Hawaii International Conference on System Sciences 2019, Maui, Hawaii (Ranking: VHB-JQ2: -, VHB-JQ3: C, CORE2018: A).
4. Falkenberg, Anne; Buchwitz, Benjamin (2019): **Enhancing Price Alert Recommendation Services – A Comparative Study**, Proceedings of the 27<sup>th</sup> European Conference on Information Systems 2019, Stockholm & Uppsala, Sweden (Ranking: VHB-JQ2: B, VHB-JQ3: B, CORE2018: A).
5. Buchwitz, Benjamin; Falkenberg, Anne; Küsters, Ulrich (2019): **Time Series Event Forecasting in Consumer Electronic Markets using Random Forests**, Working Paper, Catholic University of Eichstätt-Ingolstadt.

Preiszeitreihen sind außerhalb der Finanzökonomie nur in sehr begrenzten Rahmen sowie in kleinen Mengen verfügbar. Eine Ausnahme davon bilden die im E-Commerce relevanten Preiszeitreihen technologischer Konsumgüter wie Smartphones, Tablets oder elektronische Komponenten. Entscheidungsrelevant und von besonderem Interesse ist hier vor allem der geringste verfügbare Preis, der für die Mehrheit der preissensitiven Verbraucher maßgebend ist und oft als Referenzpreis herangezogen wird. Die zugehörigen Zeitreihen werden z. B. auf Preisvergleichsseiten abgebildet und können von Kunden im Rahmen der Festlegung ihres Kaufzeitpunktes herangezogen werden. Als Minimumpreiszeitreihe bezeichnet man dabei die emergente Datenreihe, die sich bei der Selektion des geringsten Preises aus einer Gruppe von Händlern für ein eindeutig zu differenzierendes und homogenes Produkt ergibt. Charakteristisch für die Zeitreihen im Falle von Konsumgüterprodukten ist üblicherweise (a) ein kontinuierlicher Preisverfall, (b) alternierend ruhige und aktive Preissetzungsphasen und (c) sporadisch auftretende konstante Segmente.

- a) Durch die regelmäßige Einführung von Nachfolgeprodukten sowie dem stetigen technologischen Fortschritt auf der Ebene der Konsumgüterelektronik bilden sich in Abhängigkeit der individuellen Produkteigenschaften sowie der Produktkategorie markante Preisverläufe heraus. Insbesondere stark nachgefragte High-Tech Produkte, wie Smartphones, sind dabei durch regen Preisverfall gekennzeichnet, was einem zeitvarianten Erwartungswert entspricht.
- b) Die heterogene Händlerstruktur und das intensive Wettbewerbsgefüge führen zu unterschiedlich intensiven Preissetzungsphasen, bei denen sowohl die Größe als auch die Häufigkeit der Veränderungen des Minimumpreises deutlich schwanken. Diese alternierend ruhigen und aktiven Phasen sind in der Regel dadurch gekennzeichnet, dass auf kleine bzw. große vorausgegangene Preisänderungen üblicherweise Änderungen in vergleichbarer Größe folgen. Die dadurch entstehenden Häufungen werden als Volatilitätscluster bezeichnet.
- c) Durch die kompetitive Preisgestaltung im Händlergefüge ergeben sich, verglichen mit den Preisen der einzelnen Händler, viele Preisbewegungen in der Minimumpreiszeitreihe. Es finden sich faktisch jedoch bei allen beobachteten Produkten Phasen mit konstanten Preisen. In der Sequenz der Preisänderungen, die sich aus der Anwendung des einfachen Differenzenbildungsoperators auf die Zeitreihe ergibt, werden diese konstanten Segmente zu Null-Beobachtungen. Die zugehörige statistische Eigenschaft wird daher als Nullinflationierung (engl. Zero-Inflation) bezeichnet.

Alle drei zentralen Charakteristika bieten Ansätze zur Modellierung und Fortschreibung der Preiszeitreihen, wobei die Komplexität der Modellierung mit der Berücksichtigung der Dateneigenschaften in genannter Reihenfolge ansteigt. Ziel der Dissertation ist die Entwicklung eines methodischen Rahmens, um aus historischen Minimumpreiszeitreihen Empfehlungen abzuleiten, die potenzielle Kunden bei der Bestimmung eines wirtschaftlich vorteilhaften Kaufzeitpunktes unterstützen. Die dafür notwendigen Schritte der Entwicklung, Anwendung und Konfiguration werden in den einzelnen Forschungsbeiträgen beleuchtet. Beitrag 1 legt die Grundlagen der Prognosemethodik und zeigt Möglichkeiten und Herausforderungen bei der Evaluation abgeleiteter Entscheidungen auf. Beitrag 2 verdeutlicht die entwickelte Methodik und zeigt mittels eines realen Datensatzes, dass damit erstellte Entscheidungsempfehlungen wirtschaftlichen Mehrwert

bieten. Beitrag 3 beleuchtet den Fall eines zu setzenden Schwellenwertes für eine Mindestersparnis genauer und zeigt, dass dieser Schwellenwert wirtschaftlich optimal auf Basis der Minimumpreiszeitreihe bestimmt werden kann. Zudem wird gezeigt, dass dieser Schwellenwert bereits 90 Tage nach Eintritt eines Produktes in den Markt präzise geschätzt werden kann. Beitrag 4 betrachtet einen weiteren Anwendungsfall im Kontext von Preisvergleichsportalen und entwickelt einen zusätzlichen Ansatz zur Erweiterung des als Preisalarm bezeichneten Benachrichtigungsservices. Die vorgestellte Methodik erlaubt es, bereits bei der Konfiguration der Preisbenachrichtigung zu bestimmen, wie lange Kunden auf die Realisierung des von ihnen konfigurierten Preises durchschnittlich warten müssen. Beitrag 5 erweitert das eingesetzte Methodenspektrum zur Prognose von Preisänderungsereignissen und zeigt, dass die entscheidenden Dynamiken über Produktgruppen hinweg verallgemeinerbar sind.

Im Folgenden werden die erarbeiteten Studien und ihre Ergebnisse zusammengefasst und in den Gesamtkontext der Arbeit eingeordnet. Die Datenbasis für die beschriebenen Forschungsbeiträge wurde im Rahmen eines vorbereitenden Projektes vom Autor der Dissertation mit Hilfe einer parallelisierten globalen Crawlerarchitektur über einen Zeitraum von etwa 1.5 Jahren gesammelt. Dementsprechend beziehen sich alle genannten Ergebnisse auf real verfügbare Marktpreise des deutschen E-Commerce Sektors.

### **Beitrag 1: A Time Series based Monitoring Methodology to Optimize Purchase Timing Decisions**

Ausgehend von der zentralen Forschungsfrage der Dissertation beschreibt dieser Artikel die Entwicklung eines Systems zur Entscheidungsunterstützung für Käufer von Technologieprodukten. Die in diesem Artikel erarbeiteten fünf Kernbeiträge sind die Definition von kundenrelevanten Preisänderungsereignissen für Minimumpreiszeitreihen, die Entwicklung einer Prognosemethodik auf Basis von ARIMA-GARCH Modellen, die Ableitung einer Methode zur dynamischen Bewertung und Steuerung der ökonomisch relevanten Prognosequalität, die Präsentation eines Ansatzes zur Quantifizierung des resultierenden Kundennutzens bei Einsatz des Gesamtsystems und die exemplarische Illustration des Vorgehens an einem empirischen Beispiel.

Relevante Preisänderungsereignisse ergeben sich auf Basis der zukünftigen Preisänderungskonstellationen, die einen Kunden dazu veranlassen, seinen Kaufzeitpunkt für ein spezifisches Produkt vom gegenwärtigen auf ein späteres Datum zu verlegen. Entsprechende Ereignisse werden in Abhängigkeit von zwei zentralen Größen – einer erforderlichen Mindestpreissenkung sowie der Länge des Zeitraumes, in dem diese Preisänderung realisiert werden muss – formal definiert. Trotz der vorhandenen konstanten Segmente in den Minimumpreiszeitreihen ist nicht jede zukünftig eintretende Preisänderung für alle potenziellen Kunden maßgeblich und führt nicht zwingend zu einer Änderung des geplanten Kaufzeitpunktes. Ein relevantes Preisänderungsereignis tritt demzufolge ein, wenn eine Preisreduktion in festgelegter Höhe in einer bestimmten Zeitspanne beginnend mit dem Prognoseursprung mindestens einmalig erreicht wird.

Ziel der Prognosemethodik ist die Schätzung der Eintrittswahrscheinlichkeit eines solchen spezifizierten Preisänderungsereignisses. Üblicherweise eingesetzte Methoden zur Berechnung einer Punktprognose für alle Zeitpunkte im Entscheidungshorizont können nicht unmittelbar zur Bestimmung einer Eventeintrittswahrscheinlichkeit verwendet werden und ignorieren die inhärente Unsicherheit zukünftiger Preisbewegungen. Das vorgeschlagene Verfahren involviert daher die Spezifizierung eines ARIMA(1,1,1)-GARCH(1,1) Modells auf Basis der historischen Minimumpreise. Ausgehend von dieser Modellschätzung wird ein speziell modifiziertes Residual-Bootstrap-Resampling Schema genutzt, um zahlreiche Szenarien für die zukünftige Preisentwicklung zu generieren. Dabei wird sowohl der globale Preisverfall (a) als auch die Dynamik etwaiger ruhiger oder aktiver Marktphasen (b) explizit modelliert sowie die Unsicherheit durch Einbeziehung der Schätzfehler in die Szenarien berücksichtigt.

Während die Eventeintrittswahrscheinlichkeit einen Indikator für das Auftreten des definierten Preisevents darstellt, ist unklar, ab welchem konkreten Wahrscheinlichkeitswert eine Verlegung des Kaufzeitpunktes zu wirtschaftlichen Vorteilen führt. Dieser Sachverhalt erfordert die Partitionierung des Ereignisraumes mittels eines zu kalibrierenden Schwellenwertes. Da ein global optimaler Wahrscheinlichkeitsschwellenwert für jeden Zeitpunkt weder modelltheoretisch hergeleitet werden kann noch dessen Existenz praktisch plausibel ist, muss der Schwellenwert produkt- und ursprungsabhängig revidiert und neu bestimmt werden. Eine kontinuierliche Rekalibrierung der Parameter und Schwellenwerte in Abhängigkeit der Werte der ökonomischen Zielgröße erlaubt zudem eine adaptive asymmetrische Risikosteuerung. Diese berücksichtigt das Risiko konstanter Preise (c) implizit sowie die unterschiedlichen ökonomischen Implikationen, also Ersparnisse für korrekt bzw. Verluste für fälschlicherweise gegebene Empfehlungen explizit. Der Vergleich der Eventeintrittswahrscheinlichkeit mit dem Schwellenwert erlaubt zu jedem Zeitpunkt die Klassifikation und damit die Kommunikation klarer Handlungsempfehlungen der Form „Kaufe Produkt jetzt“ oder „Warte auf Preissenkung“.

Zur Quantifizierung des ökonomischen Nutzwertes des Empfehlungssystems für Endkunden wird eine Messmethodik entwickelt, die kongruent zur Schwellenwertkalibration jeder Empfehlung faktische Ersparnisse bzw. Kosten zuweist und so einen realistischen Eindruck über die Leistung des Gesamtsystems liefert. Dabei muss zwischen zwei Fällen – Festlegung einer bzw. keiner expliziten Mindestpreissenkung – unterschieden werden. Das erste Szenario führt zu den vier klassischen Fällen der binären Klassifikatorevaluation (True Positive, False Positive, True Negative, False Negative). Im zweiten Fall wird die Konfusionsmatrix durch zwei Sonderfälle ergänzt, die die Bewertung von Situationen erlauben, in denen der Preis im Entscheidungszeitraum zwar unter den Ausgangspreis fällt, die festgelegte Mindestpreissenkung jedoch nicht erreicht wird.

### **Beitrag 2: Should I buy my new iPhone now? Predictive Event Forecasting for Zero-Inflated Consumer Goods Prices**

Aufbauend auf der entwickelten Methodik der ersten Forschungsarbeit leistet dieser Artikel einen Beitrag zur Anwendung des entscheidungsunterstützenden Prognosesystems und belegt, dass das Verfahren Käufern von Technologieprodukten im praktischen Einsatz deutlich messbare wirtschaftliche Vorteile bietet. Die vier Kernbeiträge dieser Arbeit sind eine leicht zugängliche Darstellung der entwickelten Methodik, die Vorstellung eines großen Datensatzes für den deutschen Smartphone-Markt, die Definition und ökonomische Bewertung von zwei Handlungs- bzw. Kaufstrategien und die ausführliche Evaluation von Entscheidungsempfehlungen.

Ausgangspunkt für die Analyse ist eine Teilmenge des oben erwähnten Datensatzes mit Minimumpreiszeitreihen für 238 Smartphones mit einer jeweiligen Länge von 730 täglichen Preisbeobachtungen, sodass insgesamt 173.740 Preise in die Analyse einfließen. Die Produkte repräsentieren einen Großteil der im deutschen Markt verfügbaren relevanten Mobiltelefone und decken nahezu das vollständige Marken- und Preisspektrum ab.

Verglichen wird die entwickelte Prognosemethode mit zwei Handlungsstrategien von potenziellen Käufern. Die erste Option entspricht dem trivialen Status Quo, bei dem Kunden ohne weitere Informationen das betrachtete Produkt zum jeweils tagesaktuellen Minimumpreis erwerben. Konträr dazu bildet die zweite Option die bei kontinuierlichem Preisverfall langfristig dominante Strategie ab. Hierbei warten Kunden bis zum Ende ihres definierten Entscheidungshorizontes und erwerben das Produkt zum spätmöglichen Zeitpunkt, um von dem im Mittel niedrigeren Preisen zu profitieren.

Für die Evaluation werden mit dem erläuterten Verfahren insgesamt 150.654 automatisierte Entscheidungsempfehlungen mittels rollierender Prognosesimulation generiert. Die Anzahl der erstellten Preisverlaufsszenarien wird dabei mit 10.000 und die maximale Länge des Entscheidungszeitraumes mit sieben Tagen operationalisiert. Als Zielpreis wird jeder Preis, der niedriger als der Preis am Prognoseursprung ist, akzeptiert. Die bei Befolgung der Empfehlungen entstehenden ökonomischen Implikationen können in reale und hypothetische Ersparnisse bzw. Kosten unterteilt werden. Reale Gewinne und Verluste ergeben sich dabei für Fälle, bei denen ein Kunde – ausgehend vom Status Quo des Direktkaufs – seinen Kaufzeitpunkt durch die vom Prognosesystem kommunizierte Empfehlung verlegt. Hypothetische Bewertungsgrößen resultieren aus den verbleibenden Fällen (Kunde kauft direkt) und haben keinen direkten Zahlungsstrom zur Folge, da der Produktkauf bereits getätigt wurde. Beide Ansätze führen zu vergleichbaren Resultaten, daher wird lediglich auf die für Kunden wesentlichen realen Zahlungsströme Bezug genommen. Dies hat zudem den Vorteil, dass die erste Benchmark-Option auf eine monetäre Performance von Null Euro normiert wird und so einen neutralen Referenzwert abbildet. Um die Bewertung über die Produkte hin vergleichbar zu machen, wird als zentrales Evaluationsmaß der Gewinnausschöpfungsgrad definiert, bei dem die jeweils erzielte Ersparnis in Relation zur Maximalersparnis eines perfekten Prädiktors gesetzt wird. Während das Warten bis zum letztmöglichen Kaufzeitpunkt im Entscheidungshorizont im Mittel insgesamt 45.03% der Ersparnisse erschließt, wird die Gewinnausschöpfung durch die gegebenen Empfehlungen um 9.36% auf im Mittel 54.39% gesteigert. Dabei ergeben sich für alle Marken signifikant gesteigerte durchschnittliche Ersparnisse von bis zu 26.67€ pro gegebener Empfehlung.

### **Beitrag 3: Evaluation and Optimal Calibration of Purchase Time Recommendation Services**

Das vorgeschlagene und in den Beiträgen 1 und 2 entwickelte und evaluierte Prognosesystem kann mit zwei Parametern (Entscheidungshorizont und relevante Mindestpreissenkung) konfiguriert werden. Insbesondere die Länge des Entscheidungshorizontes kann auch vom Kunden selbst subjektiv gesetzt oder durch externe Faktoren bestimmt werden. Die Festlegung einer erforderlichen Mindestersparnis und damit eines entsprechenden Zielpreises ist jedoch oft nicht eindeutig oder vom Kunden ohne Referenzwerte oder Illustration der daraus resultierenden Folgen möglich. Daher leistet diese Forschungsarbeit einen Beitrag zur Konfiguration von Prognosesystemen für Preisevents. Die drei Kernbeiträge sind dabei die Diskussion der Auswirkungen bei Setzung einer Mindestersparnis, die Definition einer

wirtschaftlich optimalen und produktabhängigen Parameterkonfiguration und die Entwicklung eines Ansatzes zur Schätzung der optimalen Konfiguration der Mindestersparnis kurz nach Eintritt eines Produktes in den Markt.

Das explizite Setzen einer Mindestersparnis hat im Wesentlichen drei Auswirkungen. Zum einen erhalten die generierten Empfehlungen einen direkten Bezug zur Situation und dadurch eine höhere Relevanz für den jeweiligen Kunden. Zum zweiten sinkt die Anzahl der auftretenden Preiseventeintritte, sobald die gesetzte Mindestersparnis größer als die kleinste beobachtete Preissenkung ist. Drittens verändert sich die zu erzielende Maximalersparnis pro Produkt. Dies kommt dadurch zustande, dass wirtschaftlich irrelevante kleine Preisänderungen bei entsprechender Kalibration übersprungen und darauffolgende, ebenfalls im Entscheidungszeitraum befindliche größere Preisänderungen als Einsparung realisiert werden können.

Der optimale Preisersparnisschwellenwert ist der Wert, bei dem die auszuschöpfende Gesamtersparnissumme, gemessen über alle Zeitpunkte für ein spezifisches Produkt, maximiert wird. Im diskutierten Beispiel steigt die Maximalersparnis bei Annahme eines perfekten Prädiktors um 35.62% im Vergleich zum Basisschwellenwert und vergrößert so das Ersparnispotenzial. Zeitgleich wird die Anzahl der Preisevents reduziert, sodass sich ein ausgeglichenes Verhältnis von Preissenkungs- zu Nicht-Preissenkungseignissen einstellt, was üblicherweise im Rahmen der Empfehlungsgenerierung ausgenutzt werden kann.

Während die retrospektive Analyse bei gegebener Minimumpreiszeitreihe verhältnismäßig einfach durchgeführt und so der optimale Wert der Mindestersparnis ex-post bestimmt werden kann, ist für die praktische Anwendung eine Konfiguration vor der Realisierung der Preise zwingend notwendig. Zur Schätzung dieses produktspezifischen Parameters wird daher ein Ansatz vorgestellt, der den auf der aktuell verfügbaren Zeitreihe bestimmten ex-post Schwellenwert adjustiert. Als Kovariate werden dafür Zeitreihenfeatures, wie die Autokorrelationsstruktur, das Preisniveau, die Nicht-Linearität, die Spektralentropie und die Markenzugehörigkeit herangezogen und der ex-post optimale Schwellenwert zwei Jahre nach Markteintritt für jedes Produkt prognostiziert. Die per Kreuzvalidierung erzeugten out-of-sample Schätzungen erschließen bei Produkten, die seit lediglich 90 Tagen angeboten werden, im Mittel bereits 97.05 % des maximal möglichen Ersparnispotenzials. Verglichen mit der Basiskonfiguration erleichtern die mit der vorgestellten Methodik bestimmten Mindestersparnisparameter die Konfiguration sowie Nutzung des Empfehlungssystems und führen zu besseren Prognosekonfigurationen bei signifikant gesteigertem Ersparnispotenzial.

#### **Beitrag 4: Enhancing Price Alert Recommendation Services – A Comparative Study**

In dieser Arbeit wird ein Ansatz entwickelt, um Preisalarme im Kontext von Preisvergleichsportalen methodisch und konzeptionell zu erweitern. Die drei Kernbeiträge sind dabei die Untersuchung der Eignung von Prognosemethoden für die Prognose von Zeitpunkten im Kontext von Konsumgüterprodukten, die generalisierte Erweiterung der Methodik um eine automatische Parameterkalibration und ein Methodenvergleich von vier Basismethoden mit und ohne Ensembleprognose für eine Teilmenge des initial vorgestellten Minimumpreiszeitreihendatensatzes.

Ein Preisalarm ist ein Benachrichtigungsservice, der Kunden die Eingabe eines Wunschpreis ermöglicht, um sie anschließend per E-Mail oder mittels mobiler Benachrichtigung einer App zu informieren, wenn dieser oder ein geringerer Preis von einem Händler angeboten wird. Prognosegegenstand ist der Zeitpunkt und damit die verbleibende Dauer bis zum Eintreten des kundenspezifischen Zielpreises, die dem Kunden bei der Konfiguration des Preisalarms angezeigt werden soll. Dies ermöglicht die individuelle Abwägung jedes Kunden, ob für die gewünschte Ersparnis die Wartedauer angemessen ist. Andernfalls kann der Kaufinteressent seinen Zielpreis adjustieren oder das Produkt direkt zum aktuell angebotenen Preis erwerben.

Alle untersuchten und tentativ berücksichtigten Punktprognoseverfahren liefern keine zufriedenstellenden Ergebnisse. Ergänzend führen sie aufgrund von Fehleinschätzungen des Preisverlaufes häufig nicht zu einer Zeitpunktprognose, da der Punktprognosepfad niemals unter den Zielpreis (5% Preisreduktion) fällt. Um die inhärente Unsicherheit der Prognose zu berücksichtigen, wird ein Residual-Bootstrap-Resampling Schema genutzt und multiple mögliche Preisverläufe generiert. Jeder simulierte Prognosepfad führt nun gegebenenfalls zu einer Zeitpunktprognose, sodass aus der Menge der generierten Preisverläufe eine diskrete Dichte für die Zeitpunkte der Zielpreiserreichung abgeleitet werden kann. Ausgehend von einem ex-post optimalen Quantil, das auf Basis der bereits eingetretenen wahren Eventzeitpunkte bestimmt wird, kann nun eine Prognose für den Zeitpunkt der Erreichung des Zielpreises abgeleitet werden. Alle vier Methoden sind – erweitert um den Ensemble-Ansatz – in der Lage, Prognosen für jeden Zeitpunkt bei gleichzeitig signifikant präziseren Ergebnissen zu erzielen.

**Beitrag 5: Time Series Event Forecasting in Consumer Electronic Markets using Random Forests**

In diesem Beitrag wird das verwendete Methodenspektrum verbreitert, um multivariate Informationen, also Preisverläufe und daraus ableitete Eigenschaften von mehreren bereits ausgelisteten sowie aktuellen Produkten zur Prognose heranziehen zu können. Die drei Kernbeiträge des Beitrags sind die Nutzung des Random Forest Klassifikationsverfahrens mit Zeitreihendaten durch Definition und Selektion prädiktiv nutzbarer uni- und multivariater Kovariate, die Präsentation einer möglichen Lösung für das Cold-Start Problem, sodass Prognosen schneller als bei den vorgestellten univariaten Ansätzen generiert werden können sowie Evidenz dafür, dass Preisänderungsdynamiken über Produkte und Marken hinweg generalisierbar sind.

Klassifikationsverfahren benötigen methodisch bedingt erklärende Variablen. Diese sind bei Zeitreihen, wie den verwendeten Minimumpreiszeitreihen, oft nicht verfügbar, können jedoch auf Basis der (endogenen) Informationen konstruiert werden. Analog zu den verwendeten univariaten Verfahren werden zunächst vergangene Preisbeobachtungen herangezogen. Aufgrund der unterschiedlichen Preisniveaus der gebündelten Produkte werden diese jedoch normiert als Renditen verwendet und die Historie auf sieben Beobachtungen gekürzt. Weitere konstruierte Variablen wie das Preisniveau, der mittlere Preis des Produktes, der Anteil der Nullbeobachtungen oder das Verhältnis des aktuellen Preises zum durchschnittlichen Preis der vergangenen Woche können ebenfalls aus der Zeitreihe eines Produktes generiert werden. Bei Vorliegen mehrerer Zeitreihen können zudem Marktindikatoren wie der mittlere Preis aller Produkte, das Niveau des Preisverfalls an einem gegebenen Tag oder die Anzahl aktueller Produkte ergänzt werden. Neben den erwähnten Renditen wurden insgesamt elf weitere Einflussgrößen identifiziert und in die Analyse mit aufgenommen.

Mit der berechneten Datenbasis wurden Modelle sowohl mit produktspezifischen Informationen und damit für jedes Produkt individuell als auch mittels globaler Informationen konfiguriert. Dadurch kann die Hypothese geprüft werden, ob multivariate Variablen und eine größere Datenbasis zusätzliche Informationen bereitstellen, die die prädiktive Güte steigern. In der durchgeführten Analyse konnte gezeigt werden, dass Random Forest Modelle für einen Großteil der analysierten Smartphones zu besseren Ergebnissen und einem größeren Anteil korrekt identifizierter Preissenkungsereignisse führen, wenn diese auf Basis der gesamten Datenbasis über Produkte und Marken hinweg konfiguriert werden. Bei der Erstellung von Prognosen können für neue Produkte so bereits nach Vorliegen der notwendigen sieben Renditen, Empfehlungen für neue Produkte nahezu unmittelbar nach Markteintritt erstellt werden. Diese frühen Empfehlungen unterscheiden sich dabei in ihrer Qualität nicht strukturell von den später generierten. Verglichen mit dem univariaten Ansatz steigt die an der Area-under-Curve (AUC) der Receiver-Operator-Charakteristik gemessenen Prognosequalität im Mittel von 63.8% auf 67.6%.

**Fazit**

Die vorliegende Arbeit stellt den methodischen Rahmen für ein prototypisches Prognosesystem bereit, das aus Minimumpreiszeitreihen von elektronischen Konsumgütern Empfehlungen ableitet, die Online-Shopper bei der Festlegung von Kaufzeitpunkten unterstützt. Durch die Forschungsbeiträge werden dabei Ansätze für die grundlegende Methodenentwicklung, die ökonomisch optimale Konfiguration von Parametern des Gesamtsystems sowie praktisch relevante Einsatzszenarien präsentiert und so gezeigt, dass trotz der schwierigen Datenlage ein profitabler Empfehlungsservice realisiert werden kann. Während Anbieter von Preisvergleichsplattformen mit der vorgestellten Methodik ihren verfügbaren Datenbestand besser nutzen und durch einen mit dem Geschäftsmodell kongruenten Service einen nachhaltigen Wettbewerbsvorteil erzielen können, werden Verbraucher dazu befähigt, methodisch fundierte und wirtschaftlich vorteilhafte Kaufentscheidungen zu treffen.

Beitrag 1

## A Time Series based Monitoring Methodology to Optimize Purchase Timing Decisions

Benjamin Buchwitz, Ulrich Küsters

**Abstract:** Businesses as well as consumers utilize price comparison portals prior to purchasing. Usually such systems provide price time series, but the transformation of the embedded information to select an appropriate purchase time is unknown. We present a methodology to forecast the probability of user customizable sufficient price change events. Four main methodological contributions are presented: (i) an economically meaningful definition of user specified price decreases, (ii) the modification of a bootstrap based ARIMA-GARCH volatility forecasting method to predict the probability of the defined events, (iii) the dynamic statistical evaluation of the forecasting accuracy and (iv) the measurement of the economic utility of the buying recommendation procedure using gain functions. Beyond this, the technique is applied to two distinct forecasting situations, which clearly show the dominance of the proposed decision theoretic framework in comparison to naive purchase strategies like always delaying or always buying immediately.

Buchwitz, Benjamin; Küsters, Ulrich (2018): **A Time Series based Monitoring Methodology to Optimize Purchase Timing Decisions**, Working Paper, SSRN ID 3179987, Catholic University of Eichstätt-Ingolstadt.

Beitrag 2

## Should I buy my new iPhone now? Predictive Event Forecasting for Zero-Inflated Consumer Goods Prices

Benjamin Buchwitz, Ulrich Küsters

**Abstract:** Price Comparison Sites enjoy great popularity because they enable customers to make better – more informed, less costly – buying decisions. We use a large dataset of daily minimum prices for 238 smartphones from Price Comparison Sites to develop a methodological approach to model the characteristics of consumer electronic goods prices and advice customers on their purchase time decision. The introduced method consists of two stages, the first one generating a multi-step price drop probability using a customized ARMA-GARCH model combined with an ensemble bootstrap. The second step constitutes a threshold for the estimated event probabilities that is optimized to achieve the highest possible savings for the customer. The derived recommendation advises users either to wait with the purchase or to buy the product of interest immediately to avoid losses due to increasing prices. In the evaluation, we compare our developed approach to two benchmark strategies: “always buy” and “always delay”, where the latter is expected to be the dominating strategy as prices show strong deterioration. The evaluation of 150,654 automated decisions shows that the proposed decision support system outperforms both benchmarks and delivers statistically as well as economically significant results. We reach savings of up to 26.67 € per recommendation and show that our fully automated decisions generate a significant added value and save money for the clients.

Buchwitz, Benjamin; Küsters, Ulrich (2018): **Should I buy my new iPhone now? Predictive Event Forecasting for Zero-Inflated Consumer Goods Prices**, Proceedings of the 39<sup>th</sup> International Conference on Information Systems 2018, San Francisco, California.



Beitrag 3

## Evaluation and Optimal Calibration of Purchase Time Recommendation Services

Benjamin Buchwitz

**Abstract:** Price Comparison Sites enable customers to make better – more informed, less costly – buying decisions through providing price information and offering buying advice in the form of prediction services. While these services differ to some extent, they are comparable regarding their prediction target and usually monitor every arbitrarily small price decrease. We use a large data set of daily minimum prices for 272 smartphones consisting of 198,560 daily price movements from a Price Comparison Site to show that the standard prediction setting is not optimal. A custom evaluation framework allows the maximization of the achievable savings by altering the calibration of the forecasting service to monitor changes that exceed a certain threshold. Additionally, we show that time series features calculated in a calibration period can be used to obtain precise out of sample estimates of the saving optimal forecasting setting.

Buchwitz, Benjamin (2019): **Evaluation and Optimal Calibration of Purchase Time Recommendation Services**, Proceedings of the 52<sup>nd</sup> Hawaii International Conference on System Sciences 2019, Maui, Hawaii.

Beitrag 4

## Enhancing Price Alert Recommendation Services – A Comparative Study

Anne Falkenberg, Benjamin Buchwitz

**Abstract:** Both online shoppers and e-commerce retailers benefit from price aggregation platforms that reduce searching costs for consumers and marketing expenses for businesses. The business model of Price Comparison Sites requires customers to frequently revisit their site. Therefore, they offer a range of services that support users in all pre-purchase stages of the buying process. A frequently used offer is the price alert service that notifies users when a customer-specific threshold price is reached. However, users are not assisted with configuring this service to find the best trade-off between waiting time and the amount of the defined saving. We use a large data set with 110,230 daily price observations for electronic consumer goods to develop a method that predicts when price alarms are triggered. The presented algorithm combines approaches from multiple fields and extends time series forecasting methodologies with a bootstrapped forecasting ensemble to generate various price development scenarios. We systematically reduce the uncertainty in the bootstrapped path space by dynamically calibrating a customised decision criterion and generate 62,061 automated predictions. Our proposed approach not only outperforms the benchmark forecasting models significantly in terms of accuracy but also produces precise estimates in cases where traditional approaches fail.

Falkenberg, Anne; Buchwitz, Benjamin (2019): **Enhancing Price Alert Recommendation Services – A Comparative Study**, Proceedings of the 27<sup>th</sup> European Conference on Information Systems 2019, Stockholm & Uppsala, Sweden.

Beitrag 5

## Time Series Event Forecasting in Consumer Electronic Markets using Random Forests

Benjamin Buchwitz, Anne Falkenberg, Ulrich Küsters

**Abstract:** Consumers are price-sensitive and opportunistic about the place of purchase when buying electronic goods. However, services that advise customers on purchase time decisions for electronic consumer goods are missing. Given the objective to provide a signal to customers to either wait or purchase immediately, classification algorithms are a direct methodological choice. Approaches like random forests allow for the derivation of a probability and class prediction but are usually not used in time series contexts. This is due to missing or non-time variant regressors and unclear prediction settings. We show how classification methods can be used to generate reliable predictions of price events and analyze if these are subject to common market dependencies. Pooling univariate random forests and enhancing them with multivariate features shows that our approach generates stable and valuable recommendations. Because dependency structures between products are transferable, multivariate forecasting increases accuracy and issues recommendations where univariate approaches fail.

Buchwitz, Benjamin; Falkenberg, Anne; Küsters, Ulrich (2019): **Time Series Event Forecasting in Consumer Electronic Markets using Random Forests**, Working Paper, Catholic University of Eichstätt-Ingolstadt.