

# Daten in Finance

Martin H. Schmidt

Zweite Jahrestagung des Arbeitskreises Finanzierung  
der Professorinnen und Professoren an Hochschulen für  
angewandte Wissenschaften  
9. Mai 2014

Humboldt University Berlin

School of Business and Economics

Office: Dorotheenstr. 1, 10117 Berlin

Mail: Unter den Linden 6, 10099 Berlin

Phone: +49 30 2093 5761

Email: [martin.h.schmidt@wiwi.hu-berlin.de](mailto:martin.h.schmidt@wiwi.hu-berlin.de)

Web: <http://lehre.wiwi.hu-berlin.de/Professuren/bwl/bb/>

# Agenda

2

## I. Datenbanken/-quellen in Finance

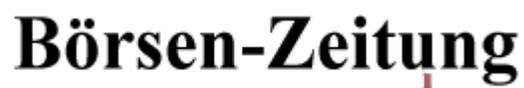
- Kostenfreie/frei zugängliche Datenbanken/-quellen
- Kostenpflichtige Datenbanken/-quellen

## II. Data Center SFB649 der Humboldt-Universität

## III. Datenqualität und Datenverarbeitung

## VI. Zusammenfassung/Diskussion

# I. Datenbanken/-quellen in Finance



# 1. Datenbanken/-quellen: Allgemein

4

- Es gibt eine mittlerweile unüberschaubare Anzahl von Datenbanken/-quellen
- Eine Einteilung kann u. a. nach folgenden Kriterien erfolgen:
  - ▣ kommerziell (e.g. Bloomberg, Thomson Reuters) vs. gemeinnützig (Kenneth French's Data Library)
  - ▣ kostenpflichtig vs. kostenfrei, komplett oder einzelne Bestandteile (e.g. stoxx.com)
  - ▣ Nischenanbieter (OptionMetrics) vs. Massenanbieter (Thomson Reuters)
  - ▣ Primär- (Data Shop Deutsche Börse) vs. Sekundärquelle (WRDS)
- Für kommerzielle Anbieter ist die Kundengruppe „Forscher“ i.d.R. uninteressant (kleiner Kreis, Rabatte, intensive Nutzung)

## 2. Kostenfreie Datenbanken/-quellen: Einführung

5

- Es gibt eine Vielzahl von Datenquellen, die kostenlos, aber für wissenschaftliche Zwecke oftmals ungeeignet sind.
- Die Gründe sind vielfältig:
  - Datenqualität ist nicht ausreichend  
(bspw. „historische Kurse“ von e.g. finanzen.net)
  - Aufbereitung der Daten ist mangelhaft (nicht konsistent)
  - Daten sind nur umständlich zu Gewinnen (kein „click-and-download“)  
(bspw. DGAP, wobei als XML-Datei auf Anfrage verfügbar)
  - Die zur Verfügung stehende Historie/Abdeckung ist zu kurz/gering
  - Die Verknüpfung mit anderen Daten fehlt (bspw. Aktiendaten mit Bilanzdaten)

...im Folgenden Datenquellen, die (teilweise) kostenlos sind und m. E. für wissenschaftliche Zwecke geeignet sind

## 2. Kostenfreie Datenbanken/-quellen: Auswahl (1)

6

### STOXX

- STOXX ist ein Indexanbieter
- Historische Zeitreihen kostenlos verfügbar für Länder/Regionen, Size, Style, Sektoren, etc.
- Zeitraum: i.d.R. ab ~2000
- [http://www.stoxx.com/data/historical/historical\\_benchmark.html](http://www.stoxx.com/data/historical/historical_benchmark.html)

### BVI

- Der Bundesverband Investment und Asset Management e.V. stellt auf Anfrage die eigene Fondsdatenbank zur Verfügung (unter anderem genutzt in Lehmann (2013)).
- Achtung: Daten enthalten Körperschaftssteuergutschrift
- Auf der Internetseite gibt es verschiedene Statistiken zur deutschen Fondslandschaft

## 2. Kostenfreie Datenbanken/-quellen: Auswahl (2)

7

**Deutsches Aktieninstitut**  
Kapital. Markt. Kompetenz.

- Das DAI veröffentlicht u.a. verschiedene Kapitalmarktstatistiken und Studien
- Besonders Hilfreich: DAI Factbooks

### **Börsen-Zeitung**

- Die Börsenzeitung (WM Gruppe) hat ein umfangreiches Archiv zu Nachrichten, Aktien, Fonds, Anleihen, Derivate, IPOs, u.v.m.
- Mit einem Abo ist der Zugang zu allen Daten kostenlos

**DGAP.DE**

- Die Deutsche Gesellschaft für Ad-hoc-Publizität mbH verbreitet Ad-hoc-Mitteilungen für börsennotierte Unternehmen (Pflichtpublizität)
- Gute Quelle für Meldungen aller Art (Achtung: nicht alle börsennotierte Unternehmen nutzen die DGAP, aber 90%?)
- Daten sind im XML-Format auf Anfrage verfügbar
- Ab ~2000, teilweise bis 1996/1997 zurück

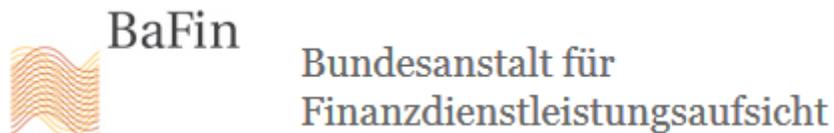
## 2. Kostenfreie Datenbanken/-quellen: Auswahl (3)

8

### Stehle-Datenbank: Fama/French & Benchmark Daten HUMBOLDT-UNIVERSITÄT ZU BERLIN



- <https://www.wiwi.hu-berlin.de/professuren/bwl/bb/data>
- Fama/French Faktoren und Benchmarkdaten: 6 Faktorenssets, Decile- und 4x4-Portfolios, 2x3 size-BM & size-MOM Portfolios, Breakpoints



- Die BaFin führt Datenbanken zu Unternehmen, Directors Dealings, Bedeutende Stimmrechte nach WpHG, Vertraglich gebundene Vermittler nach KWG, Hinterlegte Prospekte, etc.
- Auf der Internetseite (<http://www.bafin.de>) kostenlos verfügbar, teilweise nur ein Jahr rückwirkend, auf Anfrage aber komplette Daten.



- Verschiedene makroökonomische Zeitreihen, inkl. Zeitreihen zu Kapitalmärkten: CDAX, DAX, REX u. a.

## 2. Kostenfreie Datenbanken/-quellen: Auswahl (4)

9

Ungeeignet für wissenschaftliche Zwecke sind m. E.:

- finanzen.net, arriva.de, wallstreet-online.de, Yahoo-Finance (nützlich: EXCEL-Plug-in), usw.
- Die Verwendung dieser Datenquellen ist unüblich in wissenschaftlichen Arbeiten
- Wobei:
  - Zur Prüfung von Einzelfällen (Kurse etc.) durchaus hilfreich
  - Die Foren enthalten teilweise hilfreiche Details zu einzelnen Unternehmen(-sgeschichten, Stichpunkt: Neuer Markt), insbesondere wallstreet-online.de
  - Ausreichende Qualität für die Verwendung in der Lehre, Belegarbeiten, etc.

### 3. Kostenpflichtige Daten(banken): Einführung

10

- Die großen kommerziellen Anbieter wie Thomson Reuters, Bloomberg und S&P Capital IQ offerieren eine Vielzahl von Daten/-banken, primär an Unternehmen der Finanzbranche
- Große, auf Forscher ausgerichtete Datenlieferanten, die eine Vielzahl von Daten/-banken vereinen bzw. den Zugang zu verschiedenen Datenbanken unter einem Dach ermöglichen sind u.a. WRDS (Wharton Research Data Services) und IODS (INSEAD OEE Data Services).
- Zugang bieten die finanzstarken, großen Universitäten, die mehrere Lehrstühle in Finance haben; in Deutschland u.a.:
  - Johann Wolfgang Goethe-Universität Frankfurt am Main
  - Universität Mannheim
  - LMU München
  - Humboldt-Universität zu Berlin
- Die Kosten für die Nutzung variieren stark

### 3. Kostenpflichtige Datenbanken/-quellen: Auswahl (1)

11



- Deutsche Börse Datashop: <http://datashop.deutsche-boerse.com/>
- Alle Daten (auch Tickdaten), direkt von der Quelle
- Kostenlos: Liste aller handelbaren Werte (täglich aktualisiert)

#### **KKMDB - Karlsruher Kapitalmarktdatenbank**

- Adresse: <http://fmi.fbv.kit.edu/>
- Tägliche Kurse aller Aktien und Börsen zurück bis 1974, teilweise bis 1960
- Datenlieferant: Deutsche Börse AG
- Aber: Datenfehler werden nicht behoben
- Wichtigste wissenschaftliche Datenbank bis ~2000
- Ursprung des DAFOX (Deutsche Aktien-Forschungsindex)
- Datenlieferung per Anfrage, Kosten: Selbstkosten
- Gut dokumentiert, siehe Bühler/Göppl/Möller (1993), Göppl/Schütz (1995), Herrmann (1996),

### 3. Kostenpflichtige Datenbanken/-quellen: Auswahl (2)

12



- Serviceanbieter für Datenbanken in Finance, spezialisiert auf Forscher
- ...Vorstellung der DB folgt im Anschluss von Herrn Davydoff



THOMSON REUTERS™

- Unterhält eine Vielzahl von Datenbanken, bekannt durch Datastream (Börsendaten), ThomsonOne (Firmendaten), Worldscope (Bilanzdaten)
- M.E. die derzeit wichtigste und am häufigsten genutzte Datenquelle außerhalb der USA (CRSP), insbesondere für internationale Studien
- Vorteil: Verknüpfung von Aktien- & Bilanzdaten (einheitliche Identifikation)
- Daten ab 1970, aber nur geringe Abdeckung bis 1988
- Problem: Segmentzugehörigkeit fehlt (nur das derzeitige ist verfügbar)
- EXCEL-Plug-in erleichtert Datenauswahl und -sammlung
- Forschungsarbeiten zur Qualität: Brückner (2013), Ince/Porter (2006)

### 3. Kostenpflichtige Datenbanken/-quellen: Auswahl (3)

13

#### Bloomberg

- Mischung aus Datenlieferant und Nachrichtenagentur; unter Bloomberg Professional sind alle Finanzdienstleistungen zusammengefasst
- Finanzdaten jeder Art (Aktien, Anleihen, Derivate, Rohstoffe) mit Tiefgang
- EXCEL-Plug-in erleichtert Datenauswahl und –sammlung

#### Stehle-Datenbank: Aktiendaten

HUMBOLDT-UNIVERSITÄT ZU BERLIN



- Aktiendaten Frankfurt: Amtlicher Markt (seit 1953), Geregelter Markt (seit 1987), Neuer Markt; damit längste Datenbank für BRD
- Monatliche Daten bis 1995 wurden “hand-collected”
- Hohe Datenqualität, Vergleiche wurden u.a. zu KKMDB und Datastream durchgeführt; kontinuierliche Verbesserung
- Einige Details finden sich in Brückner/Lehmann/Schmidt/Stehle (2014)

### 3. Kostenpflichtige Datenbanken/-quellen: Auswahl (4)

14



- Weltweit führender Anbieter von Firmeninformationen, international mit einigen Schwerpunkten in bspw. deutsche Unternehmen
- Mehrere Produkte, größtenteils Länderbezogen



- Spezialisiert auf lange Zeitreihen zu Aktien, Anleihen, Wechselkursen, Inflation, Rohstoffen, ....
- Internationale Datenbank, Fokus auf USA und UK



- „Die“ Datenbank für wissenschaftliche Arbeiten zum US-amerikanischen Kapitalmarkt, Zugang über WRDS
- Großteil der Veröffentlichungen in den großen Journals basieren auf CRSP

# 4. Abschlussbemerkungen

15

- Grundsätzlich gilt: Umso weiter die benötigten Daten in die Vergangenheit reichen (müssen), umso
  - schwieriger sind sie zu finden;
  - schlechter die Qualität;
  - geringer die Wahrscheinlichkeit, dass sie digital verfügbar sind.
- Überprüfung der Datenqualität von Datenbanken findet gar nicht oder nur sehr selten statt bzw. ist selten dokumentiert (Ausführungen hierzu folgen in Abschnitt III).
- Oftmals haben verschiedene Anbieter die selben Datenlieferanten
- Zunehmender Wettbewerb unter den Anbietern

## **Übersichten/Filter zu Datenbanken/-quellen:**

- <http://www.bvdinfo.com/de-de/unsere-produkte/produktfilter>
- <http://www.zew.de/de/links/lifa/statpat.php>
- [http://fisher.osu.edu/cgi-bin/DB\\_Search/db\\_search.cgi?setup\\_file=finance.setup.cgi](http://fisher.osu.edu/cgi-bin/DB_Search/db_search.cgi?setup_file=finance.setup.cgi)
- <http://pages.stern.nyu.edu/~adamodar/> (private Datensammlung)



## II. Data Center SFB649 der Humboldt-Universität



HUMBOLDT-UNIVERSITÄT ZU BERLIN



# 1. Sonderforschungsbereich 649 der Humboldt-Universität zu Berlin

17

Humboldt-Universität zu Berlin >> School of Business and Economics

**Sonderforschungsbereich 649: ÖKONOMISCHES RISIKO**  
Collaborative Research Center 649: ECONOMIC RISK

**About the CRC 649**    **::: About the CRC 649**

**Publications**    We live in a world full of uncertainty. Uncertainty about the future influence present decisions. Financial market fluctuations threaten household wealth, real estate, or the value of bonds and shares.

**Awards**

**Events**

**Projects**

**Guests**

**People**

**RDC**



**News**

**3rd April 2014**  
Our new monthly newsletter is online.  
→ [more](#)

→ [former news](#)

**Events**

**7th - 9th May 2014**  
Energy Finance Workshop 2014  
→ [more](#)

**18th - 20th May 2014**  
Workshop  
"Methods and Challenges in Financial Risk Measurement"  
→ [more](#)

## Research Data Center (RDC)

- Link: <http://sfb649.wiwi.hu-berlin.de>
- Finanziert durch die DFG
- Zugang zu Datenbanken, Software und Rechenleistung



## 2. SFB 649: Zugang und Leistungen

18

- **Verfügbare Datenbanken:**
  - Datastream/Worldscope, ThomsonOne, Compustat, Bloomberg, Creditreform, Macrobond, u.v.a.
- **Installierte Software auf Leistungsstarken Rechnern:**
  - MATLAB, Mathematica, Stata 11SE, SAS, Gauss 10, C/C++ Compiler, R
- **Zugang zu Datenbanken und Software über Remotedesktopverbindung via VPN**
- **Externe Forscher willkommen**
  - Online „Bewerbungsformular“ für Gäste
  - Gastforscherprogramm (Übernahme von Reisekosten, Übernachtung, Tagespauschale) für „young researchers“ (max. 40 Jahre)



### III. Datenqualität und Datenverarbeitung

# 1. Datenqualität und deren Überprüfung

- Es gibt einige wenige wissenschaftliche Beiträge, die sich mit der Datenqualität einzelner Datenbanken beschäftigen:
  - Datastream: Brückner (2013), Ince/Porter (2006)
  - CRSP: Rosenberg/Houglet (1974), Bennin (1980), Courtenay/Keller (1994), Shumway (1997), Shumway/Warther (1999)
- Datenqualität spielt m. E. eine wesentliche Rolle in der empirischen Kapitalmarktforschung, die jedoch oftmals unterschätzt wird.
- Datenfehler können das Forschungsergebnis beeinflussen, möglicherweise völlig verändern
- Klassiker: survivorship bias, fehlende Kapitalmaßnahmen
- Die Datenqualität ist i.d.R.
  - bei großen Unternehmen besser als bei kleinen;
  - für weit in der Vergangenheit liegende Zeiträume (vor 1990) schlechter als für die jüngste Zeit (nach 2000);
  - bei älteren, etablierten Datenbanken besser als bei Newcomer

## 2. Beispiel Aktien – Renditebestandteile

21

Bestandteile bzw. zu beachtende Kapitalmaßnahmen:

- Dividenden (Bonusdividenden)
- Nennwertumstellungen
- Bezugsrechtsemissionen
- Kapitalerhöhungen aus Gesellschaftsmitteln
- Körperschaftssteuergutschrift (1977-2000)
- ...

Werden diese nicht oder nur teilweise berücksichtigt werden Renditen unterschätzt (selten überschätzt).

Auf der Ebene von Einzelaktien kann dies zu großen Verwerfungen führen (bspw. versäumen eines Aktiensplits von 1:10 → Rendite -90%, wobei real 0%)

Aber auch auf Portfolioebene: Eine Aktienanlage in deutsche Aktien erzielte je nach Definition (geometrisches Mittel, 1954-1992, siehe Stehle/Schmidt (2014)):

- FTS-Index, Performance-Index mit Körperschaftssteuergutschrift: **11.38%**
- FTS-Index, Performance-Index ohne Körperschaftssteuergutschrift **10.74%**
- Index des Statistischen Bundesamtes, reiner Preisindex: **7.08%**

## 2. Beispiel Aktien – Renditebestandteile/Datastream

22

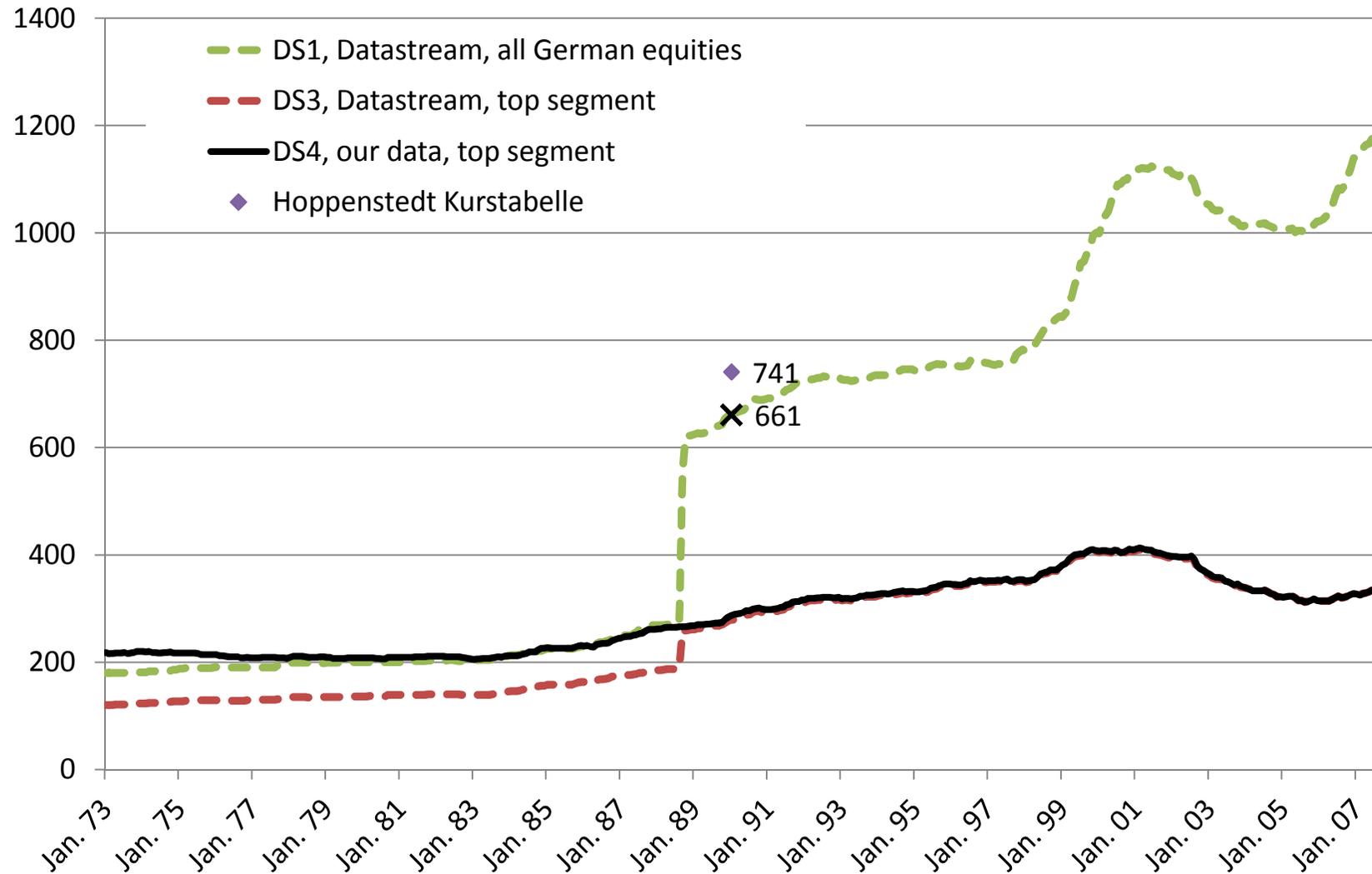
- Es ist sehr aufwendig alle Kapitalmaßnahmen zu sammeln („Der Teufel steckt im Detail“)
- Abhilfe schaffen moderne Datenbanken wie Datastream:
  - ▣ Datentypen wie „PI“ (Preisindex), RI („Total Return Index“), u.v.a.
  - ▣ Aber: Datenqualität sollte immer hinterfragt werden
  - ▣ Brückner (2013) untersuchte die Datenqualität von Datastream und verglich sie mit der Stehle-DB.

Ergebnisse:

- „[W]e cannot recommend Datastream as the primary data source before 1990”  
Gründe: unvollständige Marktabdeckung, survivorship bias, systematisch fehlende Dividenden/Kapitalmaßnahmen, u.a.
- „Only after 1990, when serious data errors are rare, the data quality is sufficient for standard academic use“  
Dennoch: Auch nach 1990 fehlende/falsche Dividenden/Kapitalmaßnahmen, aber nicht mehr systematisch

## 2. Beispiel Aktien – Datastream

23



Quelle: Brückner (2013)

### 3. Beispiel Fama/French Faktoren: Einführung

- Es gibt eine Reihe von Anbietern für Fama/French Faktoren für den deutschen Aktienmarkt, was i. G. ein Glücksfall ist, da Forscher eine Wahl haben und die Robustheit der Ergebnisse prüfen können.
- Aber: Qualität der Datensätze ist sehr verschieden
- Brückner/Lehmann/Schmidt/Stehle (2014) vergleichen die Anbieter:

Autoren	Artmann/Finter/ Kempf/Koch/ Theissen (2012)	Brückner/ Lehmann/ Schmidt/Stehle (2013)	Hanauer/ Kaserer/Rapp (2012)	Frazzini (2013)	Schmidt/Arx/ Schrimpf/ Wagner/Ziegler (2011)	Sandy Lai (2011)
Universitäten	Köln (CFR)	Humboldt	München (TUM), Marburg	Stern School NYU AQR Capital Man.	ETH Zürich	Singapore Management
Zeitraum	07/1962 - 12/2011	07/1958 - 06/2013	08/1996 - 01/2012	01/1986 - 12/2013	07/1984 - 06/2012	07/1981 - 12/2010

Sowie:

- Marktrendite von Ken French and CDAX
- MSCI (Faktoren basierend auf style indices für Deutschland)
- Marmi/Poma, Scuola Normale Superiore, Pisa

### 3. Beispiel Fama/French Faktoren: Vergleich (I)

25

Anbieter	R <sub>m</sub> -R <sub>f</sub>	SMB	HML	MOM
Our: TOP [BP:0.5]	0.46	-0.32	0.42*	1.32***
Our: ALL [BP:TOP]	0.42	-0.48*	0.64**	1.34***
Our: ALL [BP: 0.5]		-0.80***	0.76***	1.05**
Schmidt et al. [BP:0.5]	0.39	-0.28	0.55**	1.31***
Schmidt et al. [BP:0.8]		-0.27	0.77***	1.27**
Hanauer et al.	0.56	-0.71**	0.74***	1.19**
Artmann et al.	0.53	-0.55*	0.85***	1.44***
Marmi/Poma	0.31	-0.56*	1.13***	1.60***
Frazzini	0.28	-0.03	0.67*	1.24**
MSCI	0.45	-0.09	0.24	
CDAX	0.40			

Die Tabelle zeigt arithmetische Mittelwerte von Renditezeitreihen (monatlich, in %) vom Marktportfolio, SMB, HML und MOM verschiedener Anbieter.

Bemerkungen:

- R<sub>f</sub>=Monatsgeld Frankfurt
- Zeitraum: 1996/07 – 2011/06
- \*\*\*/\*\*/\* zeigt Signifikanz auf dem Niveau .01/.05/.10
- BP = Breakpoint für die Size-Sortierung.
- Unsere Daten enthalten keine Körperschaftssteuergutschrift

Quelle: Brückner/Lehmann/Schmidt/Stehle (2014)

### 3. Beispiel Fama/French Faktoren: Vergleich (II)

26

Größte Unterschiede zwischen den Zeitreihen (monatliche Renditen):

- Momentum, Okt. 2001: 30.36 %  
(Mittelwert: 1.3, Sd 6; Schmidt et al.: -40.71, Unsere Schätzung: -10.35)
- SMB, Nov. 1998: 11.69 %  
(Mittelwert 0.5, Sd 4; Hanauer et al.: 4.41, Unsere Schätzung : -7.28)
- HML, Okt. 2003: 11.54 %  
(Mittelwert 0.7, Sd 4; Artmann et al.: 1.93, Schmidt et al.: 13.47)
- Market Portf., Nov. 1998: 10.92 %  
(Mittelwert 0.5, Sd 6; Artmann et al.: 10.89, Hanauer et al.: -0.03)

Note: Marmi/Poma, Frazzini, the MSCI approximation are not included in this comparison.

## 4. Abschlussbemerkungen

27

Nach meinen Erfahrungen:

- Datenerhebung und -verarbeitung braucht *immer* länger als erwartet
- Der Teufel steckt im Detail: Die ersten 80% der Daten (Aufbereitung, Zusammenstellung, etc.) sind *schnell* erhoben, die letzten 5-10% brauchen 80% der Zeit.
- Investition in die Datenqualität lohnt sich („rubbish in, rubbish out“)
- Sehr gute Kenntnisse der Datenverarbeitung und -analyse inkl. Software- und Programmierkenntnissen (SPSS, SAS, STATA, R, etc.) sind essenziell; EXCEL allein reicht nicht
- Verwendete Daten anschauen, durchsuchen, analysieren,... („auf den Kopf stellen“)
- Keine Datenbank ist perfekt, Fehler finden sich überall

# IV. Zusammenfassung/Fazit

28

## Daten in Finance

- Unüberschaubare Vielfalt an Datenbanken/-quellen
- Frei verfügbare Daten sind selten für wissenschaftliche Zwecke geeignet
- Datenqualität spielt eine wesentliche Rolle in der empirischen Kapitalmarktforschung
- Probleme in der Datenerhebung werden oftmals unterschätzt

## Fazit

- Kooperation der FHs mit großen Universitäten, die Datenbankzugänge haben?
- Kostenpflichtige Datenbanken eventuell ALS Gruppe bestellen?
- Für Einzelkämpfer ist ein vorsichtiger Einstieg in die empirische Kapitalmarktforschung angezeigt: Thema mit geringen Datenanforderungen
- Daten austauschen empfehle ich sehr, eventuell Aufgabe des Arbeitskreises

Ende der Präsentation.

Vielen Dank.



Martin H. Schmidt  
Humboldt University Berlin  
School of Business and Economics  
Office: Dorotheenstr. 1, 10117 Berlin  
Mail: Unter den Linden 6, 10099 Berlin  
Phone: +49 30 2093 5761  
Email: [martin.h.schmidt@wiwi.hu-berlin.de](mailto:martin.h.schmidt@wiwi.hu-berlin.de)  
Web: <http://lehre.wiwi.hu-berlin.de/Professuren/bwl/bb/>

# References (I)

- Artmann, S./Finter, P./Kempf, A./Koch, S./Theissen, E. (2012): "The Cross-Section of German Stock Returns: New Data and New Evidence", *Schmalenbach Business Review*, 64, 20-43.
- Asness, C./Frazzini, A. (2013): "The Devil in HML's Details", *Journal of Portfolio Management*, 39(4), 49-68.
- Bennin, R. (1980): "Error Rates in CRSP and COMPUSTAT: A Second Look", *Journal of Finance*, 35(5), 1267-1271.
- Brückner, R. (2013): "Important Characteristics, Weaknesses and Errors in German Equity Data from Thomson Reuters Datastream and their Implications for the Size Effect", *Working Paper*.
- Brückner R./Lehmann, P./Schmidt, M. H./Stehle, R. (2013): "Fama/French Factors for Germany: Which Set Is Best?", *Working Paper*, Available at SSRN: <http://ssrn.com/abstract=2390063>.
- Brückner, R./Lehmann, P./Stehle, R. (2012): "In Germany the CAPM is Alive and Well", *Working Paper*.
- Brückner, R./Stehle, R. (2012): "Der Geregelte Markt Frankfurt: Ein ökonomischer Nachruf", *Working Paper*.
- Bühler, W. / Göppl, H. / Möller, H. (1993): "Die Deutsche Finanzdatenbank (DFDB)", *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung, Sonderheft 31: Empirische Kapitalmarktforschung*, 287-331.
- Courtenay, S.M. / Keller, S.B. (1994): "Errors in Databases Revisited: An Examination of the CRSP Shares-Outstanding Data", *The Accounting Review*, 69(1), 285-291.
- Eun, C. S./Lai, S./de Roon, F. A./Zhang, Z. (2010): "International Diversification with Factor Funds", *Management Science*, 56, 1500-1518.
- Göppl H. / Schütz, H. (1995): "Die Konzeption eines Deutschen Aktienindex für Forschungszwecke (DAFOX)", *Working Paper*.
- Hanauer, M./Kaserer, C./Rapp, M. S. (2012): "Risikofaktoren und Multifaktormodelle für den deutschen Aktienmarkt", *Betriebswirtschaftliche Forschung & Praxis (forthcoming)*.
- Hau, H./Lai, S. (2011): "The Role of Equity Funds in the Financial Crisis Propagation", Available at SSRN: <http://ssrn.com/abstract=1785514>.

# References (II)

- Herrmann, R. (1996): "Die Karlsruher Kapitalmarktdatenbank - Bilanz und Ausblick -", *Working Paper*.
- Ince, O.S. / Porter, R.B. (2006): "Individual Equity Return Data From Thomson Datastream: Handle with Care!", *Journal of Financial Research*, 29(4), 463-479.
- Lehmann, P. (2013): "More Than Fifty Years on Stage: The Performance of German Stock Mutual Funds", published in "Three Essays on the Returns of German Stocks and Mutual Funds", Dissertation, Hamburg: Dr. Kovač.
- Rosenberg, B. / Houglet, M. (1974): "Error Rates in CRSP and Compustat Data Bases and their Implications", *Journal of Finance*, 29(4), 1303-1310.
- Schmidt, P. S./Schrimpf, A./von Arx, U./Wagner, A. F./Ziegler, A. (2011): "On the Construction of Common Size, Value and Momentum Factors in International Stock Markets: A Guide with Applications", Swiss Finance Institute Research Paper, No. 10-58.
- Shumway, T. (1997): "The Delisting Bias in CRSP Data", *The Journal of Finance*, 52(1), 327-340.
- Shumway, T. / Warther, V.A. (1999): "The Delisting Bias in CRSP's Nasdaq Data and Its Implications for the Size Effect", *The Journal of Finance*, 54(6), 2361-2379.
- Stehle, R./Schmidt, M. H.(2014): "Historical Returns on German Stocks 1954 - 2013", *Working Paper*.

# Appendix