

Schreiber, Sven et al.

Research Report

Verfahren der konjunkturellen Wendepunktbestimmung unter Berücksichtigung der Echtzeit-Problematik: Welche Indikatoren geben frühzeitig zuverlässige Signale?

IMK Study, No. 27

Provided in Cooperation with:

Macroeconomic Policy Institute (IMK) at the Hans Boeckler Foundation

Suggested Citation: Schreiber, Sven et al. (2012) : Verfahren der konjunkturellen Wendepunktbestimmung unter Berücksichtigung der Echtzeit-Problematik: Welche Indikatoren geben frühzeitig zuverlässige Signale?, IMK Study, No. 27, Hans-Böckler-Stiftung, Institut für Makroökonomie und Konjunkturforschung (IMK), Düsseldorf, <https://nbn-resolving.de/urn:nbn:de:101:1-201211143755>

This Version is available at:

<https://hdl.handle.net/10419/106241>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Study

Sven Schreiber | Thomas Theobald | Christian Proaño* | Sabine Stephan | Katja Rietzler** | Daniel Detzer

Gutachten des Instituts für Makroökonomie und Konjunkturforschung (IMK) in der Hans-Böckler-Stiftung, im Auftrag des Bundesministeriums der Finanzen (Projekt I A 3 – 09/10)

Verfahren der konjunkturellen Wendepunktbestimmung unter Berücksichtigung der Echtzeit-Problematik

Welche Indikatoren geben frühzeitig
zuverlässige Signale?

* The New School for Social Research, New York

** Rietzler Economics, Berlin



Institut für Makroökonomie
und Konjunkturforschung
Macroeconomic Policy Institute

**Gutachten des Instituts für Makroökonomie und
Konjunkturforschung (IMK) in der
Hans-Böckler-Stiftung, im Auftrag des
Bundesministeriums der Finanzen (Projekt I A 3 – 09/10)**

Verfahren der konjunkturellen Wendepunktbestimmung unter Berücksichtigung der Echtzeit-Problematik

Welche Indikatoren geben frühzeitig zuverlässige Signale?

Sven Schreiber, Thomas Theobald, Christian Proaño*,
Sabine Stephan, Katja Rietzler^b, Daniel Detzer

Veröffentlichter Endbericht als IMK Study, Mai 2012

*The New School for Social Research, New York

^bRietzler Economics, Berlin

Institut für Makroökonomie und Konjunkturforschung (IMK)
in der Hans-Böckler-Stiftung
Hans-Böckler-Str. 39
40476 Düsseldorf
Telefon: 0211-7778-331
Fax: 0211-7778-266

Ansprechpartner:
PD Dr. Sven Schreiber
Telefon: 0211-7778-332
Fax: 0211-7778-4332
E-Mail: Sven-Schreiber@boeckler.de

Inhaltsverzeichnis

1	Einleitung	12
2	Datenbasis für die Wendepunktbestimmung in Echtzeit	16
2.1	Kurzer Datenüberblick	16
2.2	Echtzeitdatenauswahl und -konstruktion	17
2.3	Aufbereitung der Echtzeitdaten	23
	Anhang 2.A Weiteres Material	28
3	Deskriptive Analyse der Datenrevisionen	30
3.1	Produktion	31
3.1.1	Verteilungen	31
3.1.2	Korrelationsanalyse	37
3.2	Auftragseingänge aus dem Inland	39
3.2.1	Verteilungen	39
3.2.2	Korrelationsanalyse	44
3.3	Auftragseingänge aus dem Ausland	44
3.3.1	Verteilungen	44
3.3.2	Korrelationsanalyse	47
3.4	Konsumentenpreise	47
3.4.1	Verteilungen	47
3.4.2	Korrelationsanalyse	51
3.5	Korrelationen zwischen den Revisionen der betrachteten Indikatoren . . .	55
3.6	Zusammenfassung	55
4	Datierung von konjunkturellen Wendepunkten in Echtzeit	58
4.1	Verbreitete Methoden zur Datierung von konjunkturellen Wendepunkten .	58
4.2	Das Bry-Boschan und das Harding-Pagan Verfahren	61

4.3	Implementierung nicht-parametrischer Wendepunktdatierung in Echtzeit	66
4.3.1	Auswahl der Referenz-Zeitreihe	66
4.3.2	Detaillierte Beschreibung des Datierungsalgorithmus	70
4.3.2.1	Behandlung von Ausreißern	71
4.3.2.2	Identifikation von lokalen Hoch- und Tiefpunkten der Referenz-Zeitreihe	71
4.4	Evaluierung von Echtzeit-Aspekten	72
4.5	Zusammenfassung	79
5	Wendepunktbestimmung mit einem linearen multivariaten Einzelgleichungsansatz	87
5.1	Einleitung	87
5.2	Voruntersuchungen für die Variablenauswahl	90
5.2.1	Bivariate Schätzungen	90
5.2.2	Analyse der Korrelation zwischen den Indikatoren	99
5.3	Multivariate Prognosegleichungen	102
5.4	Schätzung der Rezessionswahrscheinlichkeit	105
5.5	Zusammenfassung	108
Anhang 5.A	Weiteres Material	109
5.A.1	Zusammensetzung des Composite Leading Indicator der OECD für Deutschland	109
5.A.2	Variablenauswahl	109
6	Wendepunktbestimmung mit Subset-VAR-Modellen	115
6.1	Einleitung	115
6.2	VAR-Modellrahmen und Modellreduktion	118
6.3	In-Sample-Eigenschaften	120
6.4	Relevante Prediktoren	122
6.5	Schätzung der Rezessionswahrscheinlichkeit	126
6.5.1	Niveau-Spezifikation	126
6.5.2	Spezifikation in Wachstumsraten	127
6.6	Out-of-Sample-Evaluation der Wendepunktprognosen	127
6.7	Schlussfolgerungen	133

7	Wendepunktbestimmung mit Markov-Switching-Modellen	136
7.1	Einleitung	136
7.2	Markov-Switching und exogene Variablen	138
7.3	Modellspezifizierung und In-sample-Ergebnisse	142
7.3.1	Datenauswahl	142
7.3.2	Modellspezifizierung	146
7.3.3	In-sample-Evaluation	148
7.3.4	Likelihood-Ratio-Test in einem MS-Modell mit 4 Regimen	156
7.3.5	Rangfolge der Frühindikatoren	157
7.4	Out-of-sample-Evaluation	158
7.4.1	Modellanpassung in Echtzeit	158
7.4.2	Echtzeitvorhersagen	163
7.5	Zusammenfassung	165
8	Wendepunktbestimmung mit dynamischen Probit-Modellen	169
8.1	Einleitung	169
8.2	Dynamische Probit-Modelle	170
8.3	Rezessionsprognosen mit Probit-Modellen in Echtzeit	174
8.3.1	Datensatz	174
8.3.2	Modellspezifikation	175
8.4	Schätzergebnisse	176
8.4.1	In-Sample-Evaluierung	177
8.4.2	Out-of-Sample-Evaluierung	195
8.5	Schlussfolgerungen	196
9	Fazit	200

Abbildungsverzeichnis

3.1	Echtzeitdatenmatrix der deutschen Industrieproduktion	31
3.2	Echtzeitdatenmatrix Revisionsdarstellung	32
3.3	Produktion – Anteil und Höhe der Revisionen	33
3.4	Produktion – kumulierte Revisionen	36
3.5	Produktion – Autokorrelation der ersten Revisionen, mit Konfidenzbändern	38
3.6	Auftragseingänge aus dem Inland – Anteil und Höhe der Revisionen . . .	41
3.7	Auftragseingänge aus dem Inland – kumulierte Revisionen	43
3.8	Auftragseingänge aus dem Inland – Autokorrelation der ersten Revisionen, mit Konfidenzbändern	45
3.9	Auftragseingänge aus dem Ausland – Anteil und Höhe der Revisionen . .	46
3.10	Auftragseingänge aus dem Ausland – kumulierte Revisionen	48
3.11	Auftragseingänge aus dem Ausland – Autokorrelation der ersten Revisio- nen, mit Konfidenzbändern	49
3.12	Konsumentenpreise – Anteil und Höhe der Revisionen	50
3.13	Konsumentenpreise – kumulierte Revisionen	53
3.14	Konsumentenpreise – Autokorrelation der ersten Revisionen, mit Konfi- denzbändern	54
4.1	Probleme der Wendepunkterkennung	63
4.2	Stilisierte Abbildung einer Rezessionsphase	65
4.3	Darstellung des Einzelhandelsumsatzes und der Erwerbstätigenstunden .	68
4.4	Ökonomische Gesamtaktivität, BIP-Quartalsdaten vs. monatliche Produk- tionsdaten	70
4.5	Produktion des Produzierenden Gewerbes ohne Bau	74
4.6	Wendepunkt datierung – Seitenansicht	76
4.7	Wendepunkt datierung – Aufsicht	77

4.8	Wendepunkt datierung im spezifischen Zyklus inländischer Auftragseingänge – Seitenansicht	80
4.9	Wendepunkt datierung im spezifischen Zyklus inländischer Auftragseingänge – Aufsicht	81
4.10	Wendepunkt datierung im spezifischen Zyklus ausländischer Auftragseingänge – Seitenansicht	82
4.11	Wendepunkt datierung im spezifischen Zyklus ausländischer Auftragseingänge – Aufsicht	83
4.12	Ökonomische Gesamtaktivität, Produktion des Produzierenden Gewerbes	84
5.1	Rezessionswahrscheinlichkeiten für die 1-Schritt- und die 2-Schritt-Prognose	106
5.2	Rezessionswahrscheinlichkeiten für die 3-Schritt-, die 4-Schritt- und die 5-Schritt-Prognose	107
5.3	Rezessionswahrscheinlichkeiten für Durchschnitte aus den Einzelprognosen	108
6.1	Prognosefehler für log-Niveau-Spezifikation	121
6.2	Prognosefehler für Wachstumsraten-Spezifikation	121
6.3	Out-of-sample-Rezessionswahrscheinlichkeiten (Rev. 0)	129
6.4	Out-of-sample-Rezessionswahrscheinlichkeiten (Durchschnitte, Rev. 0)	130
6.5	Out-of-sample-Rezessionswahrscheinlichkeiten (Rev. 1)	132
6.6	Out-of-sample-Rezessionswahrscheinlichkeiten (Durchschnitte, Rev. 1)	133
7.1	Wachstumsraten rückwärts geglätteter Industrieproduktion	144
7.2	Rezessionswahrscheinlichkeiten der MS-Spezifikationen und Verlauf der eingebetteten Frühindikatoren	153
7.3	Fortsetzung von Abbildung 7.2	154
7.4	Einmonats-Echtzeitvorhersagen für Regimewahrscheinlichkeiten der Industrieproduktion	166
8.1	In-Sample-Anpassung der geschätzten und durchschnittlichen Rezessionswahrsch., Einmonats-Prognosehorizonte	178
8.2	In-Sample-Anpassung der geschätzten und durchschnittlichen Rezessionswahrsch., Zweimonats-Prognosehorizonte	179

8.3	In-Sample-Anpassung der geschätzten und durchschnittlichen Rezessionswahrsch., Dreimonats-Prognosehorizonte	180
8.4	Beispiel der geschätzten und durchschnittlichen Rezessionswahrscheinlichkeiten	184
8.5	Echtzeit-Durchschnittsrezessionswahrscheinlichkeiten	198
9.1	Vergleich der Rezessionssignale bezogen auf die Publikationsperiode . . .	204

Tabellenverzeichnis

2.1	Variablen, Teil 1	18
2.2	Variablen, Teil 2	19
2.3	Echtzeitindikatoren und ihre Zusammensetzung	22
2.4	Basisjahrwechsel der Echtzeitindikatoren	25
2.5	Größenordnung des Erfassungsbruchs zur Veröffentlichung November 2005	26
2.6	Ergebnisse der ADF-Tests, Teil 1	28
2.7	Ergebnisse der ADF-Tests, Teil 2	29
3.1	Produktion – deskriptive Statistik zu den Revisionsdaten	34
3.2	Produktion – Durchschnitt der kumulierten Revisionen	35
3.3	Produktion – Korrelationskoeffizienten verschiedener Revisionen zu einem Berichtsmonat	38
3.4	Korrelation zwischen den Revisionen des Produktionsindikators und den Auftragseingängen aus dem Inland	39
3.5	Korrelation zwischen den Revisionen des Produktionsindikators und den Auftragseingängen aus dem Ausland	40
3.6	Korrelation zwischen den Revisionen des Produktionsindikators und den Konsumentenpreisen	40
3.7	Auftragseingänge aus dem Inland – deskriptive Statistik zu den Revisionsdaten	42
3.8	Auftragseingänge aus dem Inland – Durchschnitt der kumulierten Revisionen	43
3.9	Auftragseingänge aus dem Inland – Korrelationskoeffizienten verschiedener Revisionen zu einem Berichtsmonat	44
3.10	Auftragseingänge aus dem Ausland – deskriptive Statistik zu den Revisionsdaten	47
3.11	Auftragseingänge aus dem Ausland – Durchschnitt der kumulierten Revisionen	48

3.12	Auftragseingänge aus dem Ausland – Korrelationskoeffizienten verschiedener Revisionen zu einem Berichtsmonat	51
3.13	Konsumentenpreise – deskriptive Statistik zu den Revisionsdaten	52
3.14	Konsumentenpreise – Durchschnitt der kumulierten Revisionen	52
3.15	Konsumentenpreise – Korrelationskoeffizienten verschiedener Revisionen zu einem Berichtsmonat	53
3.16	Korrelationsmatrizen der Revisionen eines Berichtsmonats	56
4.1	Korrelationskoeffizienten zwischen den Wachstumsraten des Bruttoinlandsprodukts und ausgewählten Variablen	69
5.1	Mögliche Lagstrukturen in Abhängigkeit von der Datenverfügbarkeitslücke	91
5.2	Bivariate Schätzungen für t-1 (backcast), Sample: 1993m2-2010m9	94
5.3	Bivariate Schätzungen für t (nowcast), Sample: 1993m2-2010m9	95
5.4	Bivariate Schätzungen für t+1, Sample: 1993m2-2010m9	96
5.5	Bivariate Schätzungen für t+2, Sample: 1993m2-2010m9	97
5.6	Bivariate Schätzungen für t+3, Sample: 1993m2-2010m9	98
5.7	Ausgewählte Korrelationen zwischen Frühindikatoren	100
5.8	Ausgewählte Korrelationen zwischen dem CLI und weiteren Unterreihen	100
5.9	Ausgewählte Korrelationen zwischen Finanzmarktindikatoren	101
5.10	Ausgewählte Regressoren für die jeweiligen Prognosehorizonte	104
5.11	Zusammensetzung des Composite Leading Indicator der OECD für Deutschland	109
5.12	Bivariate Schätzungen für t-1 (backcast), Sample: 1993m2-2006m11	110
5.13	Bivariate Schätzungen für t (nowcast), Sample: 1993m2-2006m11	111
5.14	Bivariate Schätzungen für t+1, Sample: 1993m2-2006m11	112
5.15	Bivariate Schätzungen für t+2, Sample: 1993m2-2006m11	113
5.16	Bivariate Schätzungen für t+3, Sample: 1993m2-2006m11	114
6.1	Test auf Normalverteilung der Prognosefehler	122
6.2	Regressoren in den Produktionsgleichungen, log-Niveau-Spezifikationen .	124
6.3	Regressoren in den Produktionsgleichungen, Wachstumsraten-Spezif. . .	125
6.4	Out-of-sample-Evaluation (Vergleich mit ex-post-Benchmark) der VAR-Spezifikationen	134

7.1	BDS-Test für die Wachstumsraten der geglätteten Industrieproduktion . . .	145
7.2	Zusammenfassung der MS-Regressionen mit 4 Regimen – Konstante und switchender autoregressiver Term	150
7.3	Zusammenfassung der MS-Regressionen mit 4 Regimen – Restterme . . .	151
7.4	Güte der Anpassung verschiedener MS Spezifikationen	155
7.5	Rangfolge der Frühindikatoren gemäß der Standardmaße von Tabelle 7.4	158
7.6	Änderung der Koeffizienten und Welch-Testresultate der Veröffentlichung August 2008 – Teil 1	160
7.7	Änderung der Koeffizienten und Welch-Testresultate der Veröffentlichung Aug. 2008 – Fortsetzung von Tabelle 7.6	161
7.8	Maße für die Vorhersagegüte der verschiedenen MS-Spezifikationen . . .	164
8.1	Dynamische Probitregressionen, Ein-Monats-Prognosehorizont, Rechen- stand 2010m9	181
8.2	Dynamische Probitregressionen, Zwei-Monats-Prognosehorizont, Rechen- stand 2010m9	182
8.3	Dynamische Probitregressionen, Drei-Monats-Prognosehorizont, Rechen- stand 2010m9	183
8.4	Likelihood-Ratio-Tests PBY1CG 1MF – Teil 1	186
8.5	Likelihood-Ratio-Tests PBY1CG 1MF – Fortsetzung	187
8.6	Likelihood-Ratio-Tests PBY1CS 1MF – Teil 1	188
8.7	Likelihood-Ratio-Tests PBY1CS 1MF – Fortsetzung	189
8.8	Likelihood-Ratio-Tests PBY10CG 3MF – Teil 1	190
8.9	Likelihood-Ratio-Tests PBY10CG 3MF – Fortsetzung	191
8.10	Likelihood-Ratio-Tests PBY10CS 3MF – Teil 1	192
8.11	Likelihood-Ratio-Tests PBY10CS 3MF – Fortsetzung	193
8.12	Einfluss der einzelnen Regressoren auf die latente Variable im Probit-Modell beim Ein-Monats-Prognosehorizont, Rechenstand 2010m9	194
8.13	Out-of-sample Prognose-Evaluierung in Echtzeit	197
9.1	Modellvergleich	202

1 Einleitung

Eine der größten Herausforderungen für die Konjunkturforschung ist es, konjunkturelle Wendepunkte frühzeitig und zuverlässig zu erkennen. Während in stabilen Auf- und Abschwungphasen die zu erwartende Wirtschaftsentwicklung mangels Richtungswechsel recht treffsicher vorhergesagt werden kann, wird der Anfang oder das Ende einer Rezession –wie auch eines Aufschwungs– oft erst spät oder gar erst rückblickend erkannt, was eine zielgerichtete und stabilisierende Wirtschaftspolitik erheblich erschwert.

Dass die Erkennung konjunktureller Wendepunkte am aktuellen Rand besonders schwierig ist, liegt nicht nur an der Spezifikation vieler Algorithmen, die zur Datierung von Wendepunkten sowohl verzögerte als auch zukünftige Beobachtungen der Indikatorvariablen benötigen, sondern auch an der großen Unsicherheit, mit der makroökonomische Zeitreihen am aktuellen Rand behaftet sind. Diese Unsicherheit rührt zum einen daher, dass „harte“ realwirtschaftliche Indikatoren wie Industrieproduktion, Auftragseingänge oder Preisindizes erst mit einer gewissen zeitlichen Verzögerung veröffentlicht werden. Das heißt, dass es bei diesen Indikatoren eine Datenverfügbarkeitslücke (publication lag) von mehreren Monaten gibt. Zum anderen unterliegen diese Indikatoren immer wieder beträchtlichen Revisionen und es kann viele Monate dauern, bis der Revisionsprozess abgeschlossen ist. Aus diesem Grund kann sich das Konjunkturbild, das diese Indikatoren in „Echtzeit“ zeichnen, deutlich von dem unterscheiden, das sich retrospektiv auf Basis der endgültigen Daten ergibt (vgl. z.B. Stark & Croushore 2002).

Angesichts der Echtzeit-Problematik stellt sich die Frage, ob man sich bei der Erkennung bzw. bei der Prognose konjunktureller Wendepunkte verstärkt auf Indikatoren stützen sollte, die diese Probleme nicht haben. Welche Indikatoren kommen dann in Frage? Zum einen kommen „weiche“ Indikatoren wie Umfrageergebnisse in Betracht. Weder die ifo- noch die ZEW-Indikatoren haben einen publication lag und sie werden auch nicht revidiert. Allerdings bilden sie die wirtschaftliche Entwicklung nur indirekt –z.B. über Stimmungen oder Erwartungen– ab, weshalb a priori unklar ist, ob sie besser oder schlechter für die Wendepunktbestimmung geeignet sind als „harte“ Indikatoren. Zum anderen stehen Fi-

nanzmarktindikatoren wie Zinsen oder Wertpapierindizes zur Verfügung, die als „hart“ anzusehen sind, aber nicht revidiert werden. Inwieweit diese Indikatoren jedoch frühzeitig verlässliche Signale hinsichtlich der zukünftigen realwirtschaftlichen Entwicklung liefern, ist in der Literatur umstritten (vgl. z.B. Estrella et al. 2003, Rudebusch & Williams 2009). Das bedeutet, dass zur Klärung dieser Frage eine empirische Untersuchung notwendig ist.

Das Institut für Makroökonomie und Konjunkturforschung in der Hans-Böckler-Stiftung hat diese Frage im Auftrag des Bundesministeriums der Finanzen im Rahmen des Projekts IA3 – 09/10 eingehend untersucht und legt hiermit den Endbericht zu diesem Forschungsprojekt vor. Die Untersuchung wurde anhand von vier verschiedenen Modellklassen durchgeführt mit dem Ziel, eine Kombination von Schätzmethode und Indikatorenauswahl zu identifizieren, die sich am besten für die Prognose von Wendepunkten eignet. Dabei wurde der Echtzeit-Problematik explizit Rechnung getragen. Das heißt, dass sowohl für die jeweiligen Modellspezifikationen als auch für die Prognoseevaluierung immer nur der Datenstand verwendet wurde, der zum jeweiligen Zeitpunkt t tatsächlich verfügbar war. Das impliziert für die Modellspezifikation, dass die Lagstruktur der verwendeten Indikatoren die Indikator spezifischen Datenverfügbarkeitslücken berücksichtigt und dass die Prognoseevaluierung unter Echtzeit-Bedingungen durchgeführt wurde.

Das vorliegende Gutachten gliedert sich in einen „Daten“-Teil (Kapitel 2 bis 4) und einen „Modell“-Teil (Kapitel 5 bis 8). Kapitel 2 und 3 geben einen detaillierten Überblick über die verwendeten Daten, wobei das besondere Augenmerk auf den Variablen liegt, die von der Echtzeit-Problematik betroffen sind. Echtzeitdaten werden typischerweise in Form von Dreiecksmatrizen dargestellt, d.h. dass die Datenmengen beträchtlich größer sind als bei Zeitreihen. Damit sind sie aber auch anfälliger für Fehler. Aus diesem Grund wurden die Echtzeitdaten eingehend auf Konsistenz geprüft. Kapitel 2 gibt einen ausführlichen Einblick in den Umgang mit wichtigen Problemen wie Datenlücken und Benchmark-Revisionen. Im Anschluss daran erfolgt in Kapitel 3 eine statistische Untersuchung der beobachteten Revisionen mit einfachen deskriptiven Methoden. Kapitel 4 befasst sich mit verschiedenen Ansätzen zur Operationalisierung des Konzepts der konjunkturellen Wendepunkte, d.h. wie sie konkret bestimmt und datiert werden können. Nach der grundsätzlichen Erörterung dieser Frage wird dort auch untersucht, wie die Methoden an Echtzeitbedingungen angepasst werden können, und welche konkreten Auswirkungen die Verwendung von Echtzeitdaten auf die Beurteilung der Konjunktur in Deutschland im Zeitverlauf gehabt hätte. Das gewählte nicht-parametrische Datierungsverfahren liefert zudem die binäre Benchmark-Zeitreihe, die die durch Wendepunkte getrennte Auf- und

Abschwungphasen anzeigt. Diese Benchmark-Reihe wird in den späteren Kapiteln zur Evaluierung der Rezessionsprognosen der verschiedenen angewendeten Modelle verwendet.

Wie eingangs erwähnt, werden in dieser Untersuchung vier verschiedene Modellklassen verwendet: Einzelgleichungsansatz, Subset-VAR-Modell, Markov-Switching-Modell und dynamisches Probit-Modell. Für eine erste Orientierung und bessere Einordnung der verschiedenen Modelle ist es sinnvoll, bereits an dieser Stelle auf wesentliche Unterschiede bzw. auf Gemeinsamkeiten der Modelle hinzuweisen: Bei den ersten beiden Modellen handelt es sich um lineare Verfahren und die prognostizierte Zielgröße ist die Wachstumsrate bzw. das Niveau der gesamtwirtschaftlichen Outputgröße. Das heißt, dass man mit diesen Modellen die Rezessionswahrscheinlichkeit nicht direkt vorhersagen kann. Vielmehr muss man sie in einem zweiten Schritt über die Dichtefunktionen der Prognosen bestimmen. Markov-Switching und dynamisches Probit-Modell hingegen sind nichtlineare Verfahren, mit denen man die Rezessionswahrscheinlichkeit direkt vorhersagen kann.

Der „Modell“-Teil beginnt mit dem einfachsten Modelltyp, einem Ansatz, der auf linearen Einzelgleichungen beruht (Kapitel 5). Für jeden der Prognosehorizonte (Ein-Schritt-Prognose, Zwei-Schritt-Prognose, ...) wird eine eigene Prognosegleichung für die Wachstumsrate der Outputgröße spezifiziert. Anschließend werden die Dichtefunktionen der prognostizierten Werte in geeigneter Weise kombiniert, um daraus die Wahrscheinlichkeit zu bestimmen, dass das prognostizierte Wachstum eine vordefinierte Schwelle, die eine Rezession markiert, unterschreitet. Im anschließenden Kapitel 6 werden ebenfalls lineare Schätzmodelle und deren Dichteprognosen verwendet. Jedoch werden dort vektorautoregressive (VAR) Modelle herangezogen, bei denen die Mehr-Schritt-Prognosen rekursiv bestimmt werden und implizit alle aufgenommenen Variablen prognostiziert werden. Da diesen VARs diverse Nullrestriktionen für die Koeffizienten auferlegt werden, spricht man auch von *subset*-Modellen. Die nächsten beiden Kapitel 7 und 8 wenden sich dann nichtlinearen Verfahren zu, die explizit ein „Regime“-Konzept beinhalten. Zunächst werden Markov-Switching-Modelle (Kapitel 7) spezifiziert, bei denen die abhängige Variable zwar nach wie vor das Wachstum der Outputgröße ist, die Gleichungsparameter (und dadurch auch die mittelfristige Wachstumsrate) aber von dem zu schätzenden Regimeverlauf abhängen. Deshalb können die Rezessionswahrscheinlichkeiten bei diesem Modelltyp direkt abgeleitet werden. Anschließend befasst sich Kapitel 8 mit dynamischen Probit-Modellen, die ökonometrisch direkt die binäre Regimevariable als zu erklärende Variable verwenden.

den.¹ Das abschließende Kapitel 9 führt die Ergebnisse der vier „Modell“-Kapitel zusammen und nimmt eine Bewertung der verschiedenen Modelle hinsichtlich ihrer Eignung für eine frühzeitige und verlässliche Erkennung konjunktureller Wendepunkte vor, wobei das Kriterium der Eignung zwei verschiedene Dimensionen umfasst: Zum einen lässt sich vergleichen, wie groß die Übereinstimmung zwischen den prognostizierten Rezessions- und Aufschwungswahrscheinlichkeiten und der später ex post festgestellten tatsächlichen makroökonomischen Entwicklung ist. Zum anderen ist es wünschenswert, dass das Modell möglichst früh ein Signal sendet, dass ein Wendepunkt naht oder vor kurzem stattgefunden hat.

¹Eine frühere Version dieses Kapitels war bereits im Zwischenbericht enthalten. Seitdem wurden die Revisionshistorien der Variablen berücksichtigt und die Spezifikation angepasst.

2 Datenbasis für die Wendepunktbestimmung in Echtzeit

Im Folgenden wird die Datenbasis beschrieben, die zur Bearbeitung des Forschungsauftrags herangezogen wurde. Die ersten Abschnitte behandeln die verwendeten Variablen, die Datenquellen sowie einige grundsätzliche Merkmale der Daten. Im Abschnitt 2.3 werden dann jene Merkmale der Daten erläutert, die eine Aufbereitung erforderlich machen. Dazu gehören Datenlücken, die Berücksichtigung von Basisjahrumstellungen sowie weitere Sachverhalte, die unter dem Begriff der Benchmark-Revision zusammengefasst werden können.

2.1 Kurzer Datenüberblick

Für die Konjunkturanalyse und -prognose steht eine Vielzahl möglicher Indikatoren zur Verfügung. Diese reichen von „harten“ Indikatoren der Realwirtschaft wie den Auftragseingängen über Finanzmarktindikatoren wie Zinsen oder Wertpapierindizes bis hin zu „weichen“ Indikatoren wie den Ergebnissen von Konjunkturumfragen. Welcher Umfang an Indikatoren angemessen ist, hängt von der konkreten Zielsetzung der Analyse ab. Für Faktormodelle (vgl. z.B. Forni et al. 1999, Schumacher & Breitung 2008) oder Pooling (Drechsel & Scheufele 2010) ist eine möglichst umfangreiche Datenbasis sinnvoll. In der vorliegenden Studie interessiert jedoch nicht die Güte einer Prognose, die auf einer maximalen Zahl von Indikatoren beruht, sondern vor allem der Beitrag einzelner Indikatoren für eine Vorhersage, die auf einem sparsam spezifizierten Modell beruht. Hier scheint es sinnvoll, alle wesentlichen Kategorien von Indikatoren mit den wichtigsten Zeitreihen abzudecken.

In der aktuellen Literatur zu Prognosen der Industrieproduktion werden sehr unterschiedliche Herangehensweisen bei der Variablenauswahl deutlich. Drechsel & Scheufele (2010) verwenden für eine Prognose der deutschen Industrieproduktion über 90 bzw. 120 Indikatoren. Carstensen et al. (2011) beschränken sich hingegen bei ihrer Untersuchung der Vorhersagequalitäten von Indikatoren für die Industrieproduktion im Euroraum auf lediglich 7 Zeitreihen. Mit 27 Konjunkturindikatoren, die für die Industrieproduktion ohne Bau relevant sein dürften, bewegt sich die aktuelle Studie in der Mitte. Alle wesentlichen Zeitreihen sind erfasst. Ein Gesichtspunkt bei der Variablenauswahl besteht auch in einer leichten Verfügbarkeit und Handhabbarkeit der Daten. Es geht hier zudem nicht um eine einmalige Analyse, sondern um die Entwicklung eines Prognosemodells, das ohne übermäßigen Aufwand auf Dauer regelmäßig eingesetzt werden kann.

Die Tabellen 2.1 und 2.2 geben einen Überblick über die Variablen, ihre Datenquellen, die Datenverfügbarkeitslücke (*publication lag*), also die Zeitspanne zwischen dem Ende der Berichtsperiode und der Erstveröffentlichung der entsprechenden Daten, sowie über die Stationaritätseigenschaften der Zeitreihen. Sie zeigen auch an, ob die jeweilige Reihe von Revisionen betroffen ist. Mit Ausnahme der Kreditvergabe, die erst ab September 1998 vorliegt, sind alle Variablen ab Februar 1993 verfügbar. Zur Ermittlung der Stationaritätseigenschaften wurden ADF-Tests (Fuller 1976, Dickey & Fuller 1979, 1981) für alle Zeitreihen durchgeführt. Die Ergebnisse sind in den Tabellen 2.6 und 2.7 im Anhang dargestellt.

2.2 Echtzeitdatenauswahl und -konstruktion

Zur Früherkennung konjunktureller Wendepunkte eignen sich besonders monatliche Echtzeitdaten. Diese werden in der Regel in Form von Matrizen abgespeichert, so dass zeilenweise die Berichtsmonate, d.h. die Monate, für die der Indikator veröffentlicht wurde, und spaltenweise die Veröffentlichungsmonate, das sind die Monate, in denen der Indikator veröffentlicht wurde, aufgelistet werden.

Für Deutschland gibt es Echtzeitdaten bei der OECD (vgl. OECD 2010) und bei der Deutschen Bundesbank (vgl. Deutsche Bundesbank 2010b). Da die Echtzeitdatenbank der Bundesbank grundsätzlich über längere Veröffentlichungshistorien verfügt, wurden die OECD-Daten der Industrieproduktion ausschließlich dazu verwendet, jene Werte auf Übereinstimmung zu überprüfen, die in beiden Datenbanken für die gleichen Berichts-

2 Datenbasis für die Wendepunktbestimmung in Echtzeit

Tabelle 2.1: Variablen, Teil 1

Beschreibung	Quelle	Abkürzung	Pub'lag (Mon.)	Stationaritäts- eigenschaften	Revi- sionen
Zu erklärende Variable					
Industrieproduktion ohne Bau	Bundesbank	IPROD_OB	2	I(1)	ja
Realwirtschaftliche Indikatoren					
Auftragseingänge Inland	Bundesbank	AE_INL_BUBA	2	I(1)	ja
Auftragseingänge Ausland	Bundesbank	AE_AUSL_BUBA	2	I(1)	ja
Offene Stellen	Bundesbank	VAC	0	I(1)	ja*)
Baugenehmigungen	Bundesbank	BAUGEN	2	I(0) m. Trend	ja*)
Surveys					
Ifo-Index Geschäftsklima	ifo Institut	IFO_R1	0	I(0)	nein
Ifo-Index Geschäftsbeurteilung	ifo Institut	IFO_LAGE	0	I(0)	nein
Ifo-Index Geschäftserwartungen	ifo Institut	IFO_ERW	0	I(0)	nein
ZEW sentiment	ZEW	ZEW_ERW	0	I(0)	nein
ZEW situation	ZEW	ZEW_SIT	0	I(0)	nein
Industrial confidence	DG ECFIN/ Ecowin	INDCONF	1	I(0)	nein
Production expectations	DG ECFIN/ Ecowin	PRODEX	1	I(0)	nein
Economic sentiment	DG ECFIN/ Ecowin	ECSSENT	1	I(0)	nein

Tabelle 2.2: Variablen, Teil 2

Beschreibung	Quelle	Abkürzung	Pub'lag (Mon.)	Stationaritäts- eigenschaften	Revi- sionen
Zusammengesetzte Indikatoren					
OECD Composite Leading Indicator	OECD	CLI	2	I(0) m. Trend	ja
Early Bird Indikator der Commerzbank	Commerzbank	COM_EB	1	I(1)	ja*)
Preise					
Verbraucherpreise	Bundesbank	CPI	1	I(1)	ja
Real effective exchange rate DE	OECD	RAW	2	I(1)	ja*)
Brent, average, ICE, USD	Reuters/ Ecowin	BRENT	0	I(1)	nein
Zinsen und Finanzmarktindikatoren					
CDAX	Bundesbank	CDAX	0	I(1)	nein
Kreditwachstum	Bundesbank	KREDIT	1	I(1)	ja*)
Rentenindex	Bundesbank	REX	0	I(1)	nein
Fibor/ Euribor	Bundesbank	EURIBOR 3M	0	I(1)	nein
Corporate Spread	Bundesbank, eigene Berechnungen	C_SPREAD	0	I(1)	nein
Term-Spread (1y-3m)	Bundesbank - eigene Berechnungen	YC1Y	0	I(0)	nein
Term-Spread (2y-3m)	Bundesbank - eigene Berechnungen	YC2Y	0	I(0)	nein
Term-Spread (3y-3m)	Bundesbank - eigene Berechnungen	YC3Y	0	I(0)	nein
Term-Spread (5y-3m)	Bundesbank - eigene Berechnungen	YC5Y	0	I(0)	nein
Term-Spread (10y-3m)	Bundesbank - eigene Berechnungen	YC10Y	0	I(1)**)	nein

*) Revisionshistorie nicht verfügbar.

***) Wird in Analogie zu den übrigen Term-Spreads als I(0) behandelt.

und Veröffentlichungsmonate vorhanden sind. Diese Übereinstimmung war insgesamt gegeben.¹ Somit dient als primäre Datenquelle die Echtzeitdatenbank der Deutschen Bundesbank. Wie im weiteren Verlauf erläutert wird, sind die dort gespeicherten Echtzeitmatrizen für eine dem Forschungsauftrag angemessene Veröffentlichungshistorie nicht vollständig ausgefüllt. Dies bedingte einerseits die Vervollständigung der Matrizen unter der Annahme, dass keine Datenrevision stattgefunden hat. Andererseits mussten in geringem Maße sekundäre Datenquellen herangezogen werden, welche jedoch ebenfalls auf Datenmaterial der Deutschen Bundesbank beruhen. Dabei handelt es sich um die Statistischen Beihefte – Saisonbereinigte Wirtschaftszahlen, die als Einzelpublikation monatlich von der Deutschen Bundesbank veröffentlicht werden, sowie um Datensätze, die den Autoren Schumacher & Breitung (2008) für ihre Veröffentlichung von der Deutschen Bundesbank zur Verfügung gestellt wurden.²

Als nächstes war die Frage zu klären, für welche Indikatoren und welchen Berichts- und Veröffentlichungsumfang Echtzeitdaten in der vorgenannten Datenbank vorhanden sind. Damit wurden insbesondere die für die Analyse von Datenrevisionen (Kapitel 3) und für die nicht-parametrische Datierung konjunktureller Wendepunkte (Kapitel 4) zu verwendenden Daten festgelegt. Dagegen ist der in den parametrischen Prognosemodellen benutzte Datenumfang weitaus größer, da Frühindikatoren eingesetzt werden können, die nicht revidiert werden. Hierzu zählen Finanzmarktindikatoren (Deutsche Bundesbank 2011) und Umfragen wie ifo- (ifo Institut für Wirtschaftsforschung 2011) und ZEW-Indikatoren (Zentrum für Europäische Wirtschaftsforschung 2011). Zudem werden realwirtschaftliche Größen, die zwar revidiert werden, aber noch nicht in der Echtzeitdatenbank enthalten sind, unter Quasi-Echtzeitbedingungen verwendet. Dies bedeutet, dass die Verzögerung bei der Datenverfügbarkeit bei Modellspezifikationen berücksichtigt wird, jedoch werden mangels vorliegender Daten keine Revisionshistorien verwendet.

In Tabelle 2.3 werden die Indikatoren aufgeführt, deren Echtzeitmatrix eine für den Forschungsauftrag angemessene Veröffentlichungshistorie enthält. Es handelt sich um die Produktion des Produzierenden Gewerbes ohne Bau als Proxy für die ökonomische Ge-

¹Zwar erschienen vereinzelt größere Abweichungen, die sich aber auf die unterschiedliche Berücksichtigung von Benchmark-Revisionen und insbesondere Basisjahrwechsel zurückführen lassen. Siehe dazu auch die Erläuterungen der einschlägigen Sachverhalte in Abschnitt 2.3. Die grundsätzliche Übereinstimmung der Datenbanken ist plausibel vor dem Hintergrund, dass für beide die ursprüngliche Datenermittlung beim Statistischen Bundesamt sowie die Saisonbereinigung bei der Bundesbank angesiedelt ist.

²Wir danken den Autoren für die Weitergabe der Datensätze.

samtaktivität, um in- und ausländische Auftragseingänge als vorlaufende Indikatoren sowie um die Verbraucherpreisentwicklung, die im Allgemeinen als gleichlaufender Indikator eingestuft wird. Des Weiteren sind in der Tabelle enthalten: die verfügbaren Berichts- und Veröffentlichungsmonate, Variablenbezeichnung in der Datenbank der Deutschen Bundesbank sowie Zusammensetzung der Matrix, sofern mehrere Quellen herangezogen wurden.

Vergleicht man Berichts- und Veröffentlichungsmonate je Indikator, so werden die einzelnen Datenverfügbarkeitslücken deutlich, d.h. jene Anzahl von Monaten die zwischen dem Veröffentlichungszeitpunkt und dem letzten Berichtsmonat liegt. Diese Zeit wird benötigt für den Eingang der Meldungen durch die Unternehmen, die Datenaggregation und die Anwendung der Bereinigungsverfahren. Sie beträgt für die Produktion sowie für die in- und ausländischen Auftragseingänge jeweils zwei Monate und für die Verbraucherpreise einen Monat (vgl. Tabellen 2.1 und 2.2).

In der vorliegenden Untersuchung wird die Produktion im Produzierenden Gewerbe ohne Bau als Proxy für den gesamtwirtschaftlichen Output verwendet. Diese Entscheidung soll kurz erläutert werden. Da der Fokus auf der Gesamtwirtschaft liegt, scheiden die Produktion der Vorleistungsgüterproduzenten, der Investitionsgüterproduzenten, der Konsumgüterproduzenten, der Gebrauchsgüterproduzenten und der Verbrauchsgüterproduzenten als maßgebliche Größe aus. Aus dem selben Grund wird die Produktion des Produzierenden Gewerbes ohne Bau der Industrieproduktion vorgezogen, da Erstere zusätzlich die Wertschöpfung des Bergbaus enthält. Gleichwohl ist der Unterschied zwischen diesen Zeitreihen gering und verringert sich mit sinkender Wertschöpfung des Bergbaus immer mehr. Führt man dieses Argument weiter, so erscheint zunächst auch die Berücksichtigung des Baugewerbes angebracht. Allerdings führen Sondereffekte wie bspw. die Veränderung der Wohnungsbauprämie und anderer staatlicher Fördermaßnahmen dazu, dass sich der Konjunkturzyklus des Baugewerbes auf monatlicher Basis vom allgemeinen Konjunkturzyklus abhebt. Es erscheint daher sinnvoll, die Produktion des Produzierenden Gewerbes ohne Bau als maßgebliche monatliche Outputgröße zu verwenden.

Grundsätzlich wurde auf einen Rechenstand pro Veröffentlichungsperiode abgestellt, so dass etwaige Revisionen in unserem Datensatz einmal pro Monat auftreten können. Dies stellt nur eine geringfügige Vereinfachung dar. Wie bereits erwähnt sind die Echtzeitmatrizen der oben genannten Veröffentlichungshistorie nicht vollständig ausgefüllt, da die Deutsche Bundesbank die komplette Veröffentlichungshistorie erst ab Einführung der Echtzeitdatenbank angibt, d.h. ab November 2005. Zu jedem vorherigen Veröffentlichungsdatum

Tabelle 2.3: Echtzeitindikatoren und ihre Zusammensetzung

Indikator	Berichtsmonate	Veröffentlichungs- monate	Variablen- bezeichnung ^a	Anmerkungen
Produktion	02/1993-09/2010	06/1995-11/2010	IP1AA020	bezogen auf das Produzierende Gewerbe ohne Bau ^b
Auftragseingänge Inland	02/1993-09/2010	06/1995-11/2010	IO2ABA20 IO2ACM01	bezogen auf das Verarbeitende Gewerbe bis 08/2001 ^c bezogen auf die Industrieproduktion ab 09/2001 ^c
Auftragseingänge Ausland	02/1993-09/2010	06/1995-11/2010	IO3ABA20 IO3ACM01	bezogen auf das Verarbeitende Gewerbe bis 08/2001 ^c bezogen auf die Industrieproduktion ab 09/2001 ^c
Verbraucherpreise	02/1993-10/2010	01/1995-11/2010	PCIIPC100	aus Schumacher & Breitung (2008) bis 06/2005 aus Statistischen Beihften von 07/2005 bis 10/2005 aus der Echtzeitdatenbank ab 11/2005

^aDie Variablenbezeichnung bezieht sich auf die Echtzeitdatenbank der Deutschen Bundesbank.

^bDer Zusammenhang zwischen der Produktion des Produzierenden Gewerbes ohne Bau und anderen monatlichen Outputgrößen (bspw. der Industrieproduktion) wird im Folgenden erläutert.

^cDer Bezugswechsel ist hinsichtlich der Auswirkungen auf die Bruttowertschöpfung vernachlässigbar. Grundsätzlich lautet die Teilengenbeziehung gemäß Klassifizierung der Deutschen Bundesbank: Verarbeitendes Gewerbe ⊂ Industrie ⊂ Produzieren-
des Gewerbe.

sind nur Revisionen für die jeweils ca. 30 vorangehenden Berichtsmonate angegeben. Allerdings handelt es sich dabei genau um den Teil der Revisionen, der besonders relevant ist.³ Fehlende Rechenstände wurden unter der Annahme ergänzt, dass keine Revision stattgefunden hat.

Alle verwendeten Daten sind kalender- und saisonbereinigt gemäß den von der Deutschen Bundesbank benutzten Verfahren.⁴ Da zudem sowohl das Statistische Bundesamt als auch die Deutsche Bundesbank auf Anfrage eine Veränderung der Bereinigungsverfahren als maßgebliche Ursache für die im Folgenden als Benchmark-Revisionen bezeichneten Sachverhalte ausschließt, waren diesbezüglich keine weiteren Korrekturen an den Daten notwendig. Mögliche Auswirkungen der Saisonbereinigung werden auch im Kapitel 3 angesprochen.

2.3 Aufbereitung der Echtzeitdaten

Der wichtigste Grund für die Notwendigkeit einer Datenaufbereitung sind die so genannten Benchmark-Revisionen. Knetsch & Reimers (2006) fassen hierunter Basisjahrwechsel, Änderungen in der Güter- und Sektorklassifizierung sowie solche Umstellungen zusammen, die die Methodik der Datenerhebung betreffen. Allgemein kann man hierunter jede Art substanzieller Änderungen statistischer Konventionen verstehen, so dass die Veränderung der Daten von einer Veröffentlichung zur nächsten nicht dem *normalen* Schätzfehler geschuldet ist. Mit anderen Worten: Die Datenrevision entsteht nicht nur aus der Vorläufigkeit der Daten. Knetsch & Reimers (2006) identifizieren in deutschen Produktions- und Auftragsdaten abgesehen von Basisjahrwechseln zwei Ereignisse, die als Benchmark-Revisionen eingestuft werden können:

1. Juni 1995: Umstellung der Güter- und Sektorklassifizierung auf eine einheitliche europäische Basis, die Nomenclature Général des Activités Économiques dans l'Union Européenne (NACE).
2. April 1999: Umstellung des Berichtskreises in disjunkte Monats- und Quartalsmelder und die damit verbundene Änderung der Unternehmensabdeckung sowie ent-

³Werte von älteren Berichtsmonaten sollten in der Regel bis auf Umbasierung oder ähnliches nicht mehr verändert werden, da dem Statistischen Bundesamt nach einer gewissen Zeit nach Erstveröffentlichung eine sichere Datenlage gegeben ist, inklusive fast aller Nachmeldungen.

⁴Zur Saisonbereinigung setzt die Deutsche Bundesbank das Census-Verfahren X-12-ARIMA ein.

sprechender Aufschätzfaktoren.

Diese Sachverhalte werden im Folgenden als Erklärung für in den Daten festgestellte Brüche herangezogen.⁵ Zunächst soll hier aber die Behandlung der Basisjahrwechsel vorgestellt werden. Im Unterschied zur trivialen Umbasierung bei üblichen Zeitreihenanalysen tritt bei Echtzeitdaten ein Identifikationsproblem auf: Es ist aus der Datenbasis nicht eindeutig zu ermitteln, welcher Teil der Diskrepanz zwischen altem und revidiertem Wert auf den Basisjahrwechsel zurück geht bzw. wie groß die „echte“ Revision ist. Tabelle 2.4 stellt daher die Zeitpunkte und unsere unterstellten Konversionsfaktoren der Basisjahrwechsel dar. Die Konversionsfaktoren ergeben sich als Mittelwert der Verhältnisse aus den Einträgen vor und nach dem Basisjahrwechsel. Die Abweichungen vom jeweiligen Mittelwert werden also als die tatsächlichen Revisionen betrachtet, die nicht dem Basisjahrwechsel geschuldet sind. Ex post ist eine Aufteilung der Veränderungen zwischen den betroffenen Veröffentlichungen in tatsächliche und durch die Umbasierung verursachte Revisionen praktisch unmöglich, zumindest anhand der von der Bundesbank veröffentlichten Informationen.

Die Tatsache, dass der Konversionsfaktor in der Regel größer als 1 ist, ist auf die Trendinflation zurückzuführen. Allerdings können vereinzelt andere Effekte überwiegen. So geht jeder Basisjahrwechsel –und nicht nur der als Benchmark-Revision eingestufte im Zuge der europäischen Harmonisierung in 1995– mit einer Modifizierung der Güter- und Sektorklassifizierung einher.⁶

Besondere Aufmerksamkeit verdient der Konversionsfaktor für die inländischen Aufträge bzw. in abgemilderter Form auch diejenigen für die Produktion und für die ausländischen Aufträge zum Basisjahrwechsel von 2000 auf 2005. Dieser liegt nahe 1 und ist

⁵Die Autoren stellen für die beiden genannten Sachverhalte und den Basisjahrwechsel von 1995 auf 2000 im Februar 2004 fest, dass eine Nichtbeachtung oder eine einfache Methodik wie die reine Umbasierung nicht adäquat zur Behandlung von Benchmark-Revisionen sei. Dies illustrieren sie insbesondere an der durch die Benchmark Revision hervorgerufenen Änderung der Langfristbeziehung zwischen Produktion und Auftragslage in dem von ihnen verwendeten Kointegrationsansatz. Trotz dieser Feststellung benutzen wir eine Umbasierung mit einem einheitlichen Konversionsfaktor – jedoch ausschließlich für die Basisjahrwechsel. Dies lässt sich wie folgt rechtfertigen: 1. Der Ansatz von Knetsch & Reimers (2006) funktioniert nur, wenn für eine ausreichende Veröffentlichungsbreite die gleichen Daten in den verschiedenen Basisjahren vorhanden sind. Solche überlappenden Daten werden jedoch nicht von der Bundesbank veröffentlicht. 2. Der Fokus dieses Forschungsauftrags liegt auf (Kurzfrist-) Prognosen der Wendepunkte am aktuellen Rand, so dass die Konsequenzen für die Langfristbeziehung vernachlässigbar erscheinen.

⁶Vergleiche Metainformationen, die die Deutsche Bundesbank zu jedem veröffentlichten Indikator heraus gibt (Deutsche Bundesbank 2010b).

Tabelle 2.4: Basisjahrwechsel der Echtzeitindikatoren

Indikator	Veröffentlichung	Basisjahrwechsel	Konversionsfaktor
Produktion	Juli 1998	von 1991 zu 1995	0,9649
	Februar 2004	von 1995 zu 2000	1,1767
	März 2009	von 2000 zu 2005	1,0654
Auftragseingänge Inland	Juli 1998	von 1991 zu 1995	0,9286
	Februar 2004	von 1995 zu 2000	1,1012
	März 2009	von 2000 zu 2005	1,0061
Auftragseingänge Ausland	Juli 1998	von 1991 zu 1995	1,0460
	Februar 2004	von 1995 zu 2000	1,5439
	März 2009	von 2000 zu 2005	1,2170
Verbraucherpreise	September 1995	von 1985 zu 1987	1,1162
	November 1995	von 1987 zu 1991	0,9806
	März 1999	von 1991 zu 1995	1,1478
	März 2003	von 1995 zu 2000	1,0691
	März 2008	von 2000 zu 2005	1,0792

gleichzeitig mit einer relativ hohen Standardabweichung verbunden, was aus der Tatsache resultiert, dass für die erste Hälfte der Monatsmonate die Verhältnisse aus neuem und altem Wert größer 1 ist, während die Verhältnisse in der zweiten Hälfte mit etwa gleicher Größenordnung unter 1 liegen. Dies lässt den Schluss zu, dass hier die Umbasierung über den erwarteten Konversionsfaktor den Basisjahrwechsel nur ungenügend abbildet (vgl. auch Knetsch & Reimers 2006). Da es sich hierbei allerdings um einen Einzelfall handelt, wurde auf eine Umstellung der Methodik verzichtet. Statt dessen wurden in der Analyse der Datenrevisionen (z.B. in Kapitel 3) diese Beobachtungen aus der Stichprobe entfernt, um einer sonst möglichen Verzerrung der statistischen Eigenschaften der Revisionen zu begegnen.

Nachdem die Daten aus den oben genannten Quellen importiert, die Echtzeitmatrizen vervollständigt sowie mittels der oben genannten Konversionsfaktoren eine einheitliche Umrechnung auf das Basisjahr 2005 erreicht wurde, wurden die gewonnenen Echtzeitmatrizen einer Plausibilitätskontrolle in Bezug auf Betrag und Klumpenbildung der Revisionen unterzogen. Dabei wurde festgestellt, dass die Revisionen der Veröffentlichung

Tabelle 2.5: Größenordnung des Erfassungsbruchs zur Veröffentlichung November 2005

Indikator	Anzahl der Berichtsmonate	Anzahl großer Veränderungen (Differenz > 0,5)	Mittelwert der Veränderungen
Produktion	151	69	2,4
Inländische Aufträge	151	29	1,1
Ausländische Aufträge	151	35	0,8

November 2005 in verschiedenen Indikatoren⁷ ihrem Betrag nach zu den größten gehören. Zwar repräsentiert die Veröffentlichung einen Wechsel in der Erfassung, da ab diesem Zeitpunkt die Echtzeitdatenbank eingeführt wurde und seitdem die gesamte Veröffentlichungshistorie bis Januar 1991 revidiert wird, während zuvor nur etwa die letzten 30 Berichtsmonate vor Veröffentlichung revidiert wurden.⁸ Allerdings erscheint dieser Wechsel in der Erfassungstechnik auf den ersten Blick keinen inhaltlichen Grund zu liefern, warum eine Benchmark-Revision vorliegen sollte. Im Folgenden zeigt sich jedoch, dass dieser Erfassungsbruch genau den von Knetsch & Reimers (2006) gefundenen Benchmark-Revisionen entsprechen kann. Mit einem einfachen Verfahren können die Auswirkungen der Benchmark-Revisionen in den Echtzeitmatrizen abgemildert werden.

Vergleicht man die Veröffentlichungszeitpunkte der von Knetsch & Reimers (2006) als Benchmark-Revision eingestuften Sachverhalte mit dem verwendeten Veröffentlichungshorizont in Tabelle 2.3, so zeigt sich, dass die NACE-Harmonisierung im Juni 1995 keine Auswirkungen hat, weil der Veröffentlichungshorizont bereits mit den neu klassifizierten Daten beginnt. Allerdings sei daran erinnert, dass die Güter- und Sektorklassifizierung ihrerseits zu jedem Basisjahrwechsel einer Revision unterliegt, so dass sich dieser Sachverhalt sehr wohl auch in den verwendeten Echtzeitmatrizen niederschlägt. Die Änderung des Berichtskreises zum April 1999 liegt ohnehin innerhalb des Veröffentlichungshorizonts.

Um zu verstehen, wie beide Sachverhalte trotz unterschiedlicher Veröffentlichungszeitpunkte in den Erfassungsbruch des November 2005 münden können, spielt die Berichtsgrenze Ende 1998, bis zu der der Erfassungsbruch auftritt, eine entscheidende Rolle. Be-

⁷Besonders bemerkenswert ist, dass der beschriebene Sachverhalt für die Produktionsdaten nicht nur für die Berichtsmonate bis Ende 1998 vorliegt, sondern bis zu diesem Zeitpunkt auch einer Art Konvergenz in der Größe der Abweichungen unterliegt. Diese Konvergenz ist für die Auftragseingänge nicht gegeben, während für die Verbraucherpreise überhaupt kein Erfassungsbruch zu verzeichnen ist.

⁸Das entspricht in etwa dem Informationsbedarf der Statistischen Beihefte – Saisonbereinigte Wirtschaftszahlen.

rücksichtigt man nämlich die Datenverfügbarkeitslücke von zwei Monaten, so ist der April 1999, d.h. der Veröffentlichungszeitpunkt der letzten Benchmark-Revision, nicht mehr fern. Mit anderen Worten: Geht man davon aus, dass die Benchmark-Revisionen korrekt datiert sind, so treten diese in den zur Verfügung stehenden Matrizen zu den genannten Veröffentlichungszeitpunkten ausschließlich für die damals revidierten Berichtsmonate, z.B. etwa 30 Berichtsmonate vor Veröffentlichung April 1999, auf. Für den Rest der Berichtsmonate, d.h. für die älteren, wurde mangels Eintrag in der Echtzeitdatenbank unter der Annahme gearbeitet, dass keine Revision stattfindet, bis zum ersten Mal zum November 2005 die vollständige Berichtshistorie revidiert wurde. Das bedeutet, dass sich die Benchmark-Revisionen, obwohl sie inhaltlich einem anderen Datum zuzuordnen sind, zum Erfassungsbruch im November 2005 verschieben.

Um einen verzerrenden Einfluss der Benchmark-Revisionen zu vermeiden, wurde folgende Methodik gewählt: Für alle Berichtsmonate bis Ende 1998 wurden die feststellbaren Abweichungen des Erfassungsbruchs rückwirkend für alle Veröffentlichungen von Juni 1995 bis Oktober 2005 addiert (nach der vollständigen Umbasierung auf 2005). Somit kann größtenteils ausgeschlossen werden, dass die Benchmark-Revisionen irrtümlich als tatsächliche Revisionen behandelt werden. Zwar besteht dadurch die Gefahr, dass ebenfalls der Teil der Veränderungen zwischen Oktober und November 2005 aus den Daten entfernt wurde, der auf tatsächliche Revisionen zurückzuführen ist. Allerdings würde es sich dabei um den weniger gravierenden Fehler handeln.

Abschließend wurden die erstellten Echtzeitmatrizen noch einmal einer Plausibilitätskontrolle in Bezug auf Größenordnung und Anhäufung der Revisionen unterzogen. Hierzu ergaben sich keine weiteren Anmerkungen.

Anhang 2.A Weiteres Material

In den Tabellen 2.6 und 2.7 sind die Ergebnisse der Einheitswurzeltests dargestellt.

Tabelle 2.6: Ergebnisse der ADF-Tests, Teil 1

Variable	Transformation	Deterministik	Lags	Test-Statistik	Ergebnis
IPROD_OB	log level	Konstante, Trend	1-3	-3,22*	I(1)
	log difference	Konstante	1-2	-5,74***	I(0)
AE_INL_BUBA	log level	Konstante, Trend	1-2	-2,90	I(1)
	log difference	Konstante	1	-8,77***	I(0)
AE_AUSL_BUBA	log level	Konstante, Trend	1-3	-2,59	I(1)
	log difference	Konstante	1-2	6,63***	I(0)
VAC	log level	Konstante	1-3	-2,35	I(1)
	log difference		1-2	-3,92***	I(0)
BAUGEN	log level	Konstante, Trend	1	-5,11***	I(0) m. Trend
	log difference	Konstante	1	-17,50***	I(0)
IFO_R1	log level	Konstante	1-3	-3,80***	I(0)
	log difference		1-2	-4,42***	I(0)
IFO_LAGE	log level	Konstante	1-5	-4,15***	I(0)
	log difference		1-2	-4,39***	I(0)
IFO_ERW	log level	Konstante	1-3	-4,45***	I(0)
	log difference		0	-10,18***	I(0)
ZEW_ERW	level	Konstante	1	-4,07***	I(0)
	difference		0	-8,23***	I(0)
ZEW_SIT	level	Konstante	1-3	-3,42**	I(0)
	difference		1	-4,73***	I(0)
INDCONF	level	Konstante	1-3	-4,51***	I(0)
	difference		1	-4,70***	I(0)
PRODEX	level	Konstante	1-2	-3,82***	I(0)
	difference		1	-6,62***	I(0)
ECSENT	log level	Konstante	1-2	-3,43**	I(0)
	log difference		1	-5,43***	I(0)

Tabelle 2.7: Ergebnisse der ADF-Tests, Teil 2

Variable	Transformation	Deterministik	Lags	Test-Statistik	Ergebnis
CLI	log level	Konstante, Trend	1-6	-4,12***	I(0) m. Trend
COM_EB	log difference	Konstante	1-7	-4,68***	I(0)
	level	Konstante	0	-1,66	I(1)
	difference		0	-14,02***	I(0)
CPI	log level	Konstante, Trend	0	-2,70	I(1)
	log difference	Konstante	0	-15,27***	I(0)
RAW	log level	Konstante	1	-1,69	I(1)
	log difference		0	-11,41***	I(0)
BRENT	log level	Konstante, Trend	1	-3,12	I(1)
	log difference	Konstante	0	-12,11***	I(0)
CDAX	log level	Konstante	0	-2,00	I(1)
	log difference		0	-13,51***	I(0)
KREDIT	level	Konstante	0	-1,94	I(1)
	difference		0	-10,57***	I(0)
REX	log level	Konstante, Trend	0	-2,42	I(1)
	log difference	Konstante	0	-13,06***	I(0)
EURIBOR 3M	level	Konstante	1	-2,48	I(1)
	difference		0	-7,81***	I(0)
C_SPREAD	level	Konstante	0	-2,20	I(1)
	difference		0	-13,47***	I(0)
YC1Y	level	Konstante	1	-4,27***	I(0)
	difference		0	-11,93***	I(0)
YC2Y	level	Konstante	1	-3,99***	I(0)
	difference		0	-12,24***	I(0)
YC3Y	level	Konstante	1	-3,70***	I(0)
	difference		0	-11,85***	I(0)
YC5Y	level	Konstante	1	-3,26**	I(0)
	difference		0	-11,25***	I(0)
YC10Y	level	Konstante	1	-2,73*	I(1)
	difference		0	-11,06***	I(0)

3 Deskriptive Analyse der Datenrevisionen

In diesem Kapitel werden die grundlegenden statistischen Eigenschaften der beobachteten Datenrevisionen dargestellt und diskutiert. Im Vergleich zu regulären Indikatorzeitreihen, die immer nur die aktuellste Version eines Indikators wiedergeben, können durch Echtzeitdatenbanken die verfügbaren Informationsstände zu verschiedenen Zeitpunkten analysiert werden. Echtzeitdaten werden üblicherweise in Form einer zweidimensionalen Matrix dargestellt. Für die deutsche Industrieproduktion ist eine solche Matrix in Abbildung 3.1 dargestellt. In den Zeilen sind die Werte des Indikators für den jeweiligen Berichtsmonat angegeben. Die Spaltenbeschriftung gibt den Zeitpunkt an, zu welchem diese Daten von der Bundesbank veröffentlicht wurden (Veröffentlichungsmonat), beschreibt also den verfügbaren Informationsstand zum jeweiligen Zeitpunkt.¹ Der jeweils unterste Wert einer Spalte stellt eine Erstveröffentlichung dar. Bewegt man sich in der gleichen Zeile nach rechts, erhält man die revidierten Werte zum gleichen Berichtsmonat.

Eine Revision ist die Veränderung eines Wertes für einen bestimmten Berichtsmonat von einem Rechenstand zum jeweils nächsten. Da die hier untersuchten Daten nicht stationär sind, wurden die Revisionen relativ zum Wert der vorherigen Veröffentlichung berechnet, d.h. eine Revision gibt die prozentuale Veränderung des Wertes eines Berichtsmonats zu einem Rechenstand t zum vorherigen Rechenstand $t - 1$ an.

Für die Produktion sowie für die Auftragseingänge aus dem In- und Ausland standen Daten für die Berichtsmonate 1993m2 bis 2010m9, sowie die monatlichen Rechenstände von 1995m6 bis 2010m11 zur Verfügung. Daraus ergibt sich für die Berichtsmonate vor 1995m4 (wegen der Datenverfügbarkeitsverzögerung von zwei Monaten) jeweils die Revisionshistorie über 185 Rechenstände. Für die nachfolgenden Daten reduziert sich die Anzahl der Rechenstände mit jedem zusätzlichen Monat um eins. Für die Echtzeitdaten der

¹Alternative Bezeichnungen für Veröffentlichungsmonat ist Datenstand, Rechenstand oder Vintage.

Vintage/Rechenstand, Zeitpunkt zu dem die Daten bekannt wurden

	2010-05-07	2010-06-08	2010-07-08	2010-08-06	2010-09-08	2010-10-07	2010-11-08
2009-07	92,0	92,0	92,8	92,8	92,9	92,9	93,1
2009-08	93,6	93,6	94,3	94,3	94,5	94,5	94,5
2009-09	96,8	96,8	96,8	96,8	96,6	96,6	96,8
2009-10	95,2	95,2	95,7	95,7	95,9	95,9	95,9
2009-11	95,9	95,9	96,5	96,5	96,3	96,3	96,2
2009-12	95,1	95,1	95,9	95,9	95,9	95,9	95,7
2010-01	96,1	96,1	97,7	97,7	97,6	97,6	97,3
2010-02	95,9	95,9	97,4	97,4	97,3	97,3	97,0
2010-03	98,4	98,9	99,9	99,9	99,9	99,9	99,9
2010-04		99,7	101,0	101,0	101,1	100,9	100,9
2010-05			103,9	104,3	104,3	104,1	104,3
2010-06				103,8	103,7	103,6	103,7
2010-07					103,7	103,6	103,8
2010-08						105,5	105,5
2010-09							104,7

Abbildung 3.1: Echtzeitdatenmatrix der deutschen Industrieproduktion. Quelle: Deutsche Bundesbank, eigene Aufbereitung.

Preise standen die Berichtsmonate 1993m2 bis 2010m10 und Rechenstände von 1995m1 bis 2010m11 zur Verfügung. Daraus ergibt sich die maximale Anzahl von 191 Rechenständen.

Betrachtet man die Daten nun nicht nach Rechenständen, wie in Abbildung 3.1 dargestellt, sondern sortiert nach Revisionen, so dass die Revisionen nach ihrem Zeitpunkt relativ zum Berichtsmonat gruppiert werden, so kann man einige interessante Erkenntnisse gewinnen. Die veränderte Matrix ist zum besseren Verständnis als Abbildung 3.2 dargestellt. Die jeweiligen Zeilen beziehen sich weiterhin auf einen Berichtsmonat. In den Spalten sind die Revisionen nach ihrem Zeitpunkt relativ zum Berichtsmonat angeordnet.

3.1 Produktion

3.1.1 Verteilungen

Betrachtet man zunächst die Anzahl der Revisionen in Abbildung 3.3, so kann man klar erkennen, dass fast alle veröffentlichten Werte im Monat nach der Erstveröffentlichung nochmals revidiert werden und auch noch bis zu 4 Monaten nach der Erstveröffentlichung eine über 40%ige Wahrscheinlichkeit besteht, dass noch Anpassungen vorgenommen wer-

3 Deskriptive Analyse der Datenrevisionen

Vintage/Rechenstand, Zeitpunkt zu dem die Daten bekannt wurden

	1. Revision		2. Revision		3. Revision		4. Revision		5. Revision		6. Revision	
	IPROD_OB	Rev _{t+3}	IPROD_OB	Rev _{t+4}	IPROD_OB	Rev _{t+5}	IPROD_OB	Rev _{t+6}	IPROD_OB	Rev _{t+7}	IPROD_OB	Rev _{t+8}
2009-06		0,65%		0,22%		0,11%		0,00%		0,00%		0,00%
2009-07		0,00%		0,22%		0,00%		-0,22%		0,00%		0,00%
2009-08		0,32%		0,00%		-0,11%		0,00%		0,00%		0,00%
2009-09		0,41%		-0,21%		0,00%		0,00%		0,10%		0,00%
2009-10		-0,21%		0,00%		0,00%		0,21%		0,00%		0,00%
2009-11		0,00%		0,00%		0,21%		0,00%		0,00%		0,63%
2009-12		1,72%		0,32%		0,00%		0,00%		0,84%		0,00%
2010-01		-0,21%		0,00%		0,00%		1,66%		0,00%		-0,10%
2010-02		-0,10%		0,00%		1,56%		0,00%		-0,10%		0,00%
2010-03		0,51%		1,01%		0,00%		0,00%		0,00%		0,00%
2010-04		1,30%		0,00%		0,10%		-0,20%		0,00%		
2010-05		0,38%		0,00%		-0,19%		0,19%				
2010-06		-0,10%		-0,10%		0,10%						
2010-07		-0,10%		0,19%								
2010-08		0,00%										

Abbildung 3.2: Echtzeitdatenmatrix Revisionsdarstellung. Quelle siehe Abb. 3.1.

den. Dies kann durch das Eingehen neuer Informationen bei den statistischen Behörden erklärt werden. Zum Zeitpunkt der Erstveröffentlichung liegt oftmals nur ein Teil der gesamten Stichprobe vor, auf deren Grundlage der entsprechende Wert für den Berichtsmonat errechnet wird. Mit der Vervollständigung der Stichprobe besteht dann oftmals noch Revisionsbedarf. Auch wenn der Anteil an Revisionen im Zeitablauf immer geringer wird, bleibt er relativ lange auf einem hohen Niveau von über 10%. Dies lässt sich dadurch erklären, dass es sich bei den Zeitreihen um saisonbereinigte Werte handelt. Um einen Wert zu bereinigen, werden sowohl Werte vor als auch nach dieser Veröffentlichung benötigt. Da letztere erst im Zeitablauf vorliegen, kann sich durch die Änderung der sogenannten Bereinigungskomponente auch für sehr alte Veröffentlichungen noch Revisionsbedarf ergeben (Deutsche Bundesbank 1999). Interessanterweise steigt zur 12ten und zur 24sten Revision der Anteil revidierter Wert nochmals an. Dies ergibt sich durch die Tatsache, dass nach einem Jahr deutlich mehr Informationen über das Saisonmuster vorliegen und dies zu einer starken Änderung in der Saisonkomponente führen kann (Rietzler 2001). Neben diesem Rückgang der Anzahl an Revisionen, kann man aus Abbildung 3.3 auch ersehen, dass sich die durchschnittliche absolute Revisionshöhe im Zeitverlauf reduziert.

In Tabelle 3.1 ist die deskriptive Statistik zu den Revisionen der Produktionswerte der kompletten Reihen zusammengefasst. Vergleicht man zunächst die Stichprobengröße, so sieht man, dass insgesamt Daten über 21824 Revisionen vorliegen. Lässt man allerdings Nullrevisionen unberücksichtigt, d.h. man entfernt die Veröffentlichungen neuer Rechen-

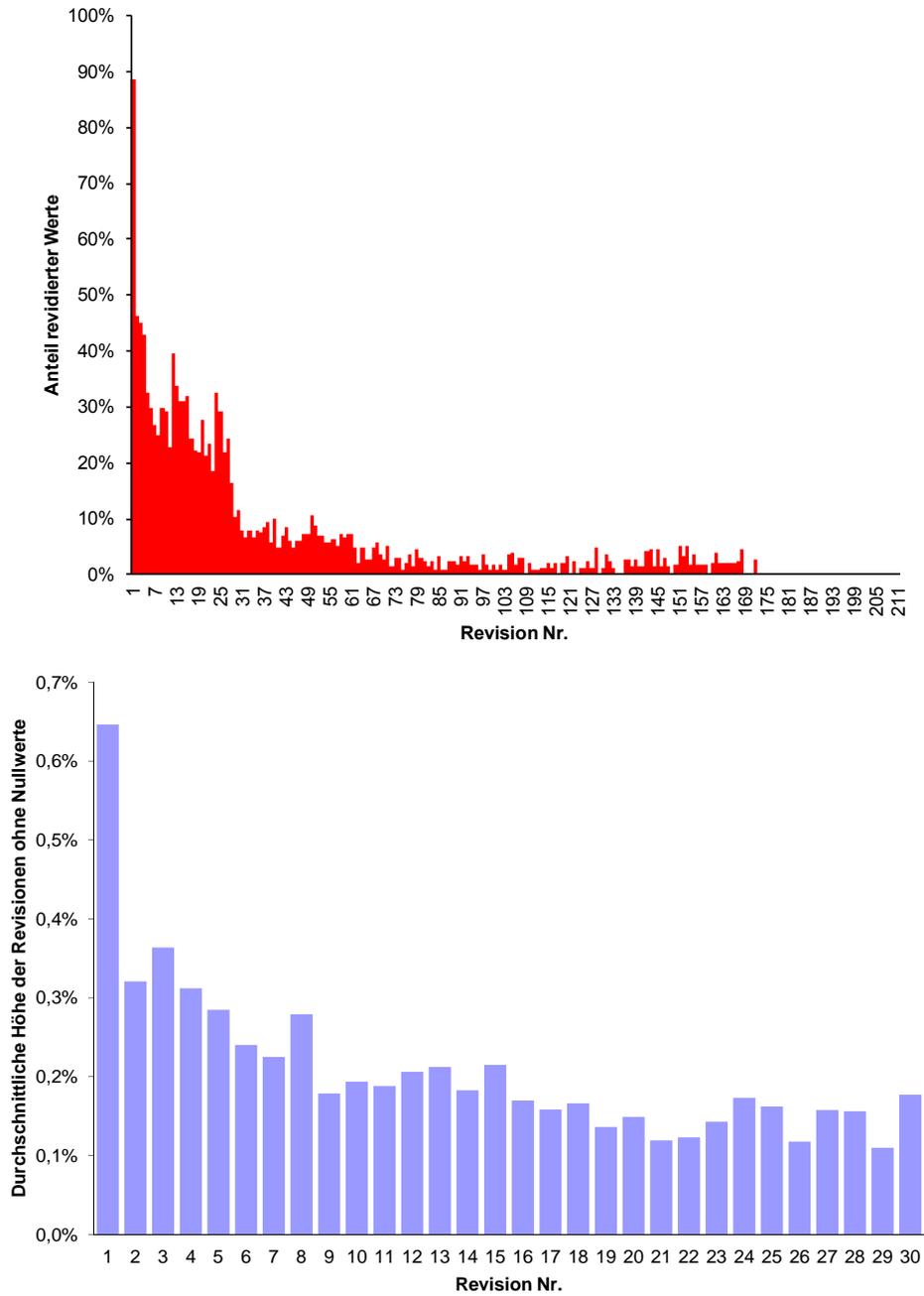


Abbildung 3.3: Produktion im Produzierenden Gewerbe ohne Bau – prozentualer Anteil der tatsächlichen Revisionen und durchschnittliche absolute Höhe der Revisionen.

Tabelle 3.1: Produktion – deskriptive Statistik zu den Revisionsdaten

	kompletter Datensatz		erste 30 Revisionen	
	mit Nullwerten	ohne Nullwerte	mit Nullwerten	ohne Nullwerte
Mittelwert	0,0057%	0,0552%	0,0187%	0,0629%
Standardabw.	0,1154%	0,3561%	0,2145%	0,3898%
Stichprobengröße	21824	2245	5514	1641
Standardfehler	7,81197E-06	7,51508E-05	2,88914E-05	9,62247E-05
t-Wert	7,267	7,344	6,475	6,533
p-Wert	3,7973E-13	2,8958E-13	1,0274E-10	8,5816E-11
Durchschnitt der abs. Revisionen	0,0235%	0,23%	0,0750%	0,25%

stände zu einem Berichtsmonat ohne Veränderung des veröffentlichten Werts, so beträgt die Zahl der tatsächlichen Revisionen mit 2245 nur noch ein Zehntel. Wie bereits erwähnt, finden die meisten Revisionen in den ersten Monaten nach der Erstveröffentlichung statt. Betrachtet man nur die Revisionen, die in den ersten 2,5 Jahren nach der Erstveröffentlichung stattfanden, bestätigt sich dies. Die Stichprobe wurde um 75% reduziert, die Anzahl der darin enthaltenen tatsächlichen Revisionen fiel aber um nur 27%.

Besonders interessant ist, dass alle gebildeten Mittelwerte positiv und statistisch signifikant sind. Geht man davon aus, dass die jeweilige Erstveröffentlichung der statistischen Ämter eine unverzerrte Schätzung des endgültigen Wertes sein sollte, so sollte der Mittelwert der Revisionen Null sein. Nach den gemessenen Mittelwerten zu urteilen scheinen die Erstveröffentlichungen die tatsächlichen Werte aber systematisch zu unterschätzen.

Wie aus der obigen Grafik und Tabelle ersichtlich, findet der Großteil der Revisionen innerhalb der ersten 30 Monate statt. Setzt man hier eine Grenze und geht davon aus, dass nach 30 Monaten der „wahre“ Wert des Index ermittelt ist und es sich bei allen weiteren Revisionen um relativ uninformative Revisionen handelt, so kann man errechnen, wie stark der erstveröffentlichte Wert von diesem „wahren“ Wert abweicht. Diese Berechnung wurde nur für den Teil der Daten durchgeführt, für die zum jeweiligen Berichtsmonat die Erstveröffentlichung und mindestens 30 Revisionen vorliegen. Diese eingeschränkten Daten beziehen sich demnach auf die Berichtsmonate 1995m4 bis 2008m3. Tabelle 3.2 zeigt

Tabelle 3.2: Produktion – Durchschnitt der kumulierten Revisionen der ersten 30 Rechenstände

	gesamter Zeitraum	seit 2000	vor 2000
Mittelwert	0,51%	0,31%	0,85%
Standardabw.	0,0107	0,0073	0,0143
Stichprobe	156	99	57
Standardfehler	0,00086	0,00073	0,0019
t-wert	5,913	4,226	4,477
p-wert	2,074E-08	5,332E-05	3,769E-05
Minimum	-2,38%	-1,21%	-2,38%
Maximum	3,94%	1,85%	3,94%

in der ersten Spalte die Ergebnisse dieser Analyse.

Geht man wie hier davon aus, dass der „wahre“ Wert nach 30 Revisionen erreicht ist, so wird eine Erstveröffentlichung im Durchschnitt um 0,51% nach oben revidiert, bevor dieser erreicht ist. Die größten Abweichungen der Erstveröffentlichung vom „wahren“ Wert betragen -2,38% und +3,94%. In Abbildung 3.4 sind die Korrekturen zwischen der Erstveröffentlichung und dem „wahren“ Wert für jeden Berichtsmonat abgetragen.

Es zeigt sich, dass die Revisionen ab dem Jahr 2000 deutlich geringer werden. Dies bestätigt sich, wenn man die Stichprobe unterteilt und die Revisionen vor und ab 2000 getrennt untersucht. Es zeigt sich, dass sowohl der Mittelwert als auch die Standardabweichung der Revisionen im zweiten Teil der Stichprobe geringer ist. Entsprechend haben sich auch die Minimal- und Maximalwerte reduziert. Diese Veränderung kann durch eine Änderung im Erhebungsverfahren im Jahr 1999 erklärt werden (Deutsche Bundesbank 2009). Die auffällig starke Revision der Werte der Berichtsmonate im Jahr 1999 ist auf die fehlende Erfahrung der statistischen Behörden mit dem neuen Erhebungsverfahren zurückzuführen (Knetsch & Reimers 2006). Insgesamt scheint sich aber die Qualität der Erstveröffentlichungen seit der Einführung des neuen Verfahrens verbessert zu haben. Gleichwohl wird noch immer eine statistisch signifikante Unterschätzung des „wahren“ Wertes um 0,31% festgestellt.

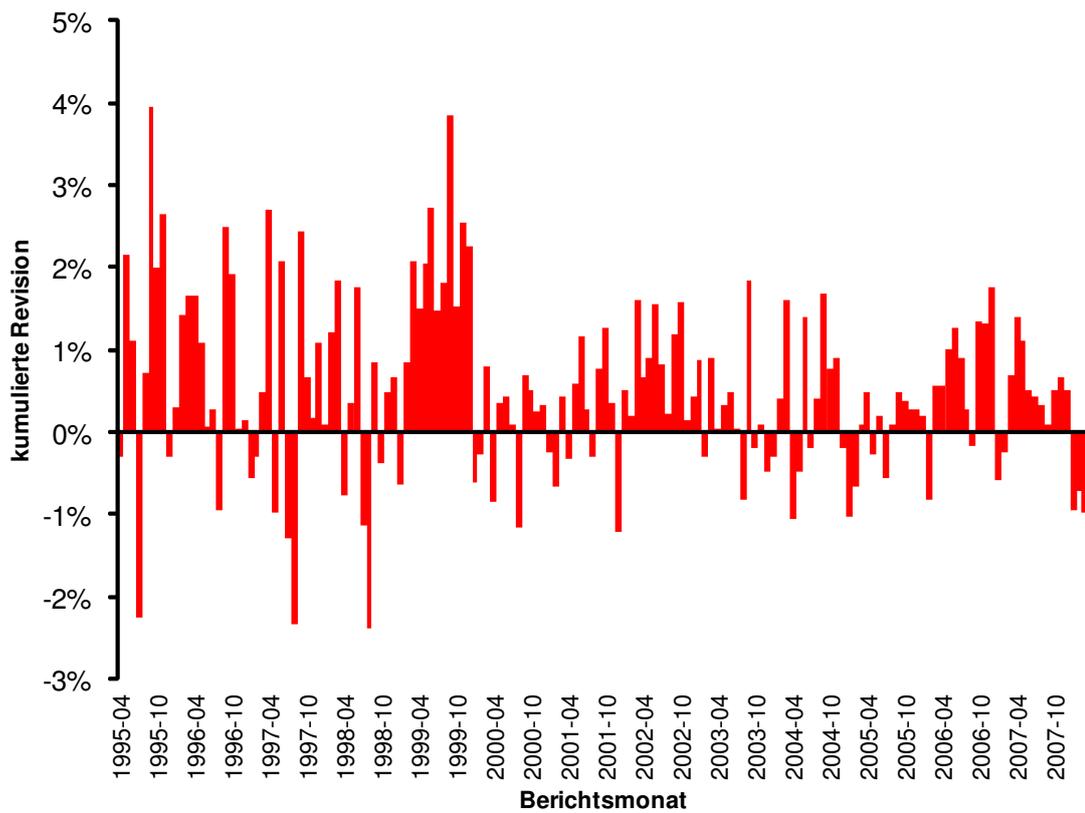


Abbildung 3.4: Produktion – kumulierte Revisionen bis zur 30sten Revision

3.1.2 Korrelationsanalyse

Im Folgenden untersuchen wir unter verschiedenen Aspekten die statistischen Zusammenhänge zwischen den Revisionshistorien der Variablen. Dafür werden im Rahmen dieser deskriptiven Analyse einfache Korrelationskoeffizienten bestimmt.

Es wäre z.B. möglich, dass zwischen der Höhe der veröffentlichten Wachstumsraten der Industrieproduktion und der Höhe der veröffentlichten Revisionen ein Zusammenhang bestehen. Man könnte vermuten, dass eine anfangs überdurchschnittlich hohe gemeldete Wachstumsrate nachfolgend unterdurchschnittlich nach oben, bzw. sogar eher nach unten revidiert wird. Dieser Zusammenhang bestätigt sich tatsächlich, es lässt sich eine Korrelation von $-0,25$ zwischen der erstveröffentlichten Wachstumsrate und ihrer anschließenden Revision feststellen. Auch hätte man vermuten können, dass eine starke (erste) Revision des Werts des vorherigen Monats einen Einfluss auf die zeitgleich erfolgte Erstveröffentlichung hat. Diese These hat sich in der Korrelationsanalyse allerdings nicht bestätigt.

Des Weiteren wurde untersucht, ob ein Zusammenhang zwischen den verschiedenen Revisionen eines Monats festzustellen ist. Es wurden die Korrelationen zwischen den ersten sechs Revisionen untersucht. Die Ergebnisse hierzu sind in Tabelle 3.3 zusammengefasst. Insgesamt zeigt sich, dass diese keine relevanten Korrelationen aufweisen.

Außerdem stellt sich die Frage, ob aufeinander folgende Revisionen (bezüglich der Werte verschiedener, ebenfalls aufeinander folgender Monatsmonate) miteinander zusammenhängen. Dabei beschränken wir uns hier auf die jeweils ersten erfolgten Revisionen. Die Ergebnisse sind in Abbildung 3.5 dargestellt. Interessant sind hier die direkt zeitlich benachbarten Revisionen: es ist eine negative Korrelation zu beobachten, d.h. sollte zu einem Monatsmonat die erste Revision überdurchschnittlich hoch sein, so fallen die ersten Revisionen zu den beiden nachfolgenden Monatsmonaten tendenziell unterdurchschnittlich aus. Der Hintergrund dieses Phänomens ist nicht geklärt, es könnte sich um institutionelle Gegebenheiten der statistischen Ämter handeln. Weiterhin ist auch ein gewisses Saisonmuster zu beobachten (z.B. bei der zwölften Verzögerung) unabhängig davon, dass die zu Grunde liegenden Reihen selbst saisonbereinigt sind.

Betrachtet man die Zusammenhänge zwischen den Revisionen des Produktionsindikators und den Revisionen der anderen Indikatoren (Tabellen 3.4 bis 3.6), dann zeigt sich, dass die erste Revision der Auftragseingänge aus dem Inland relativ stark mit der Revision der Produktion korreliert ist. Ebenfalls eine leichte Korrelation findet sich für die ersten

3 Deskriptive Analyse der Datenrevisionen

Tabelle 3.3: Produktion – Korrelationskoeffizienten verschiedener Revisionen zu einem Berichtsmonat

IPROD_OB Rev _{t+3}	IPROD_OB Rev _{t+4}	IPROD_OB Rev _{t+5}	IPROD_OB Rev _{t+6}	IPROD_OB Rev _{t+7}	IPROD_OB Rev _{t+8}	
1,0000	-0,0477	-0,0957	-0,0897	-0,0527	-0,1843	IPROD_OB Rev _{t+3}
	1,0000	-0,0007	-0,0408	0,0877	0,0701	IPROD_OB Rev _{t+4}
		1,0000	0,0268	-0,0034	0,1950	IPROD_OB Rev _{t+5}
			1,0000	-0,0045	-0,0325	IPROD_OB Rev _{t+6}
				1,0000	-0,0292	IPROD_OB Rev _{t+7}
					1,0000	IPROD_OB Rev _{t+8}

Anmerkungen: Beobachtungen 1993m2–2010m10 (Fehlwerte wurden ausgelassen), 5% kritischer Wert (zweiseitig) = 0,1345 für n = 213.

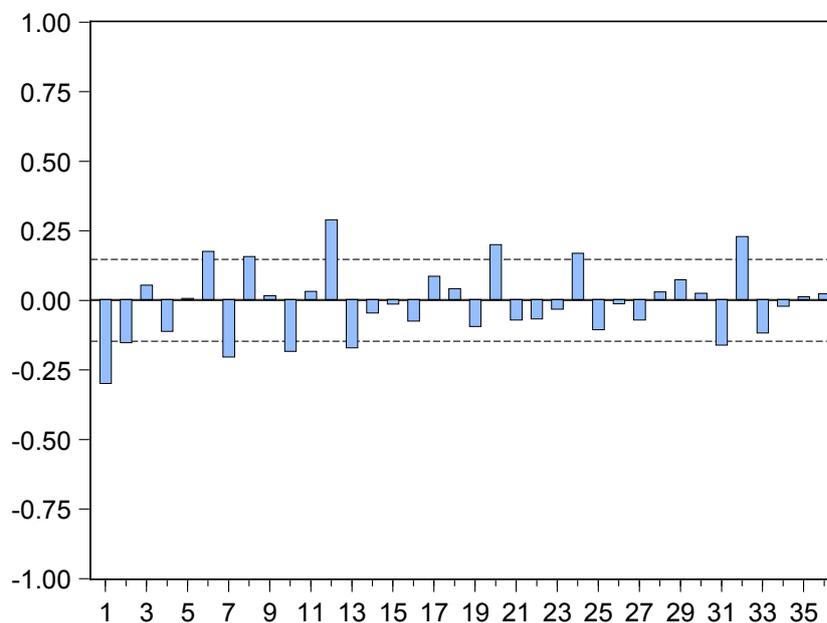


Abbildung 3.5: Produktion – Autokorrelation der ersten Revisionen, mit Konfidenzbändern

Tabelle 3.4: Korrelation zwischen den Revisionen des Produktionsindikators und den Auftragseingängen aus dem Inland

IPROD_ OB Rev _{t+3}	IPROD_ OB Rev _{t+4}	IPROD_ OB Rev _{t+5}	IPROD_ OB Rev _{t+6}	IPROD_ OB Rev _{t+7}	IPROD_ OB Rev _{t+8}	
0,4984	-0,1033	-0,0743	-0,0656	-0,0558	-0,22	AE_INL_BUBA Rev _{t+3}
0,1398	-0,0308	0,0109	0,0855	-0,0205	-0,0418	AE_INL_BUBA Rev _{t+4}
-0,093	0,0397	0,1579	-0,0543	0,045	0,0579	AE_INL_BUBA Rev _{t+5}
-0,1255	0,2621	-0,0369	0,0567	0,0621	0,1123	AE_INL_BUBA Rev _{t+6}
-0,0622	0	-0,0812	-0,0949	0,0998	-0,0664	AE_INL_BUBA Rev _{t+7}
-0,0057	0,1988	0,1259	-0,019	0,0948	0,1635	AE_INL_BUBA Rev _{t+8}

Revisionen der Auftragseingänge aus dem Ausland und der Produktion. Dies lässt sich dadurch erklären, dass wie bereits erwähnt bei der ersten Veröffentlichung die Stichprobe noch nicht vollständig vorliegt. Sollten nun die noch eingehenden Meldungen überdurchschnittliche (unterdurchschnittliche) Produktionswerte melden, liegt es nahe, dass die gleichen Firmen auch überdurchschnittliche (unterdurchschnittliche) Auftragseingänge in diesem Zeitraum zu vermelden haben. Gehen die Werte solcher Firmen verspätet ein, könnte das die Korrelation in den Revisionen erklären.

3.2 Auftragseingänge aus dem Inland

3.2.1 Verteilungen

Wie in Abbildung 3.6 zu sehen ist, werden die Auftragseingänge aus dem Inland, wie auch die Produktionsdaten, mit der ersten Revision meist revidiert (82%). Zur nächsten Veröffentlichung werden dann nochmals 34% der Werte revidiert. Der Anteil sinkt dann im Zeitverlauf weiter ab, um dann wie bei der Produktion zur 12ten und zur 24sten Revision nochmals anzusteigen. Insgesamt sind die durchschnittlichen ersten Revisionen mit 0,43% kleiner als bei der Produktion. Die zweiten Revisionen sind dann mit knapp 0,1% relativ gering und fallen von diesem Niveau weiter ab.

3 Deskriptive Analyse der Datenrevisionen

Tabelle 3.5: Korrelation zwischen den Revisionen des Produktionsindikators und den Auftragseingängen aus dem Ausland

IPROD_OB Rev _{t+3}	IPROD_OB Rev _{t+4}	IPROD_OB Rev _{t+5}	IPROD_OB Rev _{t+6}	IPROD_OB Rev _{t+7}	IPROD_OB Rev _{t+8}	
0,2151	-0,0655	0,0117	0,0835	0,1264	0,119	AE_AUSL_BUBA Rev _{t+3}
-0,1206	0,0004	0,0358	0,1341	0,0437	0,1282	AE_AUSL_BUBA Rev _{t+4}
-0,0289	0,0054	-0,0538	-0,0057	-0,0114	0,1669	AE_AUSL_BUBA Rev _{t+5}
0,0893	0,1168	0,1067	0,158	-0,0284	0,03	AE_AUSL_BUBA Rev _{t+6}
-0,131	0,0051	-0,0688	0,0659	0,121	0,0891	AE_AUSL_BUBA Rev _{t+7}
0,0031	0,0396	0,0619	0,0666	0,0248	0,1019	AE_AUSL_BUBA Rev _{t+8}

Tabelle 3.6: Korrelation zwischen den Revisionen des Produktionsindikators und den Konsumentenpreisen

IPROD_OB Rev _{t+3}	IPROD_OB Rev _{t+4}	IPROD_OB Rev _{t+5}	IPROD_OB Rev _{t+6}	IPROD_OB Rev _{t+7}	IPROD_OB Rev _{t+8}	
-0,0223	0,0325	0,0406	0,1651	0,0577	0,0617	CPI Rev _{t+2}
0,0805	-0,0126	-0,0749	-0,0575	0,0579	0,0846	CPI Rev _{t+3}
0,0599	0,0343	-0,1181	-0,0821	0,0313	0,0428	CPI Rev _{t+4}
-0,0061	0,2069	-0,0727	0,0104	-0,1667	-0,0873	CPI Rev _{t+5}
-0,1266	0,0262	0,0669	0,0692	-0,2974	0,1109	CPI Rev _{t+6}
-0,0103	0,0113	-0,0004	0,0461	0,006	-0,1626	CPI Rev _{t+7}
-0,0378	-0,1557	0,1107	0,0401	0,0119	0,0518	CPI Rev _{t+8}

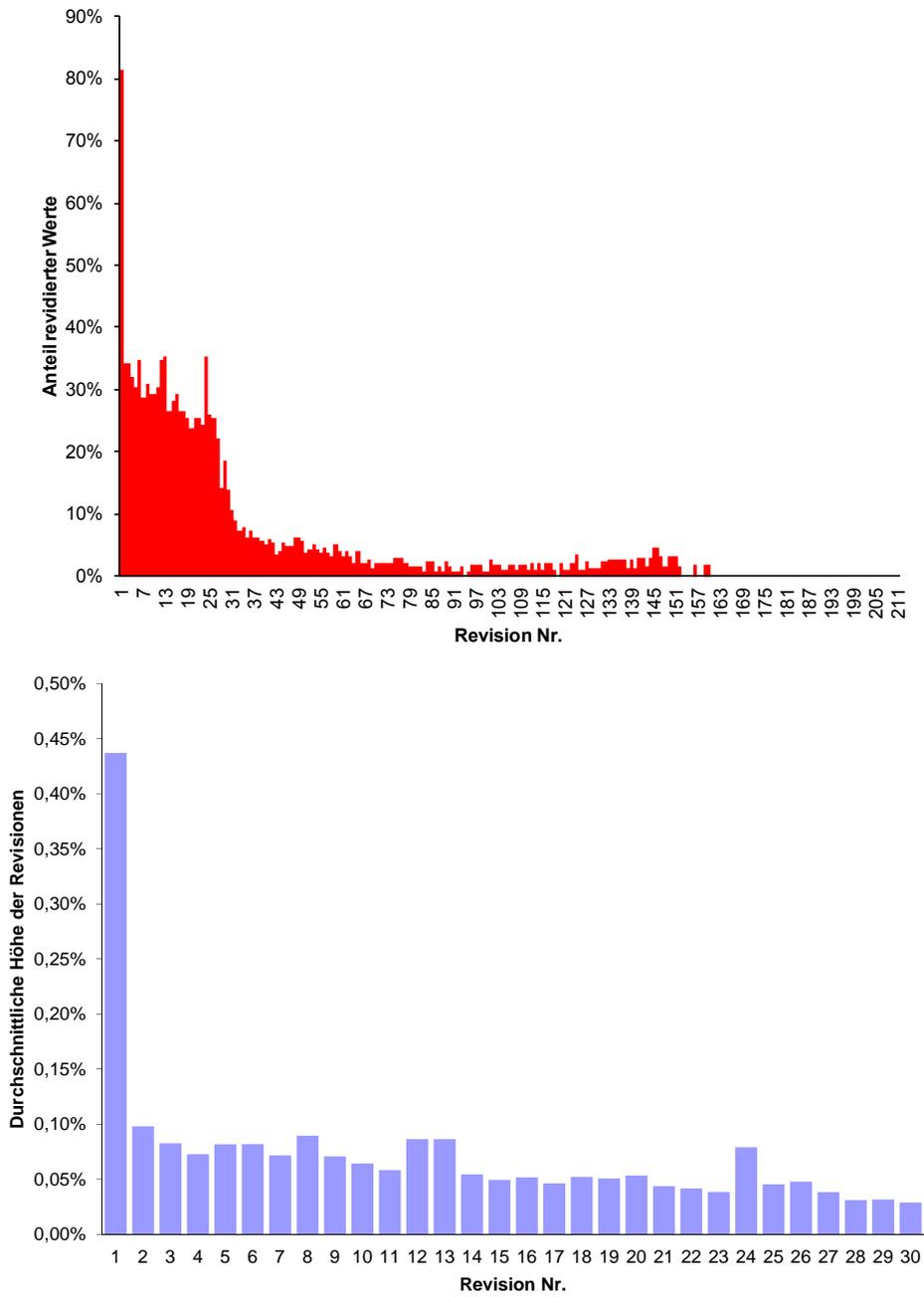


Abbildung 3.6: Auftragseingänge aus dem Inland – prozentualer Anteil der tatsächlichen Revisionen und durchschnittliche absolute Höhe der Revisionen nach Revisionsnummer

3 Deskriptive Analyse der Datenrevisionen

Tabelle 3.7: Auftragseingänge aus dem Inland – deskriptive Statistik zu den Revisionsdaten

	kompletter Datensatz		erste 30 Revisionen	
	mit Nullwerten	ohne Nullwerte	mit Nullwerten	ohne Nullwerte
Mittelwert	0,0016%	0,0164%	0,0067%	0,0226%
Standardabweichung	0,1001%	0,3223%	0,1877%	0,3451%
Stichprobengröße	21824	2100	5514	1627
Standardfehler	6,775E-06	7,034E-05	2,528E-05	8,556E-05
t-Wert	2,331	2,333	2,637	2,640
p-Wert	0,0198	0,0197	0,0084	0,0084

In Tabelle 3.7 sind die deskriptiven statistischen Kennzahlen für den kompletten Datensatz sowie für die ersten 30 Revisionen dargestellt. Auch hier zeigt sich, dass ca. drei Viertel der mittleren Gesamtrevisionen innerhalb dieser ersten 2,5 Jahre stattfinden. Ebenfalls wurde ein Mittelwert der Revisionen ermittelt, der sich signifikant von 0 unterscheidet. Demnach scheinen auch die Daten der erstveröffentlichten Auftragseingänge aus dem Inland generell den tatsächlichen Stand zu unterschätzen.

Geht man wieder davon aus, dass der „wahre“ Wert des Index nach 30 Revisionen erreicht ist, so kann man diesen wieder mit der Erstveröffentlichung vergleichen. Die Ergebnisse zur Analyse finden sich in Tabelle 3.8, eine graphische Darstellung in Abbildung 3.7. Wie bereits bei den Produktionswerten festgestellt, liegt auch die Erstveröffentlichung der monatlichen Auftragseingänge aus dem Inland systematisch unter ihrem „wahren“ Wert. Diese Unterschätzung ist mit 0,18% aber geringer als bei der Produktion. Es wurde ebenfalls untersucht, ob sich eine ähnliche Verbesserung der Schätzung bei den neueren Veröffentlichungen ergeben hat. Dies scheint bei einem ersten Blick auf die getrennt ermittelten Mittelwerte der Fall, kann allerdings auf keinem statistisch üblichen Signifikanzniveau bestätigt werden.

Tabelle 3.8: Auftragseingänge aus dem Inland – Durchschnitt der kumulierten Revisionen der ersten 30 Rechenstände

	gesamter Zeitraum	seit 2000	vor 2000
Mittelwert	0,18%	0,15%	0,22%
Standardabw.	0,0101	0,0090	0,0119
Beobachtungen	156	99	57
Standardfehler	0,00081	0,00091	0,0016
t-wert	2,190	1,674	1,413
p-wert	0,0300	0,0974	0,1631
Minimum	-2,47%	-2,44%	-2,47%
Maximum	2,97%	2,27%	2,97%

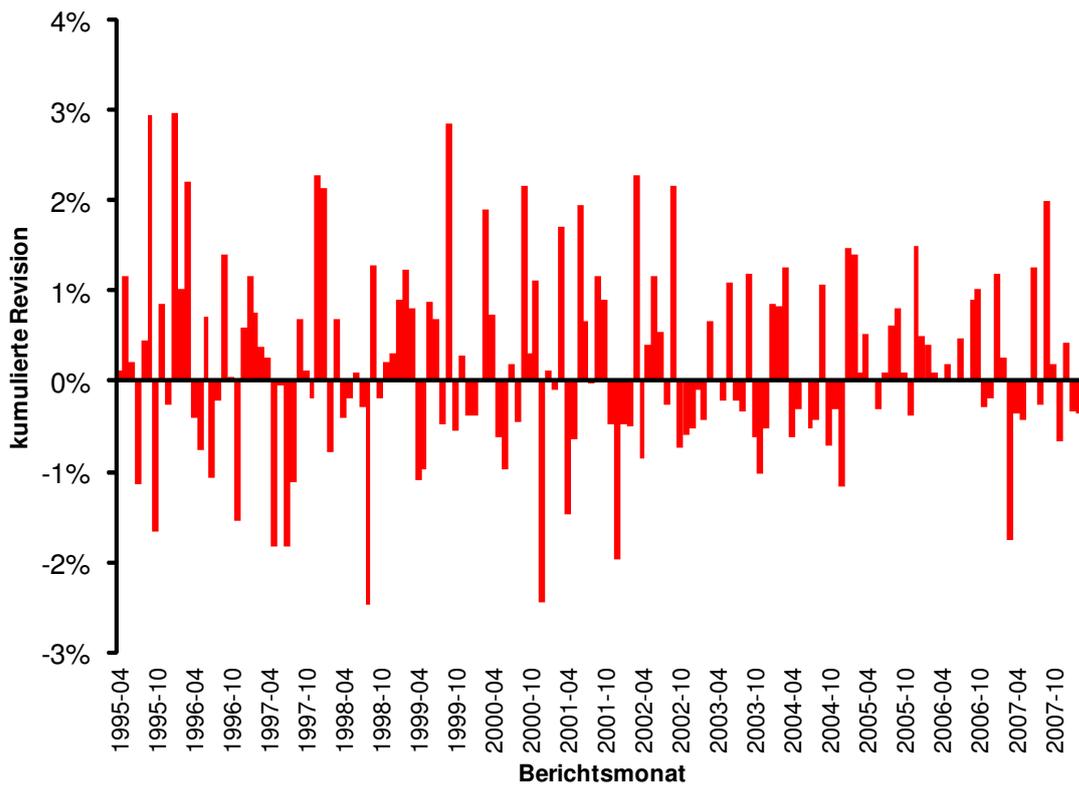


Abbildung 3.7: Auftragseingänge aus dem Inland – kumulierte Revisionen bis zur 30sten Revision

Tabelle 3.9: Auftragseingänge aus dem Inland – Korrelationskoeffizienten verschiedener Revisionen zu einem Berichtsmonat

AE_INL_BUBA Rev _{t+3}	AE_INL_BUBA Rev _{t+4}	AE_INL_BUBA Rev _{t+5}	AE_INL_BUBA Rev _{t+6}	AE_INL_BUBA Rev _{t+7}	AE_INL_BUBA Rev _{t+8}	
1,0000	0,0021	-0,0600	-0,1580	0,0238	-0,1222	AE_INL_BUBA Rev _{t+3}
	1,0000	0,0455	-0,0920	0,0422	0,0017	AE_INL_BUBA Rev _{t+4}
		1,0000	-0,0224	-0,0316	0,0527	AE_INL_BUBA Rev _{t+5}
			1,0000	-0,0460	0,1258	AE_INL_BUBA Rev _{t+6}
				1,0000	0,0207	AE_INL_BUBA Rev _{t+7}
					1,0000	AE_INL_BUBA Rev _{t+8}

Anmerkungen: Beobachtungen 1993m2–2010m9 (Fehlwerte wurden ausgelassen), 5% kritischer Wert (zweiseitig) = 0,1348 für n = 212.

3.2.2 Korrelationsanalyse

In Tabelle 3.9 sind die Korrelationen der verschiedenen Revisionen zu einem Berichtsmonat dargestellt. Es zeigt sich, dass keinerlei relevante Korrelationen bestehen.

Betrachtet man die Autokorrelationsfunktion zu den ersten Revisionen in Abbildung 3.8, so scheint es hier ein ähnliches Muster wie bei den Produktionsrevisionen zu geben. War die erste Revision des vorherigen Berichtsmonats überdurchschnittlich hoch, so wird die erste Revision des folgenden Berichtsmonats tendenziell unterdurchschnittlich hoch sein. Des Weiteren werden positive Korrelationen für die 12te, 20ste, 24ste, und 28ste Verzögerung festgestellt.

3.3 Auftragseingänge aus dem Ausland

3.3.1 Verteilungen

Wie in Abbildung 3.9 zu sehen, wird auch bei den Auftragseingängen aus dem Ausland ein Großteil der Veröffentlichungen nach einem Monat revidiert (87%). Die Revisionswahr-

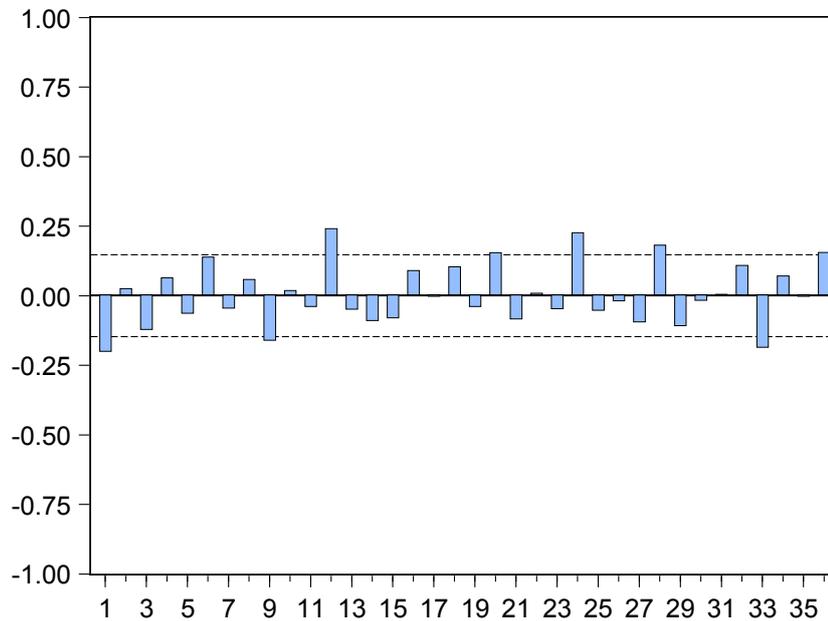


Abbildung 3.8: Auftragseingänge aus dem Inland – Autokorrelation der ersten Revisionen, mit Konfidenzbändern

scheinlichkeit fällt dann relativ schnell und hält sich bis zur 28sten Revision zwischen 30% und 40%. Wobei auch hier ein Anstieg zur 12ten und 24sten Revision zu beobachten ist. Die absolute durchschnittliche Revisionshöhe liegt bei der ersten Revision bei 0,58%, fällt dann zur zweiten Revision auf 0,14% und sinkt dann im Zeitverlauf weiter ab. Auch in der Revisionshöhe zeigt sich wieder der erhöhte Revisionsbedarf zur 12ten und 24sten Revision.

In Tabelle 3.10 sind die deskriptiven statistischen Werte des kompletten Datensatzes und der ersten 30 Revisionen dargestellt. Insgesamt finden wieder ca. drei Viertel der gesamten Revisionen innerhalb der ersten 30 Monate nach der Erstveröffentlichung statt. Hier liegen die Mittelwerte wieder signifikant über Null, so dass auch hier die Erstveröffentlichungen den tatsächlichen Wert im Durchschnitt zunächst unterschätzen.

In Tabelle 3.11 und Abbildung 3.10 sind die kumulierten Revisionen bis zum „wahren“ Wert nach 30 Revisionen dargestellt. Wie bereits erwähnt unterschätzen die Erstveröffentlichungen diesen signifikant und zwar um durchschnittlich 0,35%. Auch hier sind wieder die Werte für den Zeitraum bis 2000 und nach 2000 angegeben. Im Gegensatz zu den Pro-

3 Deskriptive Analyse der Datenrevisionen

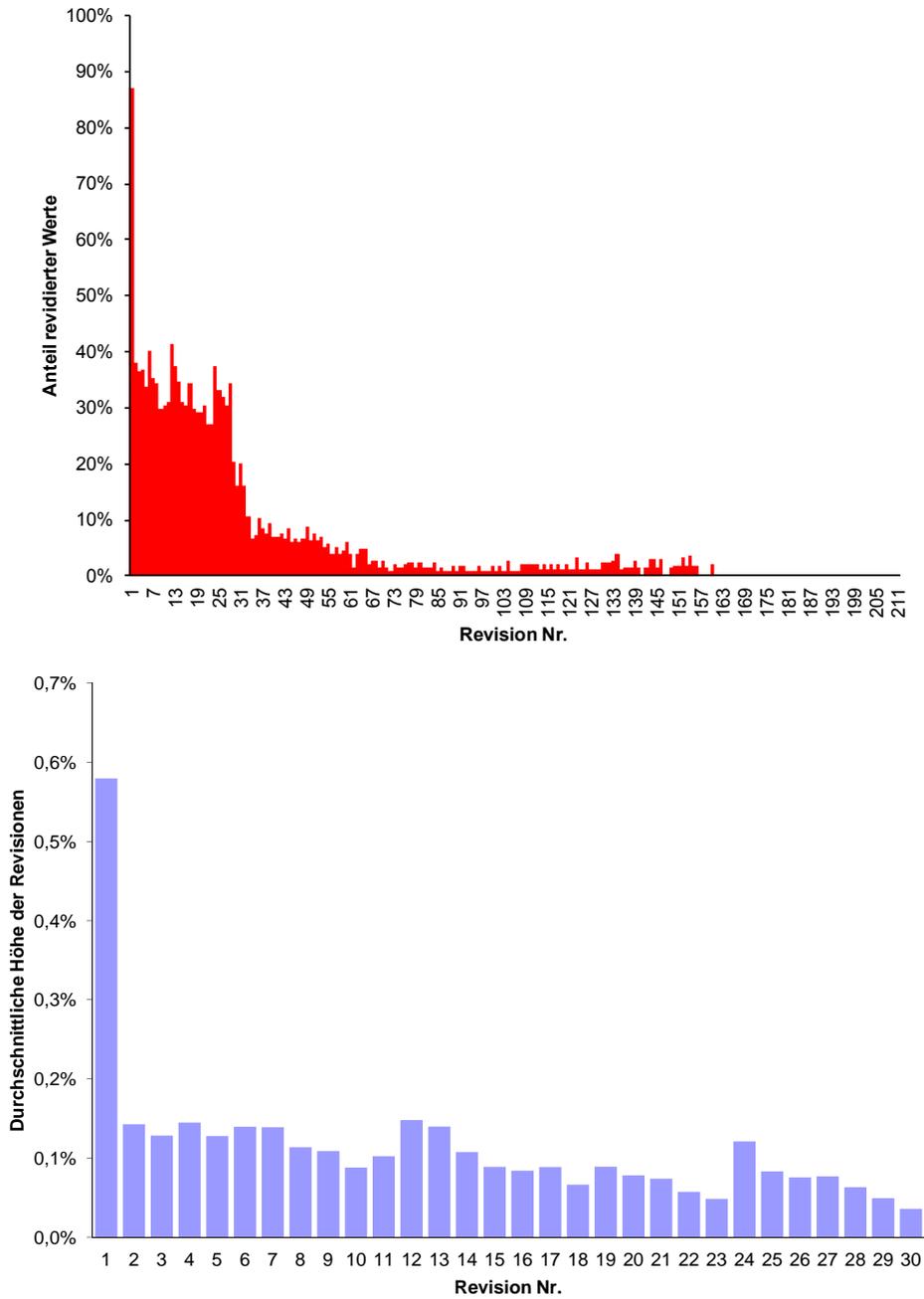


Abbildung 3.9: Auftragseingänge aus dem Ausland – prozentualer Anteil der tatsächlichen Revisionen und durchschnittliche absolute Höhe der Revisionen nach Revisionsnummer

Tabelle 3.10: Auftragseingänge aus dem Ausland – deskriptive Statistik zu den Revisionsdaten

	kompletter Datensatz		erste 30 Revisionen	
	mit Nullwerten	ohne Nullwerte	mit Nullwerten	ohne Nullwerte
Mittelwert	0,0024%	0,0215%	0,0106%	0,0312%
Standardabweichung	0,1650%	0,4932%	0,3086%	0,5287%
Stichprobengröße	21824	2441	5514	1875
Standardfehler	1,117E-05	9,982E-05	4,156E-05	0,00012
t-Wert	2,150	2,152	2,551	2,553
p-Wert	0,0315	0,0315	0,0108	0,0108

duktionsdaten haben die mittleren Revisionen bei den Auftragseingängen aus dem Ausland ab dem Jahr 2000 zugenommen.

3.3.2 Korrelationsanalyse

In Tabelle 3.12 sind die Korrelationen der verschiedenen Revisionen zu einem Berichtsmonat dargestellt. Auch hier sind die Korrelationen nur schwach ausgeprägt. Betrachtet man die Autokorrelationsfunktion der ersten Revisionen in Abbildung 3.11, konnten signifikante Korrelationen nur für die 5te und 10te Verzögerung festgestellt werden.

3.4 Konsumentenpreise

3.4.1 Verteilungen

Für die Konsumentenpreise zeigt sich in Abbildung 3.12, dass sie deutlich seltener revidiert werden und dass auch die Höhe der Revisionen deutlich geringer ist als bei den bisher untersuchten Indikatoren. Anders als bei den bisher untersuchten Kenngrößen werden bei der ersten Revision nur 11% der Werte revidiert. Es zeigt sich dann ebenfalls ein erneutes Ansteigen der Revisionen und deren Höhe zur 12ten Revision.

Anders als bei der Industrieproduktion und den Auftragseingängen gibt es Hinweise

3 Deskriptive Analyse der Datenrevisionen

Tabelle 3.11: Auftragseingänge aus dem Ausland – Durchschnitt der kumulierten Revisionen der ersten 30 Rechenstände

		seit 2000	vor 2000
Mittelwert	0,35%	0,41%	0,25%
Standardabw.	0,0161	0,0146	0,0185
Beobachtungen	156	99	57
Standardfehler	0,0013	0,0015	0,0025
t-wert	2,712	2,787	1,006
p-wert	0,0074	0,0064	0,3186
Minimum	-5,56%	-5,56%	-3,68%
Maximum	4,43%	3,90%	4,43%

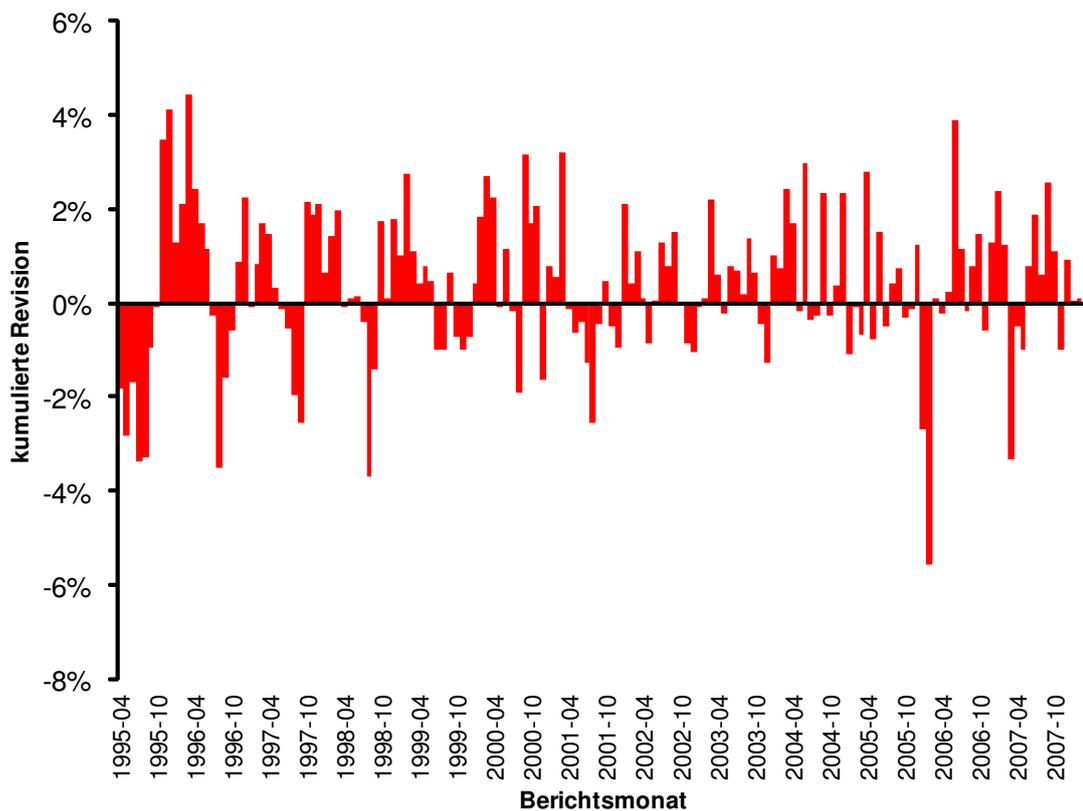


Abbildung 3.10: Auftragseingänge aus dem Ausland – kumulierte Revisionen bis zur 30sten Revision

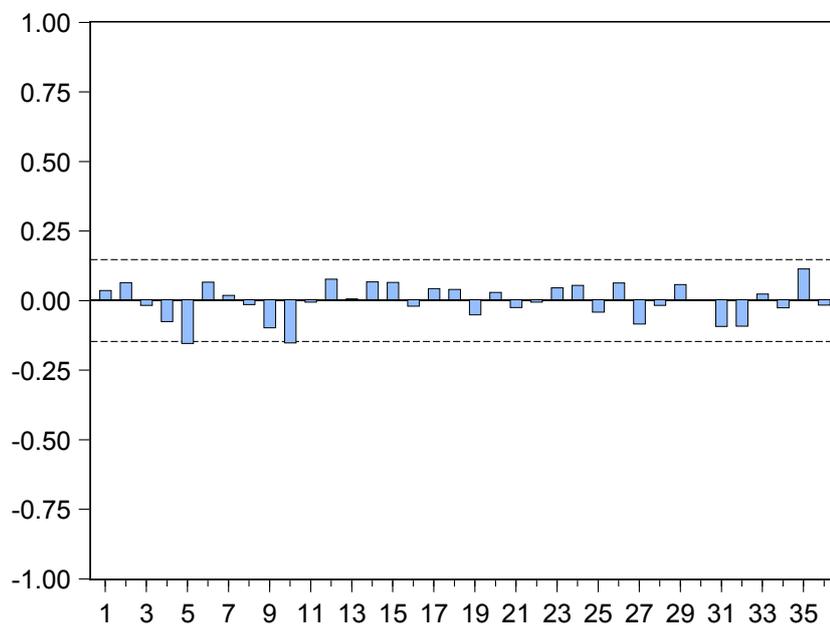


Abbildung 3.11: Auftragseingänge aus dem Ausland – Autokorrelation der ersten Revisionen, mit Konfidenzbändern

3 Deskriptive Analyse der Datenrevisionen

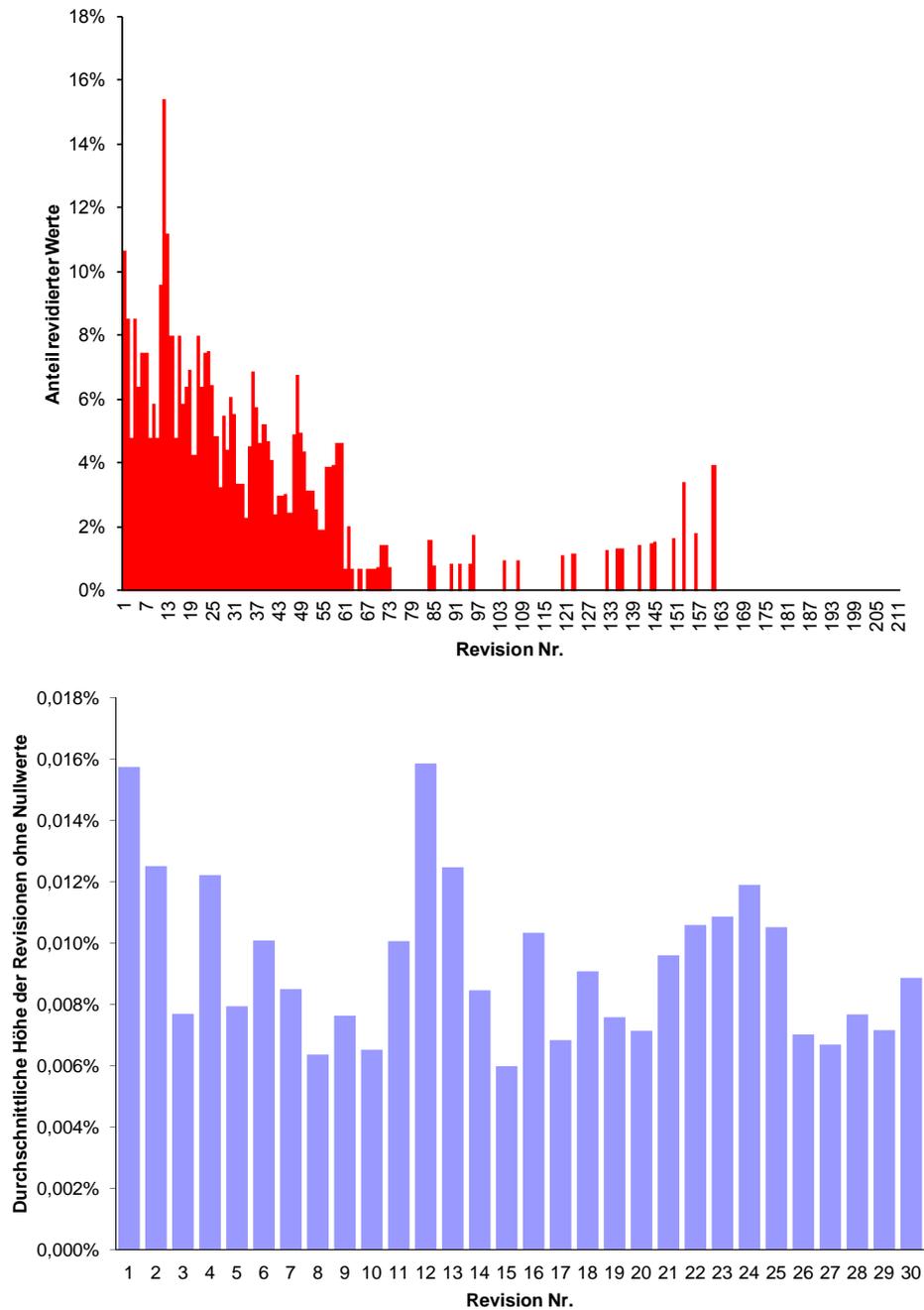


Abbildung 3.12: Konsumentenpreise – prozentualer Anteil der tatsächlichen Revisionen und durchschnittliche absolute Höhe der Revisionen nach Revisionsnummer

Tabelle 3.12: Auftragseingänge aus dem Ausland – Korrelationskoeffizienten verschiedener Revisionen zu einem Berichtsmonat

AE_AUSL_BUBA Rev _{t+3}	AE_AUSL_BUBA Rev _{t+4}	AE_AUSL_BUBA Rev _{t+5}	AE_AUSL_BUBA Rev _{t+6}	AE_AUSL_BUBA Rev _{t+7}	AE_AUSL_BUBA Rev _{t+8}	
1,0000	0,0255	0,1598	0,1252	0,0491	0,0324	AE_AUSL_BUBA Rev _{t+3}
	1,0000	-0,0141	0,0198	0,1993	0,0130	AE_AUSL_BUBA Rev _{t+4}
		1,0000	-0,0468	0,0368	0,0279	AE_AUSL_BUBA Rev _{t+5}
			1,0000	-0,0265	0,0070	AE_AUSL_BUBA Rev _{t+6}
				1,0000	0,0000	AE_AUSL_BUBA Rev _{t+7}
					1,0000	AE_AUSL_BUBA Rev _{t+8}

Anmerkungen: Beobachtungen 1993m2–2010m9 (Fehlwerte wurden ausgelassen), 5% kritischer Wert (zweiseitig) = 0,1348 für n = 212.

darauf, dass die Preissteigerungen bei der Erstveröffentlichung eher über- als unterschätzt werden, siehe Tabelle 3.13. Geht man auch hier davon aus, dass der „wahre“ Wert nach 30 Revisionen erreicht ist, so wird die Preisentwicklung durchschnittlich um 0,11% überschätzt (Tabelle 3.14). Bemerkenswert ist, dass sich starke Korrekturen auf den Zeitraum Mitte 2000 bis Ende 2002 konzentrieren (Abbildung 3.13). Dies könnte mit der Einführung des Euro und damit einhergehenden nominellen Preisänderungen zusammenhängen.

3.4.2 Korrelationsanalyse

In Tabelle 3.15 sind die Korrelationen der verschiedenen Revisionen zu einem Berichtsmonat dargestellt. Es zeigt sich, dass keinerlei relevante Korrelationen bestehen. Betrachtet man die Autokorrelationsfunktion zu den ersten Revisionen in Abbildung 3.14, sind bis auf die zweite Verzögerung keine signifikanten Werte zu finden.

3 Deskriptive Analyse der Datenrevisionen

Tabelle 3.13: Konsumentenpreise – deskriptive Statistik zu den Revisionsdaten

	kompletter Datensatz		erste 30 Revisionen	
	mit Nullwerten	ohne Nullwerte	mit Nullwerten	ohne Nullwerte
Mittelwert	-0,00015%	-0,0054%	-0,0026%	-0,0374%
Standardabweichung	0,0331%	0,1958%	0,0481%	0,1784%
Stichprobengröße	21993	628	5612	392
Standardfehler	2,230E-06	7,812E-05	6,415E-06	9,012E-05
t-Wert	-0,693	-0,693	-4,069	-4,147
p-Wert	0,4881	0,4886	4,783E-05	4,141E-05

Tabelle 3.14: Konsumentenpreise – Durchschnitt der kumulierten Revisionen der ersten 30 Rechenstände

	gesamter Zeitraum	seit 2000	vor 2000
Mittelwert	-0,11%	-0,14%	-0,06%
Standardabw.	0,0022	0,0024	0,0017
Stichprobe	161	100	61
Standardfehler	0,00017	0,00024	0,00021
t-wert	-6,344	-5,767	-2,839
p-wert	2,201E-09	9,193E-08	0,0062
Minimum	-0,0068	-0,0068	-0,0067
Maximum	0,0019	0,0019	0,0017

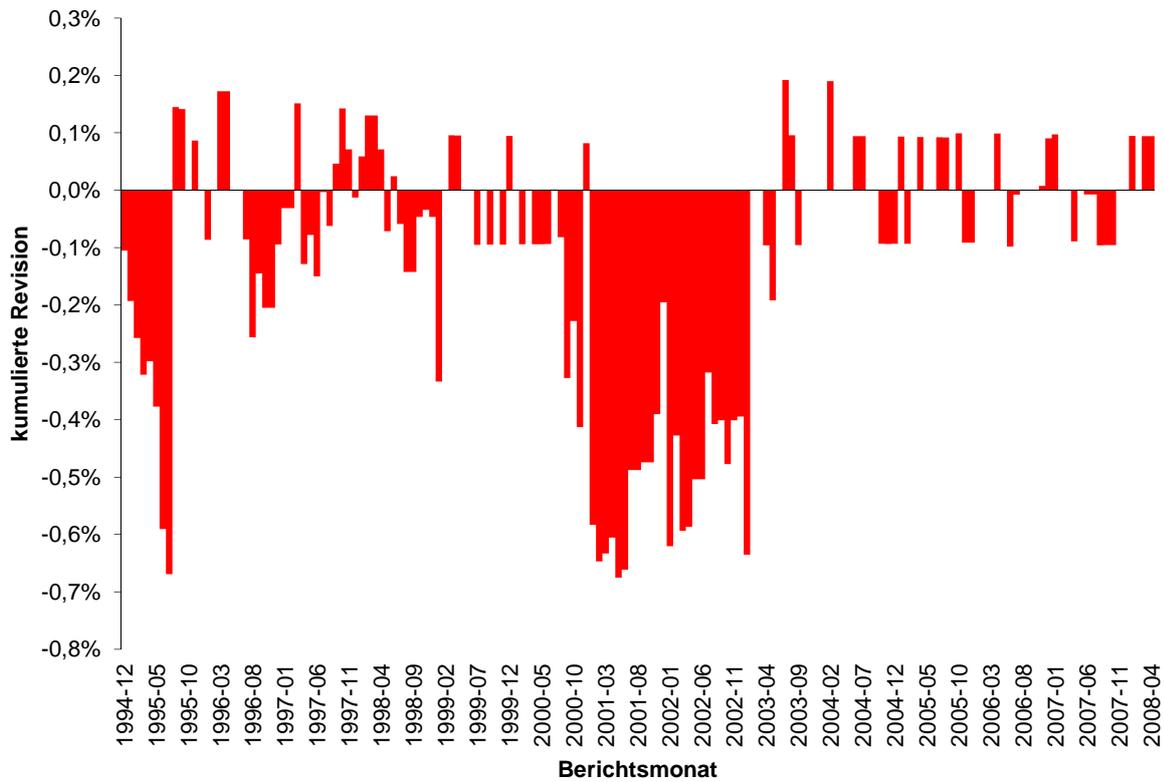


Abbildung 3.13: Konsumentenpreise – kumulierte Revisionen bis zur 30sten Revision

Tabelle 3.15: Konsumentenpreise – Korrelationskoeffizienten verschiedener Revisionen zu einem Berichtsmonat

CPI Rev _{t+2}	CPI Rev _{t+3}	CPI Rev _{t+4}	CPI Rev _{t+5}	CPI Rev _{t+6}	CPI Rev _{t+7}	
1,0000	0,0247	-0,0701	-0,0252	-0,0240	-0,0503	CPI Rev _{t+2}
	1,0000	-0,0069	-0,0014	-0,0067	0,0106	CPI Rev _{t+3}
		1,0000	-0,0127	0,0107	-0,0304	CPI Rev _{t+4}
			1,0000	-0,0509	-0,0752	CPI Rev _{t+5}
				1,0000	-0,0538	CPI Rev _{t+6}
					1,0000	CPI Rev _{t+7}

Anmerkung: Beobachtungen 1993m2–2010m9 (Fehlwerte wurden ausgelassen), 5% kritischer Wert (zweiseitig) = 0,1459 für n = 181.

3 Deskriptive Analyse der Datenrevisionen

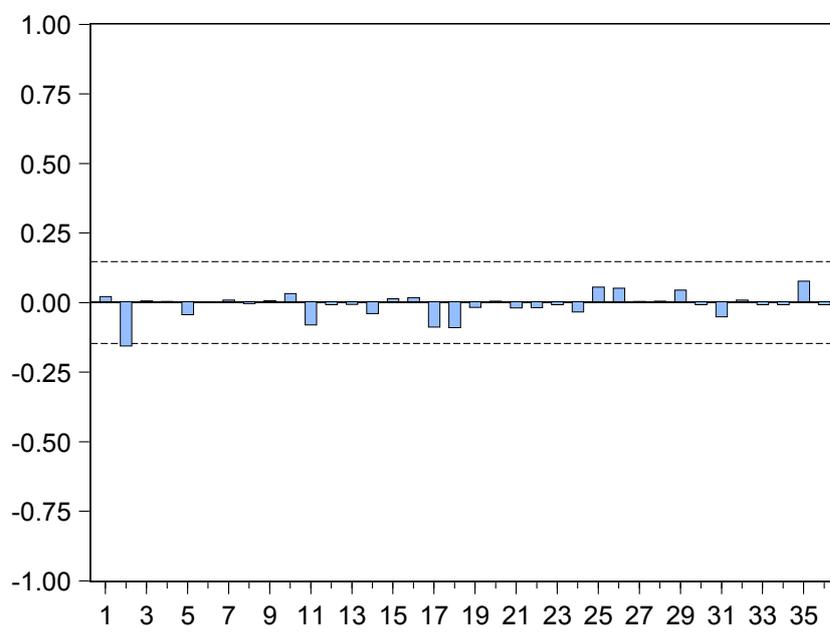


Abbildung 3.14: Konsumentenpreise – Autokorrelation der ersten Revisionen, mit Konfidenzbändern

3.5 Korrelationen zwischen den Revisionen der betrachteten Indikatoren

Die Korrelation zwischen den Revisionen der verschiedenen Indikatoren ist in Tabelle 3.16 dargestellt. Die Korrelationen mit den Revisionen des Produktionsindex wurden bereits früher im Text besprochen. Insgesamt sind zwischen den Revisionen der anderen Indikatoren keine besonders hohen Korrelationen festzustellen. Auffällig ist, dass die Revisionen der Auftragseingänge aus dem Ausland eine vergleichsweise hohe positive Korrelation zwischen 0,16 und 0,27 mit den Revisionen der Preise aufweisen. Ebenfalls positiv korreliert sind die ersten Korrekturen der In- und Auslandsaufträge, wobei die zweiten Revisionen dann auf einen negativen Zusammenhang hinzuweisen scheinen. Insgesamt lässt sich jedoch feststellen, dass die gefundenen Korrelationen vergleichsweise gering sind.

3.6 Zusammenfassung

Im Rahmen dieser Studie liegen die Revisionshistorien der folgenden mit monatlicher Frequenz gemessenen Variablen bzgl. der deutschen Volkswirtschaft vor (siehe auch Kapitel 2): Industrieproduktion, Auftragseingänge aus dem Inland bzw. aus dem Ausland und die Konsumentenpreisindizes. Die systematische statistisch-deskriptive Analyse dieser empirisch beobachteten Datenrevisionen hat bereits einige interessante Ergebnisse erbracht.

So sind bei den ersten dreien der genannten Variablen bei der jeweils zweiten Veröffentlichung (d.h. bei der ersten möglichen Revision) die Wahrscheinlichkeiten relativ hoch, dass tatsächlich Revisionen erfolgen. Auch die Revisionsausmaße sind generell durchaus relevant, weitere Datenveröffentlichungen weisen dann allerdings typischerweise deutlich weniger und geringere Revisionen auf. Dagegen zeigen sich die veröffentlichten Konsumentenpreisindizes insgesamt weit weniger revisionsanfällig.

Durch die verwendeten amtlichen Verfahren der Saisonbereinigung ist es möglich, dass auch lange Zeit nach dem Monatsmonat noch Revisionen erfolgen. Dies zeigt sich in den Daten z.B. auch darin, dass die Revisionshäufigkeit bei den 12ten und teils auch 24sten Revisionen etwas erhöht ist.

Auffällig ist bei allen Variablen außer den Preisen, dass die Erstveröffentlichungen im Laufe der Zeit durchschnittlich signifikant nach oben revidiert werden. Anders ausgedrückt liefern die Erstveröffentlichungen nur verzerrte Schätzungen für die später abgeschlosse-

3 Deskriptive Analyse der Datenrevisionen

Tabelle 3.16: Korrelationsmatrizen der Revisionen eines Monats

IPROD_OB Rev _{t+3}	AE_INL_ BUBA Rev _{t+3}	AE_AUSL_ BUBA Rev _{t+3}	CPI Rev _{t+2}	
1,0000	0,4984	0,2151	-0,0223	IPROD_OB Rev _{t+3}
	1,0000	0,3290	0,0619	AE_INL_BUBA Rev _{t+3}
		1,0000	0,1634	AE_AUSL_BUBA Rev _{t+3}
			1,0000	CPI Rev _{t+2}

IPROD_OB Rev _{t+4}	AE_INL_ BUBA Rev _{t+4}	AE_AUSL_ BUBA Rev _{t+4}	CPI Rev _{t+3}	
1,0000	-0,0308	0,0004	-0,0126	IPROD_OB Rev _{t+4}
	1,0000	-0,2041	0,0638	AE_INL_BUBA Rev _{t+4}
		1,0000	0,2199	AE_AUSL_BUBA Rev _{t+4}
			1,0000	CPI Rev _{t+3}

IPROD_OB Rev _{t+5}	AE_INL_ BUBA Rev _{t+5}	AE_AUSL_ BUBA Rev _{t+5}	CPI Rev _{t+4}	
1,0000	0,1579	-0,0538	-0,1181	IPROD_OB Rev _{t+5}
	1,0000	-0,1133	-0,0292	AE_INL_BUBA Rev _{t+5}
		1,0000	0,2076	AE_AUSL_BUBA Rev _{t+5}
			1,0000	CPI Rev _{t+4}

IPROD_OB Rev _{t+6}	AE_INL_ BUBA Rev _{t+6}	AE_AUSL_ BUBA Rev _{t+6}	CPI Rev _{t+5}	
1,0000	0,0567	0,1580	0,0104	IPROD_OB Rev _{t+6}
	1,0000	0,0525	0,2989	AE_INL_BUBA Rev _{t+6}
		1,0000	0,2721	AE_AUSL_BUBA Rev _{t+6}
			1,0000	CPI Rev _{t+5}

Anmerkungen: Jeweils Korrelationen zu den ersten bis vierten (von oben nach unten) Revisionen. Beobachtungen 1993m2–2010m10 (Fehlwerte wurden ausgelassen), 5% kritischer Wert (zweiseitig) = 0,1345 für n = 213.

nen Messungen der betreffenden Variablen. Bei der Industrieproduktion jedoch scheint die Umstellung der Erfassung um das Jahr 2000 herum tatsächlich dazu geführt zu haben, dass diese Unterschätzung abgenommen hat. Bei den Konsumentenpreisen beobachten wir dagegen eher eine anfängliche systematische Überschätzung der später festgeschriebenen Messungen.

Eine Analyse verschiedener Korrelationen der beobachteten Revisionen hat nur wenige deutliche Zusammenhänge aufgezeigt. So gibt es bei keiner der Variablen signifikante Korrelationen zwischen verschiedenen Revisionen bzgl. des selben Monats. Allerdings sind für aufeinander folgende Monatsberichte die jeweils ersten Revisionen der Industrieproduktionswerte korreliert; gleiches gilt auch für die Auftragseingänge aus dem Inland, sowie mit Abstrichen auch für die Konsumentenpreise. Für die Industrieproduktion gilt zudem, dass die Erstveröffentlichung selbst (d.h. die entsprechende implizite Wachstumsrate der Produktion) mit der zeitgleich veröffentlichten ersten Revision des vorherigen Monats negativ korreliert ist. Variablenübergreifend finden sich vor allem Zusammenhänge zwischen den (ersten) Revisionen der Auftragseingänge aus dem Inland mit der Industrieproduktion und auch mit den Auftragseingängen aus dem Ausland. Insgesamt bestätigen die Ergebnisse dieses Kapitels die Notwendigkeit, bei der Wendepunktbestimmung und -prognose die Echtzeitproblematik zu berücksichtigen.

4 Datierung von konjunkturellen Wendepunkten in Echtzeit

4.1 Verbreitete Methoden zur Datierung von konjunkturellen Wendepunkten

In einer der frühesten Untersuchungen der Eigenschaften von Konjunkturzyklen definierten Burns & Mitchell (1946) eine Rezession als „a substantial prolonged decline in economic activity that occurs broadly across various sectors of the economy“. Diese Definition beschreibt trotz ihrer Ungenauigkeit drei fundamentale Eigenschaften von ökonomischen Rezessionen: die Deutlichkeit des wirtschaftlichen Abschwungs, ihre nicht-triviale Dauer und die Tatsache, dass dieser Abschwung in vielen Teilen der Volkswirtschaft bemerkbar ist. Seitdem wurde diese grundsätzliche Definition der Hauptmerkmale einer Rezession mit weiterentwickelt. So ist die heute allgemein anerkannte Definition einer Rezession durch das NBER Business Cycle Dating Committee zu nennen (Business Cycle Dating Committee 2008, S.1), welche sich an der synchronisierten Bewegung (*Comovement*) ausgewählter makroökonomischer Größen orientiert:

A recession is a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale retail sales. A recession begins just after the economy reaches a peak of activity and ends as the economy reaches its trough. Between trough and peak, the economy is in an expansion.

Obwohl die NBER-Definition von konjunkturellen Wendepunkten auf den ersten Blick wie eine genaue und operationale Beschreibung konjunktureller Wendepunkte in den USA erscheint, ist die NBER-Methodik keineswegs frei von Kritik. Wie z.B. Chauvet & Pi-

ger (2008) anmerkten, ist die Datierung von Wendepunkten des NBER zum einen weder transparent noch replizierbar, da sie auf einem Konsens von Experten beruht, die verschiedene Techniken benutzen und verschiedene Ansichten bezüglich der jüngst vergangenen Wirtschaftslage vertreten. Zum anderen nimmt das NBER Dating Committee keine Revisionen seiner Datierung vor, obwohl die meisten Wirtschaftsdaten deutlichen Revisionen unterliegen. Darüber hinaus werden die konjunkturellen Wendepunkte seitens des NBER Business Cycle Dating Committee mit einer deutlichen Verzögerung veröffentlicht. Zum Beispiel wurden die Tiefpunkte der U.S.-Konjunktur von März 1991 und November 2001 mit nahezu zwei Jahren Verzögerung bekannt gegeben. Aufgrund all dieser Gründe ist es sinnvoll, eine klar definierte Methodik zur Datierung von konjunkturellen Wendepunkten zu verwenden, die vor allem durch ihre Transparenz die Nachvollziehbarkeit und Replizierbarkeit ihrer Ergebnisse gewährleistet.

Eine mögliche Art, das Phänomen des Konjunkturzyklus zu erfassen, stellt der parametrische Ansatz dar. Demnach wird der Konjunkturzyklus (und implizit die Datierung seiner Wendepunkte) anhand ökonometrischer Methoden geschätzt. Diesem Ansatz zu Grunde liegt die Vorstellung, dass der Konjunkturzyklus von einem stochastischen Prozess erzeugt wird, der nicht direkt beobachtbar ist. Gegeben ist demnach nur jeweils ein Pfad möglicher Realisierungen der beteiligten Variablen. Anhand ökonometrischer Verfahren zur Schätzung der unbeobachteten Variablen ist es dann möglich, aus beobachtbaren makroökonomischen Daten wie der Industrie-Produktion, dem BIP oder den Auftragseingängen Rückschlüsse auf den zu Grunde liegenden stochastischen Prozess zu ziehen. Durch die eindeutige Identifikation des zu Grunde liegenden stochastischen Prozesses lässt sich erschließen, wann ein konjunktureller Wendepunkt stattfindet. Auf dieser Vorstellung basieren z.B. sogenannte dynamische Faktormodelle (Geweke 1977, Sargent & Sims 1977, Stock & Watson 1989).

Gemäß dieser Literatur entsteht die synchronisierte Bewegung von Referenz-Zeitreihen, die langfristig durch ökonomische Beziehungen aneinander gekoppelt sind, durch gemeinsame Faktoren. Liegt nur ein solcher unbeobachtbarer Faktor vor, so kann dieser als gemeinsamer Zustand der Wirtschaft interpretiert werden. Technisch gesehen lassen sich Schätzungen für diesen Zustand mit dem Mittel der Hauptkomponentenanalyse gewinnen. In Stock & Watson (1989) zum Beispiel wird eine Rezession dann identifiziert, wenn die erste Differenz der Zustandsvariable für eine bestimmte Anzahl aufeinander folgender Monate unter eine gewisse Grenze fällt.

Es ist anzumerken, dass mit diesem parametrischen Ansatz wiederum auch die Vorstellung vereinbar ist, dass verschiedene stochastische Prozesse den verschiedenen Phasen des Konjunkturzyklus zu Grunde liegen (siehe z.B. Neftci 1982), d.h. dass wirtschaftliche Auf- und Abschwünge von zwei verschiedenen datengenerierenden Prozessen (data generating processes, DGP) erzeugt werden und je nach Konjunkturlage einer dieser Prozesse gültig ist. Nach Neftci (1982) würde sich der Konjunkturzyklus dann in einem Wendepunkt befinden, wenn der stochastische Prozess seine Wahrscheinlichkeitsstruktur (Dichte) ändert. Solche Probleme lassen sich anhand der Optimal-Stopping-Theorie lösen, so dass eine Erwartung für den Eintritt des Wendepunkts gewonnen werden kann. Diese Betrachtungsweise hat insbesondere den Vorteil, mögliche Asymmetrien in den verschiedenen Phasen des Konjunkturzyklus abbilden zu können.

Im Gegensatz zum parametrischen Ansatz versucht der nicht-parametrische Ansatz die Datierung von Ab- und Aufschwüngen der Wirtschaftsaktivität durch vorgegebene Algorithmen festzulegen, welche auf bestimmten Mustern in der beobachteten Variable basieren. Dieser Beobachtungsansatz wurde durch den grundlegenden Beitrag von Bry & Boschan (1971) eingeführt und von Harding & Pagan (2002) zu einem multivariaten Ansatz erweitert. Bei den hier vorgestellten Ansätzen zur ex post-Datierung handelt es sich in Abgrenzung zu den im weiteren Projektverlauf benutzten Methoden um nicht-parametrische. Dies bedeutet, dass kein die Zeitreihe generierender Prozess inklusive der Schätzung seiner Parameter zur Wendepunktbestimmung benutzt wird, sondern vielmehr ein Algorithmus, dessen Entscheidungsregeln mehr oder weniger den Konsens in der Literatur abbilden. Grundsätzlich kommen dabei als Wendepunkt zwischen Rezession und Expansion nur lokale Minima und als Wendepunkt zwischen Expansion und Rezession nur lokale Maxima in Frage. Umgekehrt bedeutet dies jedoch nicht, dass alle lokalen Extrema Wendepunkte darstellen. Betrachtet man einen Wendepunkt in der Gegenwart, so wird zur Bestimmung der Eigenschaft als Extremum ein Blick in die Zukunft notwendig. Dementsprechend sind nicht-parametrische Ansätze ausschließlich zur Expost-Datierung geeignet.

In Hinblick auf die Datierung konjunktureller Wendepunkte unter Echtzeitbedingungen hat der nicht-parametrische Ansatz einen entscheidenden Vorteil gegenüber dem parametrischen Ansatz: Durch die direkte Evaluierung ökonomischer Zeitreihen werden mögliche Fehler, die aus der falschen Wahl eines ökonometrischen Schätzverfahrens und/oder aus einer möglichen Fehlspezifikation eines solchen resultieren, vermieden. Aus diesem Grund werden im Folgenden zwei der wichtigsten nicht-parametrischen Methoden im Detail dis-

kutiert.

4.2 Das Bry-Boschan und das Harding-Pagan Verfahren

Zu den am weitesten verbreiteten nicht-parametrischen Datierungsverfahren konjunktureller Wendepunkte zählen die Ansätze von Bry & Boschan (1971) sowie Harding & Pagan (2002), die uns einerseits als Basis für die Erzeugung einer binären Rezessionsvariable (vgl. Kapitel 8) und andererseits als Benchmark für die Qualität geschätzter Regimewahrscheinlichkeiten dienen. Daher sollen diese Ansätze hier zunächst eingehend erläutert werden, bevor auf die Veränderung ihrer Datierungsergebnisse unter Echtzeitbedingungen eingegangen wird. Ein Schwerpunkt der Beschreibung liegt dabei auf den ökonomischen Grundlagen der Methodik sowie gewählter Schwellenwerte.

In diesem Abschnitt werden zunächst Probleme der Wendepunkterkennung und die daraus resultierenden Konsequenzen für den Aufbau eines Datierungsalgorithmus erläutert. Details zu dem dabei verwendeten Vorgehen, insbesondere zur Wahl der Schwellenwerte, werden im Abschnitt über die implementierte Version der *ex post*-Datierung behandelt. Tatsächlich müssen bei der Entscheidung, ob ein Wendepunkt vorliegt oder nicht, die im Folgenden vorgestellten Probleme oft gleichzeitig beachtet werden. Der Übersichtlichkeit halber werden diese Probleme und die Mittel zu ihrer Behebung jedoch anhand der in Abbildung 4.1 stilisierten Form nacheinander beschrieben.

Teilabbildung A betrachtet das Problem des zusätzlichen Zyklus, der im Gesamtverlauf jedoch nur als Teil einer Phase, hier einer Expansion, zu verstehen ist. Bry & Boschan (1971) nennen zwei Kriterien, um ein fehlerhaftes Erkennen der Wendepunkte im Extrazyklus zu vermeiden: Erstens muss jeder erkannte Zyklus eine gewisse Mindestdauer erreichen. Zweitens soll die Amplitude eines strittigen Zyklus die geringste Amplitude der unstrittigen Zyklen übertreffen.

Teilabbildung B stellt das Gegenteil des vorherigen Problems, den ausgelassenen Zyklus, dar. Klassisches Beispiel für ein solches Auftreten wäre die Betrachtung eines Wachstumssektors, so dass sich hierfür trotz rückläufiger gesamtwirtschaftlicher Entwicklung positive Wachstumsraten ergeben. Da das Produzierende Gewerbe ohne Bau einen Großteil der Bruttowertschöpfung Deutschlands über verschiedene Sektoren hinweg abdeckt, kann eine fehlerhafte Datierung aufgrund dieses Phänomens in der vorliegenden Untersuchung

ausgeschlossen werden.

Teilabbildung C behandelt zwei Phänomene: Erstens Ausreißer und zweitens Punkte des gleichen Betrags. Zunächst gilt, dass Ausreißer nur dann fehlerhaft als Wendepunkte erkannt werden können, wenn sie als lokale Extrema erkannt werden. Das bedingt, dass sie in der Umgebung der tatsächlichen Wendepunkte diese betragsmäßig überschreiten. Liegen Ausreißer nicht in der Nähe der tatsächlichen Wendepunkte, so ist das Phänomen dem des Extrazyklus sehr ähnlich. Eine Möglichkeit, Ausreißern zu begegnen, ohne die zu Grunde liegende Zeitreihe nur an einzelnen Punkten zu verändern, besteht in der Glättung der Reihe. Um dem Phänomen zu begegnen, dass mehrere Punkte den selben Wert haben, empfiehlt sich die Verlängerung des aktuellen Regimes bis zum letzten dieser Punkte. Demgegenüber wird jedoch bei aufeinander folgenden Extrema der gleichen Art aber unterschiedlicher Höhe der betragsmäßig größere ausgewählt (Teilabbildung D).

Die in den Teilabbildungen E und F dargestellte Problematik, dass Reihen eher stufen- als wellenförmig verlaufen, spielt bei den gewählten Daten monatlicher Frequenz eine untergeordnete Rolle (vgl. Abbildung 4.4). Schließlich bleibt die Frage, ob der Datierungsalgorithmus eher auf den Originaldaten oder einer trendbereinigten Version arbeiten sollte (Teilabbildung G). Obwohl die Produktion des Produzierenden Gewerbes ohne Bau einen stark positiven Trend aufweist, zumindest bis zur Finanzmarktkrise (Abbildung 4.5), gibt es gute Gründe, die gegen die Verwendung einer trendbereinigten Version der Reihe sprechen. Zunächst einmal gilt, dass unterschiedliche Trendbereinigungsverfahren, etwa die Verwendung der ersten Differenz oder des Hodrick-Prescott (HP) Filters (Teilabbildung G) auch Verschiebungen bei der Wendepunktatierung bedingen. Des Weiteren sei die Tatsache betont, dass die einfach zu definierende Menge lokaler Extrema in der Levelbetrachtung, etwa bestimmt durch

$$\{y_t : y_t > y_{t \pm k}, k = 1, \dots, 5\} \cup \{y_t : y_t < y_{t \pm k}, k = 1, \dots, 5\}, \quad (4.1)$$

keineswegs genauso einfach in der ersten Differenz, etwa bestimmt durch

$$\begin{aligned} & \{y_t : \Delta y_{t-k} > 0, k = 0, \dots, 4 \quad \text{und} \quad y_t : \Delta y_{t-k} < 0, k = 1, \dots, 5\} \cup \\ & \{y_t : \Delta y_{t-k} < 0, k = 0, \dots, 4 \quad \text{und} \quad y_t : \Delta y_{t-k} > 0, k = 1, \dots, 5\}, \end{aligned} \quad (4.2)$$

gegeben ist. Dazu folgendes Beispiel: Gegeben seien etwa zwei aufeinander folgende Extrema der gleichen Art, zwischen denen eine kurzzeitige Niveaumäßigung vorliegt. Nach

PROBLEMS OF TURNING POINT DETERMINATION

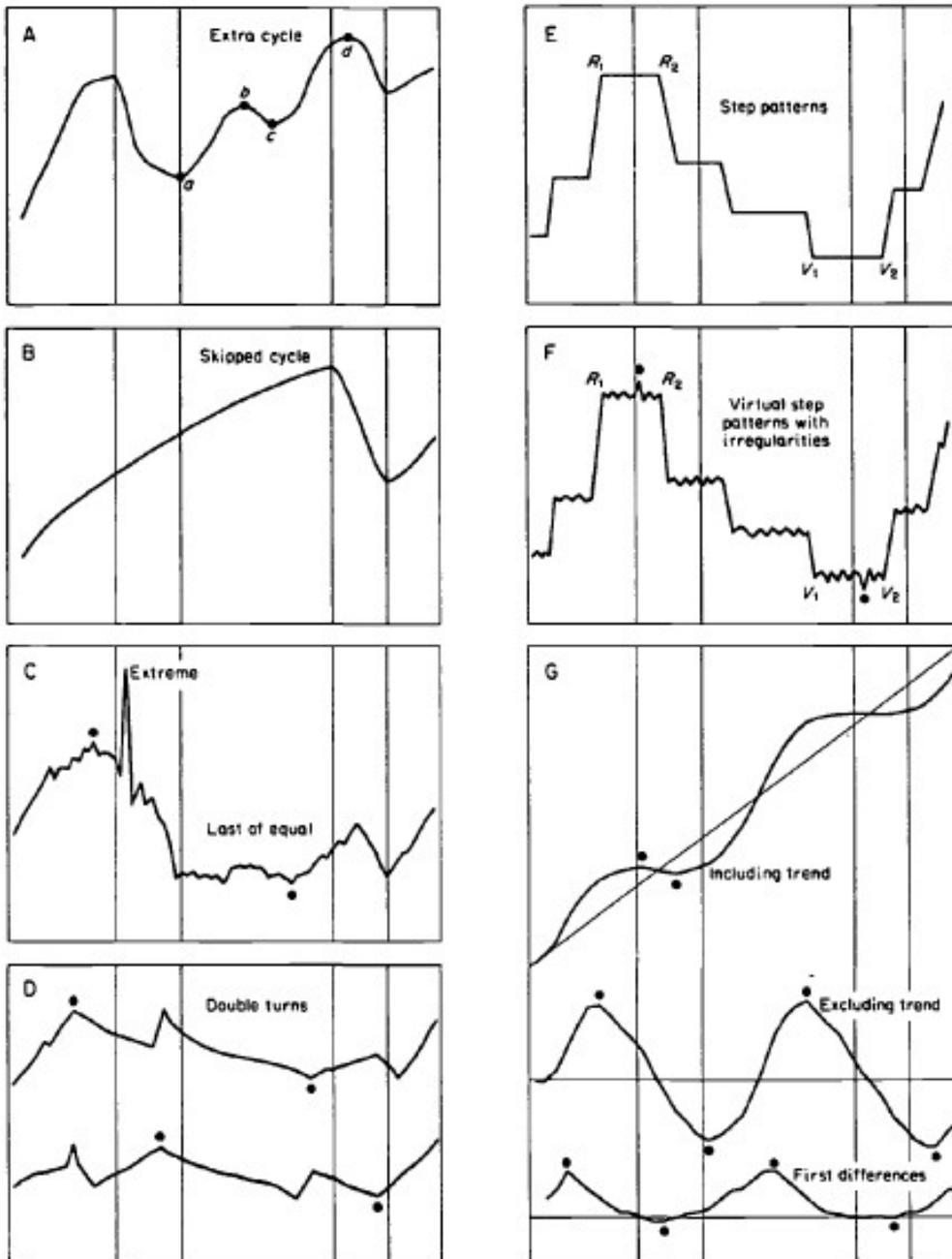


Abbildung 4.1: Probleme der Wendepunkterkennung. Die Punkte geben die korrekte Datierung in stilisierten spezifischen Zyklen wieder, während die Linien den Wechsel von Hoch- und Tiefpunkten in der allgemeinen Wirtschaftsaktivität ausdrücken. Quelle: Bry & Boschan (1971).

(4.1) handelt es sich bei einem dieser Punkte sehr wohl um einen potentiellen Wendepunkt, während (4.2) dies ausschließt. Würde man andererseits auf Punkte des größten Wachstums, bei typischem Verlauf mathematische Wendepunkte, abzielen, wäre dies nicht in Einklang mit der ökonomisch intuitiven Definition zu bringen, dass eine Rezession unmittelbar nach einem Hochpunkt gesamtwirtschaftlicher Aktivität und eine Expansion unmittelbar nach einem Tiefpunkt, also bei den mathematischen Extrempunkten, beginnt. Zusammenfassend führt dies zu der Entscheidung, die algorithmische ex post-Datierung auf dem nicht trendbereinigten Niveau des Outputs auszuführen. Dies schließt jedoch nicht aus, dass für die später angewendeten, parametrischen Verfahren Differenzen oder Wachstumsraten verwendet werden, sofern es das jeweilige Schätzverfahren erfordert.

In diesem Abschnitt werden Entscheidungsregeln der algorithmischen ex post-Datierung vorgestellt, die dazu führen, dass nicht jedes lokale Maximum oder Minimum automatisch einen Wendepunkt zwischen den Regimen darstellt. Als maßgebliche Kenngrößen dieser Selektion sind Dauer und Amplitude des Zyklus zu nennen, die in dem von Harding & Pagan (2002) eingeführten Dreiecksansatz gebündelt werden können (vgl. Abbildung 4.2).

Dieser Ansatz wird im Folgenden grundsätzlich beschrieben, während die exakte Umsetzung, insbesondere die zu wählenden Schwellenwerte, im Abschnitt über die implementierte Version der ex post-Datierung, behandelt wird. Während es für die Festlegung der Minstdauer einer Rezessionsphase sowie des gesamten Konjunkturzyklus relativ leicht ist, sich innerhalb eines Konsensrahmens zu bewegen, ist die Bestimmung einer Mindesthöhe der Rezessionsamplitude mit Schwierigkeiten verbunden, da diese letztlich von der in ihrer zeitlichen Umgebung vorherrschenden Volatilität abhängen muss. Daher wird im Dreiecksansatz statt einer Mindestamplitude das Überschreiten eines Schwellenwerts durch den aus Amplitude und Dauer gebildeten Flächeninhalt,

$$\Delta = \frac{1}{2} |t_P - t_T| |y_P - y_T|, \quad (4.3)$$

gefordert. Dabei bezeichnet P den Hochpunkt (*englisch*: peak), T den Tiefpunkt (*englisch*: trough), t die Zeit und y den Output, also die Produktion des Produzierenden Gewerbes ohne Bau. Statt vom Flächeninhalt sprechen Harding & Pagan (2002) vom kumulierten Verlust, wobei sich ihr Verlustbegriff sowohl auf die Tiefe des Abschwungs, also den eigentlichen Verlust, als auch auf die zeitliche Dauer bezieht. Dahinter steckt die Idee, dass eine kürzere Rezession bei gleicher Amplitude weniger gravierend ausfällt, da auf den längeren Zeitraum gesehen schon eine Erholung im Outputniveau einsetzt.

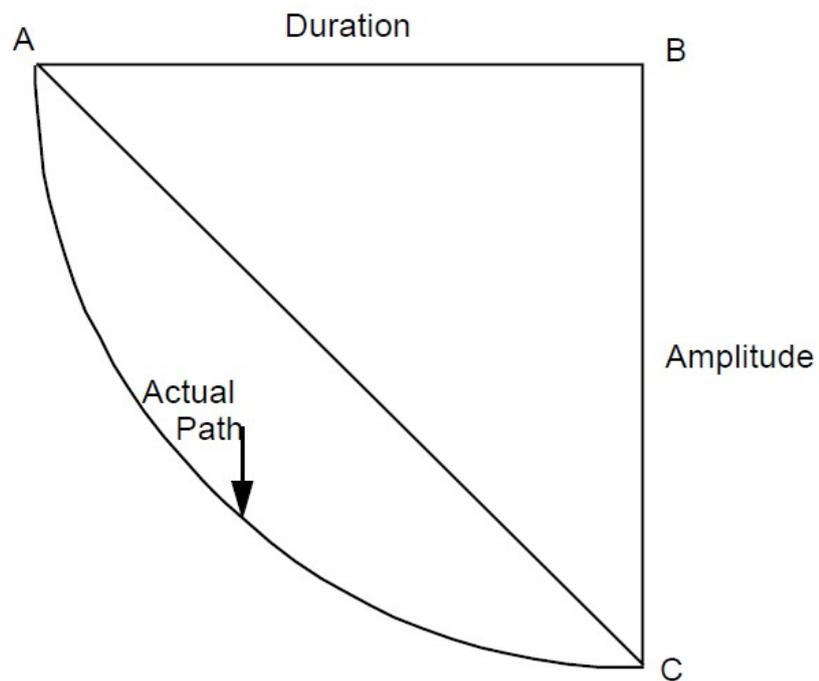


Abbildung 4.2: Stilisierte Abbildung einer Rezessionsphase. Die Rezession verläuft von A nach C. Dauer und Amplitude formen die Seiten eines Dreiecks. Quelle: Harding & Pagan (2002).

Viele Untersuchungen zeigen die Asymmetrien zwischen Rezessions- und Expansionsphasen (etwa Friedman 1993). Dies spiegelt sich auch im Dreiecksansatz wieder, wenn man den sogenannten überschüssigen kumulierten Verlust einer Phase i ,

$$E_i = \Delta_i - \tilde{\Delta}_i,$$

betrachtet, wobei Δ_i der Dreiecksapproximation und $\tilde{\Delta}_i$ der tatsächlichen Fläche unter dem aktuellen Pfad, die sich bspw. durch Integration bestimmen lässt, entspricht. So stellen Harding & Pagan (2002) für mehrere Volkswirtschaften fest, dass der überschüssige kumulierte Verlust in Kontraktionsphasen deutlich geringer ausfällt als in Expansionsphasen. Auch der Tatsache vorliegender Asymmetrien soll daher die implementierte Version gerecht werden.

4.3 Implementierung nicht-parametrischer Wendepunktatierung in Echtzeit

Als allgemein anerkannte Rezessionsdefinition eignet sich die oben erwähnte des National Bureau of Economic Research (NBER 2010) der Vereinigten Staaten, nach der eine Rezession eine signifikante Abwärtsbewegung gesamtwirtschaftlicher Aktivität darstellt, die sich über die meisten Wirtschaftsbereiche erstreckt, über mehrere Monate anhält und normalerweise sichtbar ist in den makroökonomischen Größen reales Bruttoinlandsprodukt, reales Einkommen, Beschäftigung, Industrieproduktion sowie Groß- und Einzelhandelsumsätze. Daher können die Verläufe dieser Größen auch als Referenzzyklen (*reference cycles*) bezeichnet werden. Demgegenüber versteht man unter den spezifischen Zyklen (*specific cycles*) die Pfade vorlaufender Indikatoren, die die frühzeitige Erkennung einer Rezession ermöglichen. Diese Unterscheidung zwischen Referenz- und spezifischem Zyklus, die erstmals von Burns & Mitchell (1946) eingeführt wurde und seitdem die Diskussion prägt, welche (einzelne) makroökonomische Größe am besten dem Konjunkturzyklus (*business cycle*) entspricht, ist ebenfalls wesentlich für die hier vorgestellte ex post-Datierung.

4.3.1 Auswahl der Referenz-Zeitreihe

Im Hinblick auf die Zielsetzung des vorliegenden Forschungsauftrags ist es zwar grundsätzlich sinnvoll, die Definition von Rezessions- und Expansionsphasen auf das Brutto-

inlandsprodukt zu beziehen. Jedoch eignen sich Zeitreihen monatlicher Frequenz besser für die Wendepunkterkennung am aktuellen Rand, da eine Datenaktualisierung nicht erst zum Ende des Quartals erfolgt, so dass zwischenzeitlich überhaupt eine Früherkennung möglich ist. Da Daten für das BIP in Deutschland nur auf Quartalsbasis verfügbar sind, ist es notwendig herauszufinden, welche der oben genannten alternativen Referenzzyklen, die auf Monatsbasis veröffentlicht werden, am besten als Proxy für die gesamtwirtschaftliche Entwicklung geeignet sind. Hierzu dient eine Korrelationsanalyse der entsprechenden Variablen.

Entsprechend der Definition des NBER kommen als Referenzzyklen außer dem BIP also Einkommen, Beschäftigung, Industrieproduktion und Einzelhandelsumsätze in Frage. Aus der Echtzeitdatenbank der Bundesbank eignen sich die folgenden Monatszeitreihen als Proxys für die genannten Indikatoren: Um die Beschäftigung abzubilden, bieten sich die Erwerbstätigenstunden im Verarbeitenden Gewerbe (BE2AA022), die sozialversicherungspflichtig Beschäftigten (BG1CA010) oder die Arbeitslosenquote (USCC02) an.

Da konjunkturelle Spitzen bei den Unternehmen oftmals durch Überstunden der Beschäftigten, und konjunkturelle Eintrübungen zunächst durch einen entsprechenden Rückgang der Arbeitsstunden abgefangen werden, sind konjunkturelle Schwankungen nicht immer in vollem Ausmaß in der Statistik der Arbeitslosenquote oder der sozialversicherungspflichtig Beschäftigten zu sehen. Demgegenüber wirken sich sowohl Variationen in der Anzahl der Stunden als auch in der Anzahl der Beschäftigten auf die Erwerbstätigenstunden aus, weshalb diese am ehesten eine starke Korrelation mit den konjunkturellen Schwankungen des BIP vermuten lassen.

Zur Darstellung der Einzelhandelsumsätze stehen in der Datenbank Zeitreihen über die Umsätze des Einzelhandels ohne den Einzelhandel mit Kraftfahrzeugen (AGA01) als auch Umsätze des Einzelhandels inklusive des Einzelhandels mit Kraftfahrzeugen (IT1AGC02) zur Verfügung. Letztere ist allerdings in ihrem Verlauf stark von nichtzyklisch oder antizyklisch wirkenden Sondereffekten, wie z.B. der Abwrackprämie oder der Erhöhung der Umsatzsteuer, geprägt, weshalb die Reihe ohne KFZ vorzuziehen ist. Anstelle der Industrieproduktion wird, wie in Kapitel 2 erläutert, die Zeitreihe Produzierendes Gewerbe ohne Bau (IP1AA020) herangezogen.

Grundsätzlich sind alle betrachteten Zeitreihen kalender- und saisonbereinigt. In Abgrenzung zu zeitraumbezogenen Veränderungen eines Bestandes (Flussgrößen) handelt es sich bei den sozialversicherungspflichtig Beschäftigten und bei der Arbeitslosenquote direkt um eine zeitpunktsbezogene Bestandsgröße. Die übrigen Zeitreihen sind Flussgrößen.

4 Datierung von konjunkturellen Wendepunkten in Echtzeit

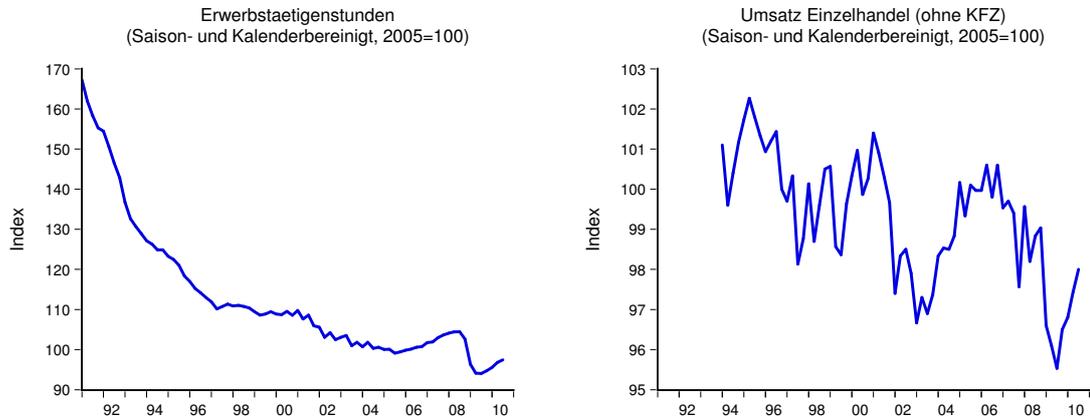


Abbildung 4.3: Darstellung des Einzelhandelsumsatzes (ohne KFZ) und der Erwerbstätigenstunden im Bereich Verarbeitendes Gewerbe von Q1:1991 bis Q3:2010. Quelle: Deutsche Bundesbank.

Zusammenfassend kommen also die Zeitreihen Erwerbstätigenstunden, Einzelhandelsumsätze ohne den Einzelhandel mit Kraftfahrzeugen und Produzierendes Gewerbe ohne Bau am ehesten als Proxies für das BIP in Frage und werden auf Quartalsbasis umgerechnet, um ihre Korrelation mit dem BIP zu testen.¹

Bevor die Korrelationen zwischen den gewählten Proxyvariablen und dem BIP gemessen werden können, müssen die Monatsdaten in Quartalsdaten umgerechnet werden. Für die Flussgrößen (Erwerbstätigenstunden, Umsätze des Einzelhandels und Produktion) wurden aus den jeweiligen Monatsindizes durch Durchschnittsbildung die Quartalsindizes berechnet.

In Abbildung 4.3 sind die Zeitreihen für den Umsatz im Einzelhandel sowie für die Erwerbstätigenstunden dargestellt. Bei Ersterem fällt ein ungebrochener Abwärtstrend auf. Intuitiv würde man in einer Volkswirtschaft mit einem wachsenden Bruttoinlandsprodukt aber einen Aufwärtstrend, d.h. einen Anstieg der Umsätze erwarten. Der hier zu beobachtende Abwärtstrend könnte sich durch die Erfassungsmethode erklären. Da für die Zeitreihe nur Unternehmen mit Umsätzen über 250.000 EUR erfasst werden, spiegelt sich eine Verlagerung von Umsätzen in kleinere Unternehmen in einem Rückgang der Umsätze in dieser Zeitreihe wieder.

Bei den Erwerbstätigenstunden ist bis 1997 ein starker Abwärtstrend festzustellen, der

¹Die Korrelationen wurden auch für die anderen Variablen getestet; diese waren, wie erwartet, durchgehend geringer.

Tabelle 4.1: Korrelationskoeffizienten zwischen den Wachstumsraten des Bruttoinlandsprodukts und ausgewählten Variablen

	Produktion	UmsatzEHoK	Erwerbstätigenstunden
BIP	0,7944	0,4685	0,4921

sich in den Folgejahren deutlich abschwächt. Der Abwärtstrend in der Stundenzahl im produzierenden Gewerbe ergibt sich maßgeblich durch Produktivitätssteigerungen, die stärkere Fokussierung auf die Erstellung von Dienstleistungen, den Verlust von Industriearbeitsplätzen in den neuen Bundesländern sowie durch die Reduzierung der durchschnittlichen tariflichen Wochenarbeitszeit. Die beschriebenen Trends könnten in der Korrelationsanalyse zu Scheinkorrelationen führen. Um dies zu vermeiden wurden die Log-Differenzen der gewählten Variablen in der Korrelationsanalyse genutzt, sodass nicht die Indexwerte selbst, sondern deren Wachstumsraten verglichen werden.

In Tabelle 4.1 sind die errechneten Korrelationskoeffizienten dargestellt. Mit einer Korrelation von fast 0,8 werden die Schwankungen im BIP-Wachstum am Besten von der Zeitreihe Produktion im Produzierenden Gewerbe wiedergegeben, weshalb diese als Proxy für die BIP-Entwicklung genutzt werden wird. Dies deckt sich auch mit der Beobachtung von Fritsche & Stephan (2002), nach der die Nutzung der Industrieproduktion als beste Approximation für Konjunkturbewegungen geeignet ist, da hierin der volatile Teil des BIP (Investitionen und Exporte aus der Industrieproduktion) aufgefangen wird.

Des Weiteren ist zu beachten, dass bei der monatlichen Produktionsreihe bereits solche Phasen als Rezession wahrgenommen werden, die im BIP als Trendunterbrechung erscheinen. So ist mit Ausnahme des starken Abschwungs während der Finanzmarktkrise im BIP kein wirklich zyklisches Verhalten feststellbar. Dies ist für eine algorithmische Wendepunkterkennung insofern von Bedeutung, da aufgrund der geringen Volatilität etwa die Identifikation lokaler Extrema nicht immer eindeutig möglich ist (vgl. E und F in Abbildung 4.1).

Schließlich bleibt zu erwähnen, dass die algorithmische ex post-Datierung auch auf die Auftragseingänge aus dem In- und Ausland (spezifische Zyklen) angewendet wird,² um den Vorlauf ihrer Wendepunkte unter Echtzeitbedingungen zu untersuchen. Die Verläufe

²Bezeichnung in der Echtzeitdatenbank der Bundesbank: IO2ABA20/ACM01 und IO3ABA20/ACM01. Die Deutsche Bundesbank führt hier bis 2001 das Verarbeitende Gewerbe und ab 2001 die Industrieproduktion als Bezugsgröße. Der Bezugswechsel ist allerdings hinsichtlich der Auswirkungen auf die Bruttowertschöpfung vernachlässigbar.

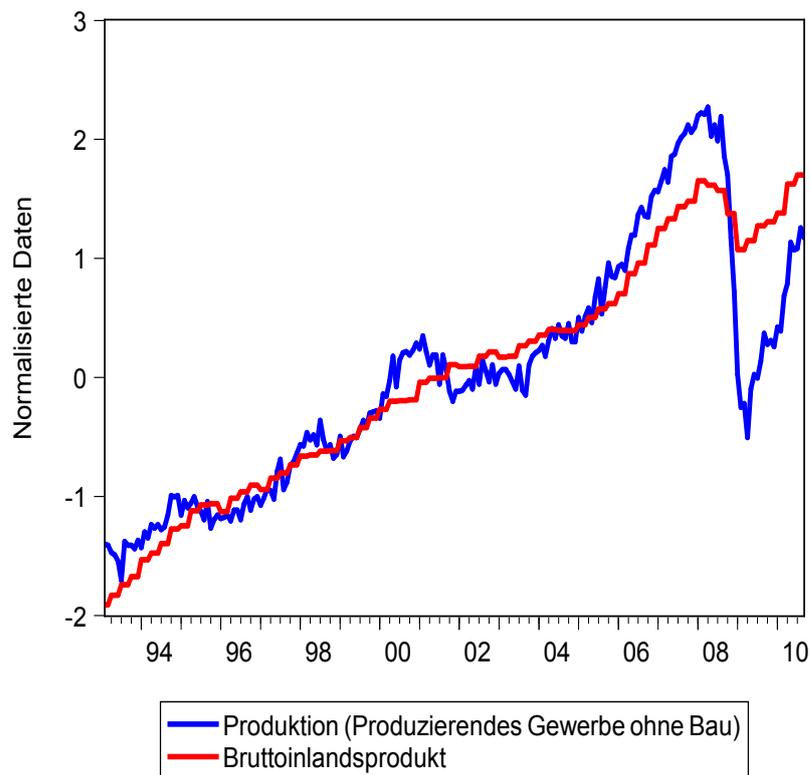


Abbildung 4.4: Ökonomische Gesamtaktivität. Normalisierte Darstellung von Bruttoinlandsprodukt (Quartalsdaten, AG1CA010) und Produktion des Produzierenden Gewerbes ohne Bau (Monatsdaten, IP1AA020) in 2005er Preisen. Beide Zeitreihen sind kalender- und saisonbereinigt. Stand November 2010. Quelle: Deutsche Bundesbank, Echtzeitdatenbank.

dieser Reihen stellen die einzigen hier untersuchten spezifischen Zyklen dar, da nur für sie die beiden Kriterien, Indikatoreigenschaft mit vorlaufendem Charakter sowie ausreichende Echtzeitdatenhistorie, erfüllt waren.

4.3.2 Detaillierte Beschreibung des Datierungsalgorithmus

In diesem Abschnitt wird die implementierte Version vorgenannter Entscheidungsregeln zur ex post-Wendepunkt-datierung vorgestellt. Dazu werden zunächst die Schritte des Datierungsalgorithmus aufgelistet, ehe im Anschluss eine weitere Auseinandersetzung mit den dabei maßgeblichen Schwellenwerten erfolgt. Grundsätzlich ist der Algorithmus auf eine sensitive Rezessionserkennung angelegt, d.h. dass das Nichterkennen einer Rezessi-

on als der maßgeblichere Fehler gegenüber dem Nichterkennen einer Expansion gewertet wird.

4.3.2.1 Behandlung von Ausreißern

Um der Problematik von Ausreißern gerecht zu werden, wird die Produktionsreihe durch einen gleitenden Durchschnitt der kleinstmöglichen Ordnung 2 geglättet. Auf dieser Grundlage erfolgt die Bestimmung lokaler Extrema, während darauf folgende Selektionen nach der Amplitude die Größenordnung der Originaldaten berücksichtigen. Das Vorgehen stellt einen Kompromiss zwischen der unerwünschten Verwendung von Ausreißern und der unerwünschten Glättung von Phasen höherer Volatilität dar.

4.3.2.2 Identifikation von lokalen Hoch- und Tiefpunkten der Referenz-Zeitreihe

Lokale Extrema werden anhand der Kriterien von (4.1) bestimmt. Zusätzlich wird eine an Phasen geringer Volatilität angepasste Mindestamplitude gemäß dem Dreiecksansatz verlangt. Um der Asymmetrie zwischen den Regimen gerecht zu werden, liegt diese für Rezessionen bei $-0,01$ und für Expansionen bei $+0,10$. Die genannten Werte beziehen sich direkt auf den von der Deutschen Bundesbank verwendeten Index (Basisjahr 2005). In diesem Schritt reagiert der Algorithmus dem Betrag nach also wesentlich sensitiver auf mögliche Rezessionen (vgl. auch Abbildung 4.12).

1. Die Phasen zwischen lokalem Maximum und lokalem Minimum stellen mögliche Rezessionen dar. Werden zwei aufeinander folgende Extrema der gleichen Art berichtet, greifen die im Abschnitt über die Probleme der Wendepunkterkennung erwähnten Regeln, so dass Hoch- und Tiefpunkte stets alternieren.
2. Durch Festlegen der Rezessionen mittels geforderter Mindestdauer der Phase (4 Monate), geforderter Mindestdauer des gesamten Zyklus (14 Monate) und Überschreiten des kumulierten Verlustes in Höhe von $2,5\%$ werden für jede Veröffentlichung die entsprechenden Berichtsmonate den Expansions- und Rezessionsphasen (vgl. Abbildungen 4.6 und 4.7) zugeordnet.
3. Die Datierung am Berichtsrand erfolgt unter Verwendung leicht abgeänderter Re-

geln,³ um der Tatsache gerecht zu werden, dass der Zyklus hier in der Regel nicht vollständig vorliegt.

Die Mindestdauer der Phase und des Zyklus lehnen sich an die in Bry & Boschan (1971) formulierten Vorgaben des National Bureau of Economic Research (NBER) an. Um jedoch der Tatsache Rechnung zu tragen, dass das Nichterkennen einer Rezession gravierender wirkt als die fehlende Datierung einer Expansion, wurde jeweils ein Monat weniger an Mindestdauer gefordert. Damit bewegen sich die Kriterien aber weiter im Konsensrahmen der Literatur.

Bezüglich des in Gleichung (4.3) aufgeführten kumulierten Verlustes wird eine Änderung von $|y_P - y_T|$ zu $|y_P - y_T|/y_P$ vorgenommen, d.h. statt der absoluten Tiefe der Rezession wird die relative betrachtet. Dementsprechend wird das Dreieck in Abbildung 4.2 gestaucht, ohne dass die illustrative Aussagekraft an Wirkung verliert. Da jedoch für den Schwellenwert des kumulierten Verlustes in der Literatur kein eindeutiger Richtwert (vgl. etwa Darne & Ferrara 2009) besteht und Harding & Pagan (2002) diese Größe vorrangig komparativ verwenden, erhält der gewählte Schwellenwert durch dieses Vorgehen eine anschauliche Bedeutung. So muss etwa die relative Amplitude am Ende einer Rezession mit einer Dauer von 10 Monaten mindestens 0,5% betragen. Dauert die Rezession nur 5 Monate, so muss entsprechend der Rückgang der gesamtwirtschaftlichen Aktivität bei mindestens 1% des Ausgangsniveaus (lokales Maximum) liegen. Wie den Abbildungen 4.6 und 4.7 zu entnehmen ist, handelt es sich dabei um ein trennscharfes Vorgehen, das gleichzeitig versucht, den Fehler der Nichterkennung einer Rezession als denjenigen erster Art zu behandeln.

4.4 Evaluierung von Echtzeit-Aspekten

In diesem Abschnitt werden die Resultate der Implementierung eines nicht-parametrischen Algorithmus zur ex post-Datierung von konjunkturellen Wendepunkten im Geiste der Algorithmen von Bry & Boschan (1971) und Harding & Pagan (2002) unter Echtzeitbedin-

³Die vorhandene Echtzeitdatenhistorie der Bundesbank startet zum Berichtsmonat Februar 1993, d.h. in einer Rezessionsphase. Um das Ende dieser Phase zu datieren, wird das Kriterium der Länge des Gesamtzyklus vernachlässigt, da die entsprechende Information vor Berichtsbeginn liegt. Der Schwellenwert des relativen, kumulierten Verlustes wird am Berichtsanfang und -ende um 0,5% auf 2% herabgesetzt. Wird zudem mindestens vier Monate vor Ende einer Veröffentlichung ein Hochpunkt ohne darauf folgenden Tiefpunkt detektiert, so werden alle Beobachtungen ab diesem unmittelbar einer Rezession zugeordnet, da die Phase zur Einhaltung der weiteren Kriterien noch nicht lange genug währt.

gungen evaluiert. Der Begriff *ex post* lässt sich dabei konkretisieren. Da Produktionsdaten einer Datenverfügbarkeitslücke von 2 Monaten unterliegen und lokale Extrema über eine Umgebung von 5 Monaten gefunden werden, kann der Algorithmus genau 7 Monate später anhand der vorhandenen Echtzeitwerten entscheiden, ob ein konjunktureller Wendepunkt vorliegt. Hierzu betrachte man zunächst den Verlauf der Echtzeitproduktionswerte, welche die Grundlage der Datierung bilden (vgl. Abbildung 4.5).

Wie den Abbildungen 4.6 und 4.7 zu entnehmen ist, werden über den Berichtsverlauf sechs durchgängige Rezessionen, eine nicht durchgängige sowie eine im Veröffentlichungsverlauf nur kurzfristig als Rezession eingestufte Phase (Veröffentlichungen zu Beginn 1999) erkannt. Grundsätzlich sind im Echtzeitverlauf Datierungsveränderungen den beiden folgenden Quellen zuzuordnen:

- Durch Revisionen verursachte Änderungen sind so groß, dass sich lokale Extrema verschieben oder aufheben.
- Durch Revisionen verursachte Änderungen sind so groß, dass die aus Dauer und Amplitude zusammengesetzten Kenngrößen des Dreiecksansatzes die zuvor überschrittenen Schwellenwerte unterschreiten (oder umgekehrt).

Beide Gründe werden als gleichberechtigt hinsichtlich ihrer Bedeutung für die Datierungsentscheidung eingestuft, so dass es in den Abbildungen 4.6 und 4.7 letztlich keine Rolle spielt, aufgrund welcher Ursache die Datierungsveränderung im Echtzeitverlauf zustande kommt. Die Beschreibung der Echtzeitveränderungen aller Rezessionsphasen wird mit folgender Unterscheidung der wirtschaftspolitischen Auswirkungen vollzogen:

1. Als wirtschaftspolitisch am relevantesten werden Datierungsveränderungen eingestuft, welche den Rezessionsbeginn früher deklarieren. Ebenfalls der ersten Prioritätsstufe zugeordnet werden Rezessionsphasen, die im Veröffentlichungsverlauf zusätzlich deklariert werden.
2. Zu der zweiten Prioritätsstufe gehören Tiefenveränderungen (relative Amplitude) der Rezession, welche die Stärke notwendiger wirtschaftspolitischer Maßnahmen tangieren.
3. Schließlich sind Verlängerungen der Rezessionsphasen im Veröffentlichungsverlauf zu nennen. In Abgrenzung zu 1 wird hierbei das Rezessionsende durch die Revision verschoben. Dieser Sachverhalt wird hinsichtlich seiner wirtschaftspolitischen Auswirkungen als am wenigsten wichtig eingestuft.

4 Datierung von konjunkturellen Wendepunkten in Echtzeit

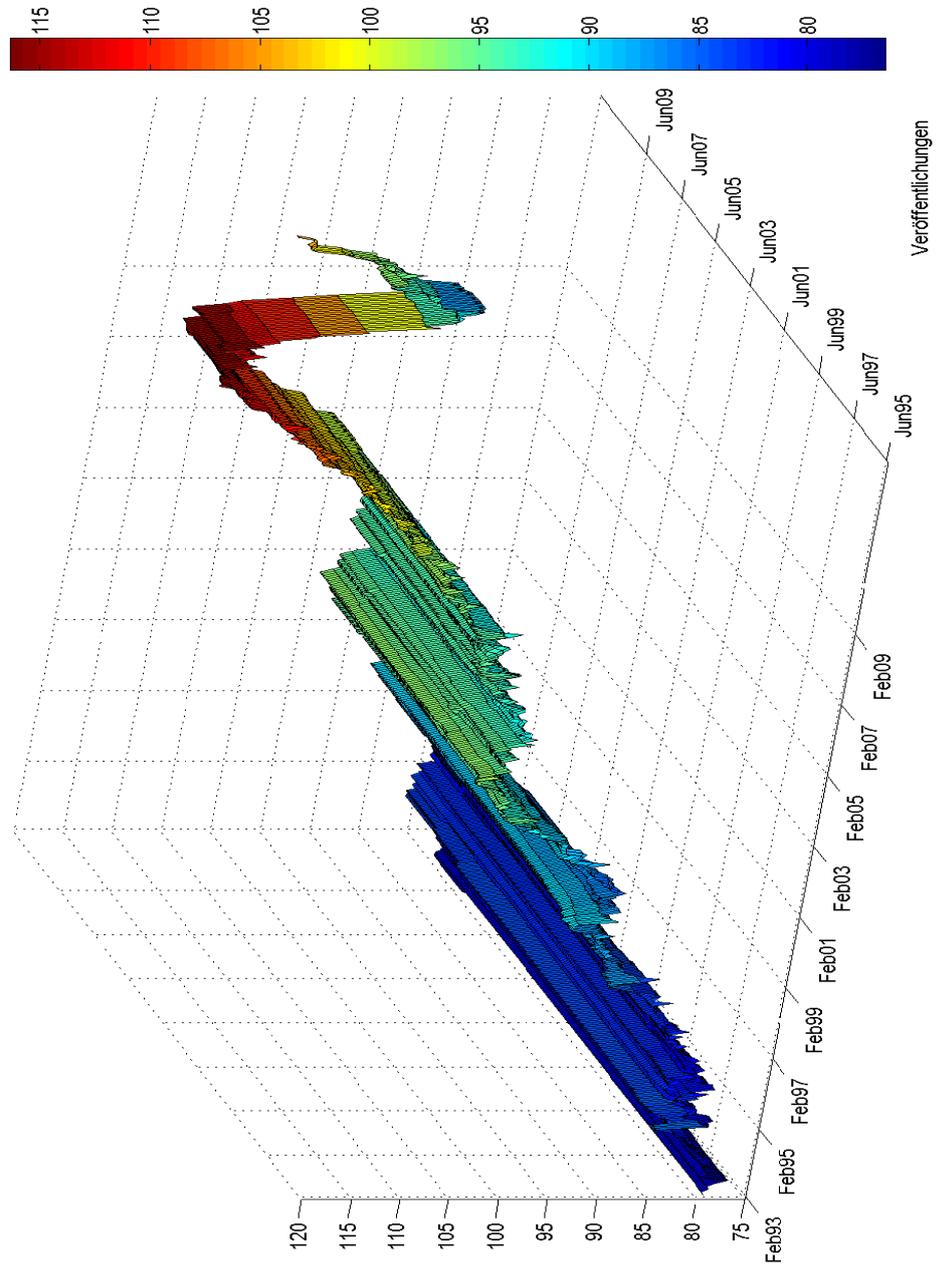


Abbildung 4.5: Produktion des Produzierenden Gewerbes ohne Bau, saison- und kalenderbereinigt, 2005=100. Quelle: Echtzeitdatenbank der Deutschen Bundesbank und eigene Berechnungen.

Anhand dieser qualitativen Unterscheidung werden nun die aufgedeckten Regime einzeln hinsichtlich ihres Veröffentlichungsverlaufs erläutert. Dabei wird insbesondere auf ihr Verhalten am aktuellen Rand eingegangen.

Die erste Rezession währt durchgängig vom Berichtsbeginn im Februar 1993 bis zum Juli 1993 mit zwischenzeitlich einmonatigem früher angezeigtem Ende. Im Veröffentlichungsverlauf stellen sich drei Tiefenveränderungen der zweiten Prioritätsstufe ein, so dass die zugewiesene relative Amplitude zwischen -1,5% und -3,5% schwankt, wobei die höchsten Werte am Ende des Veröffentlichungszeitraums erscheinen.

Die zweite Rezession zwischen November 1994 und November 1995 besitzt ebenfalls einen durchgängigen Rezessionsstart mit Ausnahme einer leicht verzögerten Wahrnehmung in der ersten Veröffentlichung vom Juni 1995. Die relative Rezessionstiefe (Prioritätsstufe 2) schwankt zwischen -2% und -4%, wobei die aktuelleren Werte zu den niedrigeren gehören. Im Veröffentlichungsverlauf weist die Rezession drei verschiedene Endzeitpunkte auf, wobei das Ende am aktuellen Rand um bis zu 8 Monate vorzeitig und zwischenzeitlich mehrmals um 3 Monate nachgelagert vermeldet wird (Prioritätsstufe 3).

In der anschließenden Expansion zwischen Dezember 1995 und Juni 1998 wird kurzfristig von September 1996 bis Januar 1997 für 6 Veröffentlichungen eine Rezession berichtet. Zwar ist der Sachverhalt der höchsten Priorität zuzuordnen. Man beachte aber, dass sich die Nichterkennung am aktuellen Rand im weiteren Veröffentlichungsverlauf nicht als Fehler erweist.

Die dritte Abschwungphase zwischen Juli 1998 und November 1998 ist neben leichten Abweichungen am aktuellen Rand geprägt von einem längeren Veröffentlichungszeitraum, in dem die Phase als länger und weniger tief erachtet wird. Insgesamt schwankt die relative Amplitude zwischen -2,5% und -4%.

Die vierte Abschwungphase zwischen Februar 2001 und November 2001 deutet sich am Veröffentlichungsrand im April 2001 genau wie bei den beiden darauf folgenden Rezessionen an. Sie verläuft mit Ausnahme einer Tiefenveränderung von etwa -2,5% am Beginn des Veröffentlichungszeitraums zu etwa -5,5% am Ende desselben sehr stabil.

Die fünfte Rezession zwischen August 2002 und September 2003 stellt insoweit eine Besonderheit dar, dass der Rezessionsbeginn mehrmals wechselt; insbesondere am Ende des betrachteten Veröffentlichungszeitraums wird der Beginn 6 Monate früher datiert als zwischenzeitlich (Prioritätsstufe 1). Die Tiefenschwankungen verlaufen zwischen -2% und -3,5%.

Die sechste Rezession zwischen Juli 2004 und November 2004 wird erst ab Veröffent-

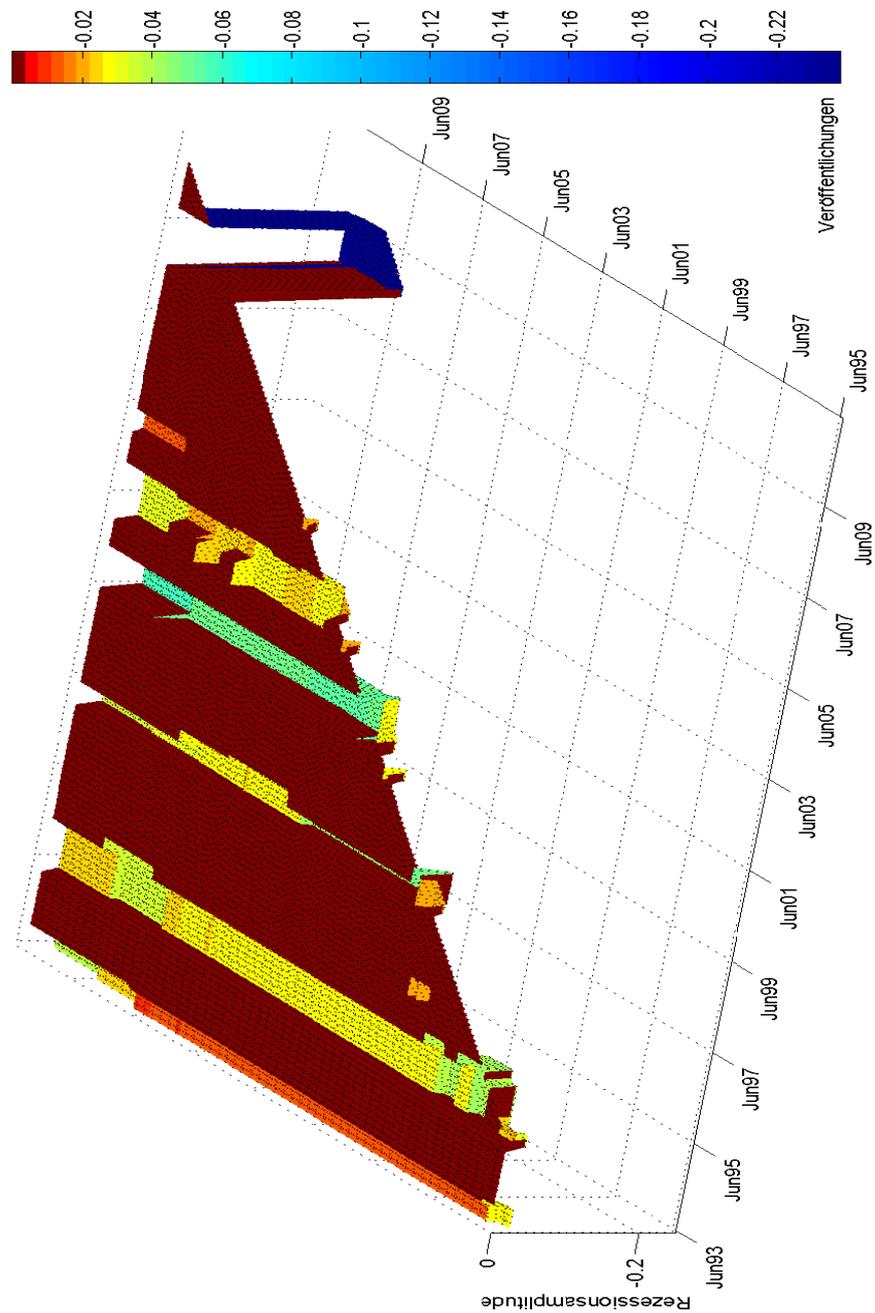


Abbildung 4.6: Wendepunktatierung und Rezessions-Tiefendarstellung – Seitenansicht. Expansionsphasen werden identisch 0 gesetzt und Rezessionsphasen mit ihrer relativen Tiefe angegeben.

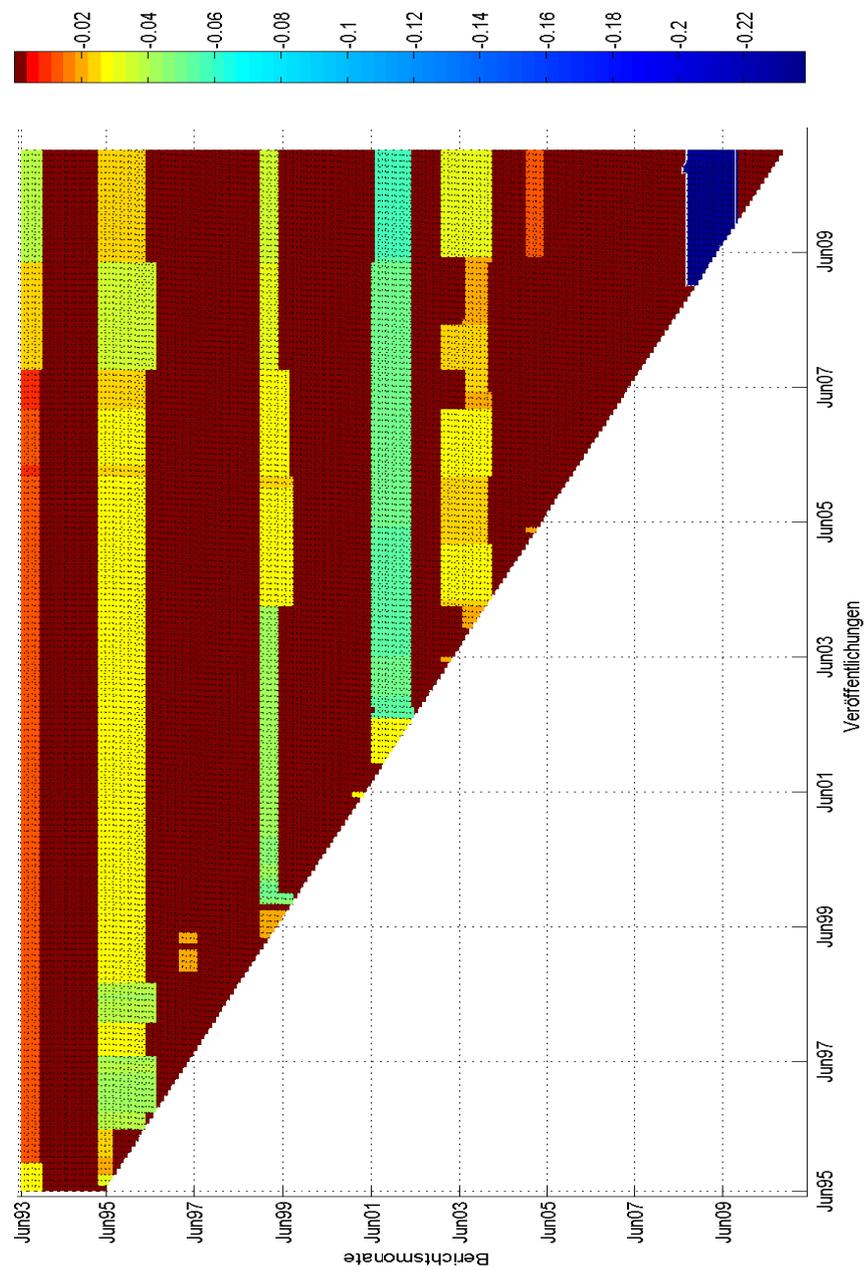


Abbildung 4.7: Wendepunkt-datierung und Rezessions-Tiefendarstellung – Aufsicht. Expansionsphasen werden identisch 0 gesetzt und Rezessionsphasen mit ihrer relativen Tiefe angegeben.

lichung April 2009 erkannt. Dann erscheint sie konstant mit einer relativ schwachen Amplitude von -1,5%. Die realwirtschaftlichen Auswirkungen der Finanzmarktkrise werden schließlich konstant zwischen März 2008 und April 2009 mit einer massiven relativen Tiefe von über -23% datiert. Bevor die Ergebnisse der ex post-Datierung im Echtzeitverlauf zusammengefasst werden, wird nun der Datierungsalgorithmus auf in- und ausländische Auftragseingänge angewendet (Abbildungen 4.8 bis 4.11). Betrachtet man die inländischen Auftragseingänge, so lassen sich zwar die zuvor in der Produktionsreihe detektierten sieben (und sogar die kurzfristig erscheinende achte) Abschwungphase wiederfinden; es stellen sich aber markante Unterschiede ein.

So werden die erste und zweite Abschwungphase ab Veröffentlichung März 2009 nicht mehr als solche erkannt. Zusätzliche Phasen der Auftragsrückgänge ergeben sich in der Expansion zwischen Dezember 2004 und Februar 2008 für insgesamt sieben Veröffentlichungen. Ferner schmelzen die vierte und fünfte Rezession für etwa 40 Veröffentlichungen zu einer Gesamtrezession zusammen. Die wichtigste Tiefenveränderung ergibt sich am aktuellen Rand der Finanzmarktkrise, in der der Rückgang der Auftragslage so rasant erfolgt, dass vor Ende der Veröffentlichung kein lokales Minimum erreicht wird. Bei solchen Fällen greift der Algorithmus auf die Tiefe von Abschwüngen vorhergegangener Phasen zurück. Da diese von geringerer Volatilität gekennzeichnet waren, führt dies zu einer Unterschätzung der relativen Amplitude bis zu dem Zeitpunkt, in dem ein lokales Minimum zur Berechnung der Tiefe zur Verfügung steht. Nichtsdestotrotz verläuft die Datierung an dieser Stelle korrekt. Insgesamt erscheint die Datierung im Vergleich zum Referenzzyklus stärker der Revisionsproblematik unterworfen zu sein.

Die Abweichungen sollten jedoch auch nicht überbewertet werden, da es sich bei den Aufträgen, wie erwähnt, um spezifische Zyklen handelt. Zwar decken sie unterschiedliche Wirtschaftssektoren ab, nicht jedoch alle Teile der Produktion aufgrund der Erwartungsbildung und Lagerhaltung der Unternehmen. Diesbezüglich ist zu erwähnen, dass die im parametrischen Indikatorvergleich verwendeten Modelle eben weitere Regressoren zur Vorhersage der gesamtwirtschaftlichen Aktivität und ihrer Wendepunkte verwenden. Im Mittelpunkt der Abbildungen 4.8 bis 4.11 soll daher vielmehr die Frage stehen, wie sich der Vorlauf der Auftragslage, sofern überhaupt feststellbar unter Echtzeitbedingungen verhält.

Der Vorlauf inländischer Auftragsrückgänge vor den Rückgängen der Produktion (Rezession) ist, sofern er sich einstellt, im Veröffentlichungsverlauf stabil. Da der Berichtsbeginn hinsichtlich des Vorlaufs keine Rolle spielt, beginnt die Betrachtung mit der zweiten

Rezession. Hier lässt sich jedoch nur ein Gleichlauf zu dem in der Produktionsreihe bestimmten Start feststellen.

Bei der dritten Rezessionsphase erhält man einen konstanten Vorlauf von fünf Monaten, für die vierte liegt dieser sogar bei neun Monaten. Dort, wo kein Zusammenschmelzen mit der fünften Rezession stattfindet, gilt erneut ein Gleichlauf. Die sechste Phase der Auftragsrückgänge verfügt gegenüber der korrespondierenden Rezession schließlich wieder über einen Vorlauf von konstant drei Monaten. Auch in der Finanzmarktkrise ist ein Vorlauf von drei Monaten erkennbar.

Betrachtet man die ausländischen Auftragseingänge, so korrespondieren die Phasen der Auftragsrückgänge durchgängig mit den sieben Rezessionen des Output. Eine Ausnahme stellt hier nur die Rezessionsphase am Berichtsbeginn dar, die sich nur partiell in den Auftragsrückgängen widerspiegelt. Auffällig ist, dass die Auftragsrückgänge, welche der fünften Rezessionsphase zuzuordnen sind, analog zum Output mehrmals einige Monate früher detektiert werden. Am aktuellen Veröffentlichungsrand ergibt sich ein ähnliches Bild wie bei den inländischen Aufträgen zu Berichtszeiten der Finanzmarktkrise. Die relative Amplitude wird auch hier unterschätzt, da der Algorithmus kein lokales Minimum erkennen kann und demzufolge auf Phasen geringerer Volatilität zurückgreift.

Betrachtet man nun den Verlauf, so lässt sich für die zweite Rezessionsphase nur ein Gleichlauf feststellen. Dagegen besitzt die Phase des dritten Auftragsrückgangs einen konstanten Vorlauf von drei Monaten. Für die vierte Rezession liegt dieser ebenfalls konstant in gleicher Höhe. Genauso für die fünfte, wenn man sich an den überwiegenden Veröffentlichungen des früheren Auftragsrückgangs orientiert. Für die sechste Rezession beträgt der Vorlauf zwei Monate. In der Finanzmarktkrise ist schließlich der größte Vorlauf ausländischer Auftragseingänge in Höhe von vier Monaten feststellbar.

Abschließend wird in Abbildung 4.12 das Ergebnis der ex post-Datierung vom Stand November 2010 festgehalten, also für die letzte der in dieser Untersuchung berücksichtigten Veröffentlichungen.

4.5 Zusammenfassung

Die Ergebnisse der ex post-Datierung unter Echtzeitbedingungen lassen sich wie folgt zusammenfassen:

1. Insgesamt verhalten sich die datierten Rezessionsphasen stabil über den Veröffent-

4 Datierung von konjunkturellen Wendepunkten in Echtzeit

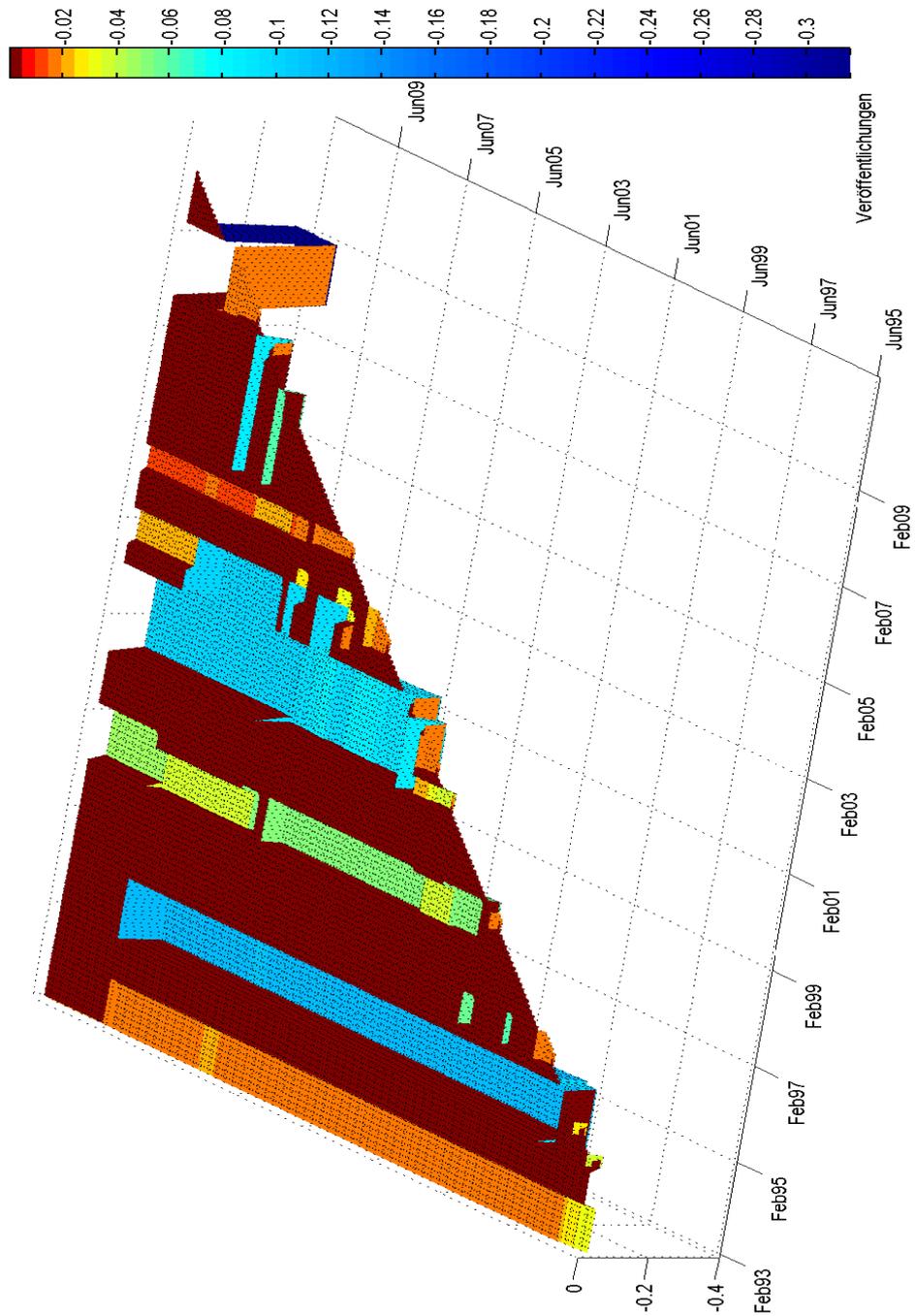


Abbildung 4.8: Wendepunktatierung im spezifischen Zyklus inländischer Auftragseingänge – Seitenansicht. Auftragsausweitungen werden identisch 0 gesetzt und Auftragsrückgänge mit ihrer relativen Tiefe angegeben.

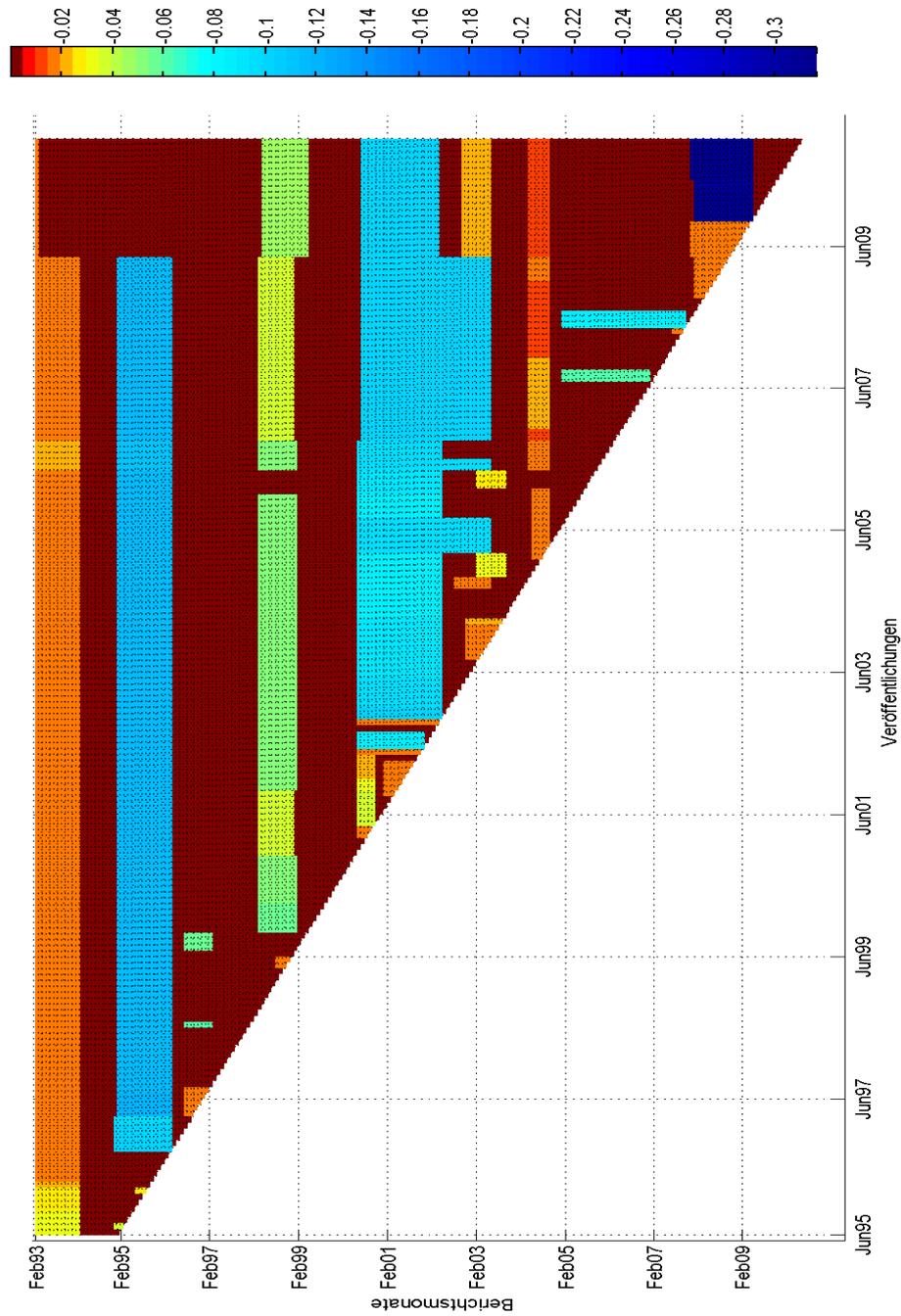


Abbildung 4.9: Wendepunkt datierung im spezifischen Zyklus inländischer Auftragseingänge – Aufsicht. Auftragsausweitungen werden identisch 0 gesetzt und Auftragsrückgänge mit ihrer relativen Tiefe angegeben.

4 Datierung von konjunkturellen Wendepunkten in Echtzeit

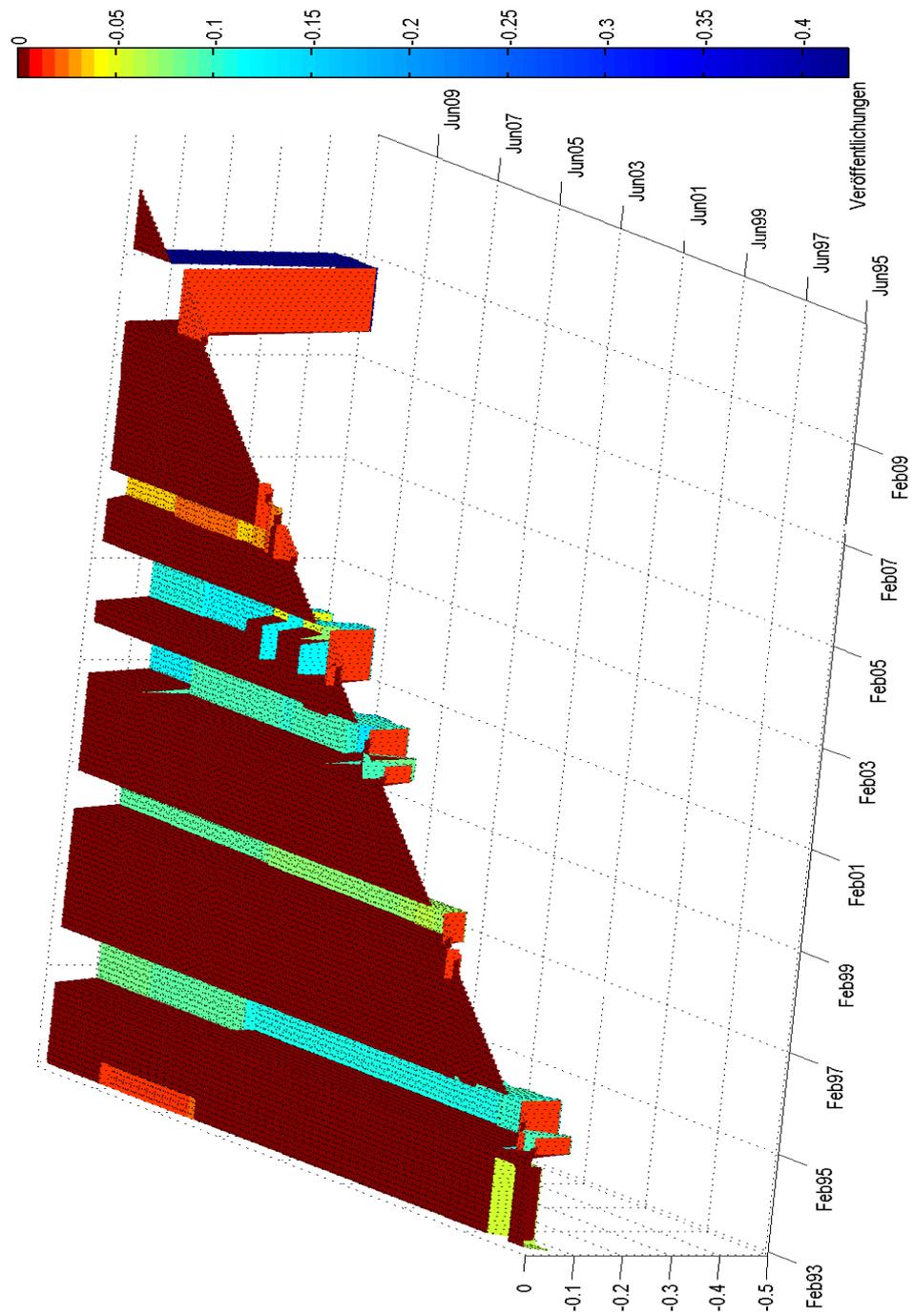


Abbildung 4.10: Wendepunktatierung im spezifischen Zyklus ausländischer Auftragseingänge – Seitenansicht. Auftragsausweitungen werden identisch 0 gesetzt und Auftragsrückgänge mit ihrer relativen Tiefe angegeben.

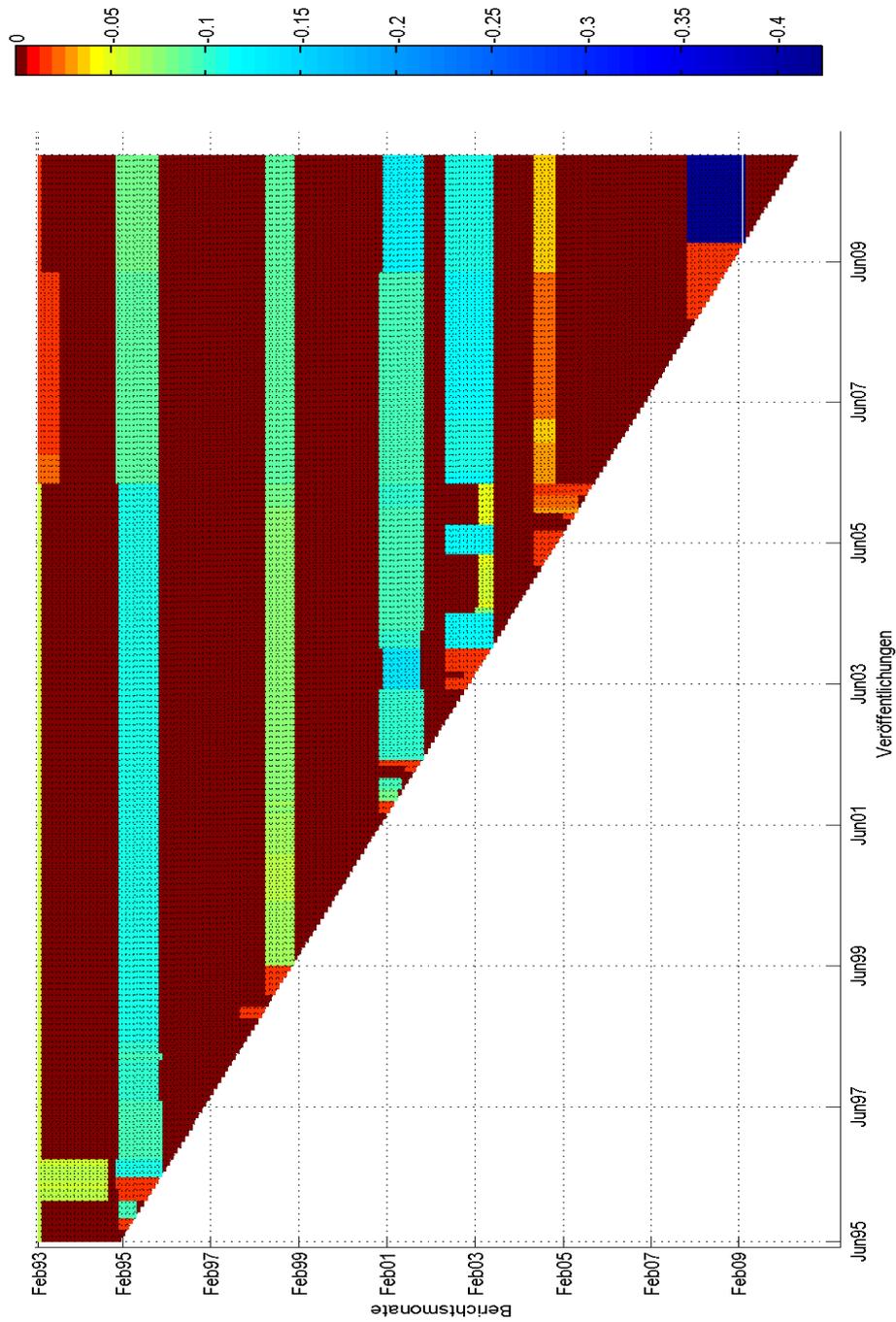


Abbildung 4.11: Wendepunkt datierung im spezifischen Zyklus ausländischer Auftragseingänge – Aufsicht. Auftragsausweitungen werden identisch 0 gesetzt und Auftragsrückgänge mit ihrer relativen Tiefe angegeben.

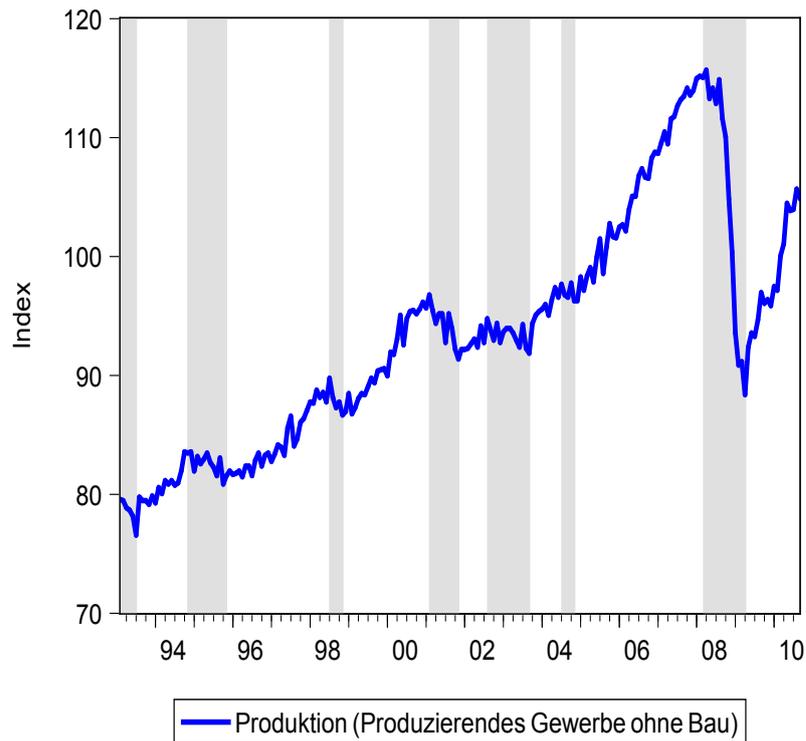


Abbildung 4.12: Ökonomische Gesamtaktivität. Darstellung der Produktion des Produzierenden Gewerbes ohne Bau (IP1AA020) in 2005er Preisen. Diese monatliche Zeitreihe ist kalender- und saisonbereinigt. Die grau hinterlegten Phasen stellen Rezessionen dar, detektiert anhand der Reihe IP1AA020. Stand November 2010. Quelle: Deutsche Bundesbank, Echtzeitdatenbank und eigene Berechnung.

lichungszeitraum. Lediglich zweimal führt die Datenrevision im Veröffentlichungsverlauf zu Echtzeitveränderungen der Prioritätsstufe 1 (kurzzeitige Meldung einer Rezession von September 1996 bis Januar 1997 sowie veränderter Beginn der Rezession August 2002 bis September 2003). Häufiger hingegen kommt es zu Veränderungen der zweiten Prioritätsstufe. So treten mehrere als mittelschwer einzustufende Veränderungen der relativen Amplitude auf, die zeigen, dass die Stärke der Rezession durchaus der Revisionsproblematik unterworfen ist, was wiederum Auswirkungen hat auf die Stärke der einzuleitenden wirtschaftspolitischen Maßnahmen. Als nachrangig sind dagegen die durch die Revisionen verursachten Veränderungen des Rezessionsendes einzustufen (Prioritätsstufe 3).

2. Phasen des Rückgangs der gesamtwirtschaftlichen Aktivität korrespondieren grundsätzlich in allen betrachteten Zeitreihen. Das bedeutet, dass sich die Phasen des Rückgangs im Output auch bei der Auftragslage wiederfinden. Das gilt jedoch nicht vollständig, weil zu einem geringen Teil auch zusätzliche Phasen steigender oder fallender wirtschaftlicher Aktivität erkannt werden. Der Vorlauf der Auftragseingänge bewegt sich in der Regel nur knapp über der Datenverfügbarkeitslücke, d.h. jenem Zeitraum zwischen dem letzten von Amts wegen veröffentlichten Datenstand und dem gegenwärtigen Zeitpunkt t . Ein Spitzenwert von neun Monaten Vorlauf wird in der Untersuchung nur für eine Rezessionsphase erreicht. Zudem ist der Vorlauf nicht durchgängig gegeben.
3. Insgesamt vermitteln die Ergebnisse der ex post-Datierung, dass im Hinblick auf die frühzeitige Wendepunkterkennung die Revisionsproblematik als nachrangig gegenüber der beschränkten Prognosefähigkeit vorlaufender Indikatoren zu beurteilen ist, wobei betont werden muss, dass ausschließlich die Auftragseingänge –stellvertretend für die vorlaufenden Indikatoren– mit ausreichender Echtzeitdatenhistorie zur Verfügung standen. Dabei umfasst beschränkte Prognosefähigkeit die Tatsache, dass falls ein eindeutiger Vorlauf erkennbar war, dieser auf Grund der Datenverfügbarkeitslücke eher für eine Gegenwarts- (Nowcasting) als für eine Zukunftsprognose (Forecasting) geeignet ist.

Grundsätzlich ist auch der Einsatz parametrischer Modelle, wie sie im weiteren Projektverlauf zur Früherkennung benutzt werden, zur ex post-Datierung denkbar. Die hier vorgestellten nicht-parametrischen Ansätze sind jedoch im Gegensatz zu den parametrischen

Ansätzen zur *Vorhersage* von Wendepunkten eher ungeeignet. So werden für die Festlegung lokaler Extrema am aktuellen Rand in der implementierten Version 5 Monate der Zukunft benötigt, so dass nach Berücksichtigung der Datenverfügbarkeitslücke von 2 Monaten mit dem vorliegenden Algorithmus 7 Monate später entschieden werden kann, ob zum betrachteten Zeitpunkt ein konjunktureller Wendepunkt vorlag oder nicht.

5 Wendepunktbestimmung mit einem linearen multivariaten Einzelgleichungsansatz

5.1 Einleitung

Die folgenden Kapitel 5 bis 8 befassen sich mit vier verschiedenen Modelltypen. Als erstes wird ein einfacher Ansatz vorgestellt, der auf linearen multivariaten Einzelgleichungen beruht. In diesem Ansatz wird die Wachstumsrate der Outputgröße¹ anhand eigener Verzögerungen und geeigneter (Früh)Indikatoren erklärt. Im Gegensatz zu Faktormodellen (vgl. z.B. Schumacher & Breitung 2008) und Pooling-Ansätzen (vgl. z.B. Drechsel & Scheufele 2010 und Drechsel & Scheufele 2011), die die relevante Information aus riesigen Datenmengen extrahieren, wird hier ganz bewusst ein kleines Modell spezifiziert. Es hat gegenüber komplexen Modellen, die schnell zu „black boxes“ werden, den großen Vorteil, dass man an der Schätzgleichung direkt ablesen kann, welche Indikatoren einen Erklärungsbeitrag leisten, und dass die Ergebnisse leicht nachvollziehbar sind. Darüber hinaus ist es interessant zu sehen, wie gut dieser einfache Ansatz im Vergleich zu den drei anderen weitaus komplexeren Modelltypen –Subset-VAR-, Markov-Switching- und dynamisches Probit-Modell– abschneidet.

Es ist wichtig, an dieser Stelle noch einmal explizit darauf hinzuweisen, dass die vorliegende Untersuchung zwei strengen Restriktionen unterliegt, die andere einschlägige Untersuchungen üblicherweise nicht beachten (müssen). Zum einen wird streng unter Echtzeitbedingungen gearbeitet, zum anderen gilt für alle Modelltypen die Maßgabe, dass sie

¹Die abhängige Variable ist die Produktion im Produzierenden Gewerbe ohne Bau, für die im Folgenden auch der Begriff Industrieproduktion verwendet wird. Die Zeitreihe für die Produktion wurde zunächst logarithmiert. Anschließend wurden die Veränderungen zum Vormonat berechnet. Bei kleinen Änderungen kann man mit diesem Vorgehen die Berechnung von Wachstumsraten approximieren.

so programmiert sind, dass eine (spätere) Anwendung in der Praxis möglich ist. Das bedeutet:

- die verwendeten Indikatoren müssen ohne größeren Aufwand zeitnah aktualisiert werden können,
- Modellspezifikation und Prognoseevaluierung müssen weitgehend automatisiert sein,
- die Modellspezifikation muss bei allen verwendeten Variablen die jeweiligen Datenverfügbarkeitslücken berücksichtigen und
- es dürfen nur solche Informationen verwendet werden, die zum jeweiligen Zeitpunkt t tatsächlich zur Verfügung standen.

Im Gegensatz zu anderen empirischen Studien, die ihre Untersuchung üblicherweise nur einmal für einen bestimmten Beobachtungszeitraum durchführen und ihre Prognosemodelle möglichst gut an diesen einen Datensatz anpassen, ist das Augenmerk in dieser Untersuchung darauf gerichtet, mit einem begrenzten Zeit- und Arbeitsaufwand und einer weitgehenden Automatisierung möglichst gute Prognosen unter Echtzeitbedingungen zu machen.

Die vier Modelltypen, die in dieser Studie hinsichtlich ihrer Eignung, konjunkturelle Wendepunkte vorherzusagen, untersucht werden, sind sehr unterschiedlich. Für eine erste Orientierung und bessere Einordnung der Modelle ist es daher sinnvoll, kurz auf wesentliche Unterschiede bzw. auf Gemeinsamkeiten der Modelle hinzuweisen: Einzelgleichungsansatz und Subset-VAR-Modell sind lineare Verfahren und die prognostizierte Zielgröße ist die Wachstumsrate bzw. das Niveau des gesamtwirtschaftlichen Outputs. Das heißt, dass man mit diesen Modellen die Rezessionswahrscheinlichkeit nicht direkt vorhersagen kann. Vielmehr muss man sie in einem zweiten Schritt über die Dichtefunktionen der Prognosen bestimmen. Markov-Switching- und dynamisches Probit-Modell hingegen sind nichtlineare Verfahren, mit denen man die Rezessionswahrscheinlichkeit direkt vorhersagen kann. Aufgrund der Tatsache, dass die linearen Modelle keine Prognosen für die Rezessionswahrscheinlichkeit liefern, sondern die Entwicklung der Outputgröße vorhersagen, braucht man eine explizite Regel für die Bestimmung einer Rezession. Diese Rezessionsdefinition muss praktikabel sein und zeitnah angewendet werden können. Folglich kann sie nicht identisch sein mit dem ausgefeilten Datierungsalgorithmus, der der ex post-Datierung zugrunde liegt (Abschnitt 4.2). Das Verfahren, das im Folgenden angewendet

wird, stützt sich auf die Dauer und die Stärke (Amplitude) eines Produktionsrückgangs innerhalb von fünf Monaten. Aus diesem Grund beträgt unser Prognosehorizont (Anzahl der Monate, für die Prognosen erstellt werden) fünf Monate.

Da in dieser Studie die Echtzeitbedingungen streng beachtet werden und im Einzelgleichungsansatz keine Prognosen für die erklärenden Variablen gemacht werden,² müssen bei der Modellspezifikation bei allen Variablen die jeweiligen Datenverfügbarkeitslücken (publication lag) beachtet werden. Das hat zwei wichtige Implikationen für die Vorgehensweise:

1. Bei der abhängigen Variablen, der Produktion im Produzierenden Gewerbe ohne Bau, beträgt der publication lag 2 Monate. D.h. zum Zeitpunkt t liegt der letzte amtliche Wert für die Outputgröße für den Zeitpunkt $t - 2$ vor. Daraus folgt, dass die erste Prognose für den Zeitpunkt $t - 1$ gemacht werden muss. Es besteht also die ungewöhnliche Situation, dass ein Wert, der vom Zeitpunkt t aus betrachtet bereits in der Vergangenheit liegt, vorausgesagt werden muss. Oder anders ausgedrückt: hier ist der backcast eine echte Prognose. Die zweite Prognose wird für den Zeitpunkt t gemacht. Das heißt, dass auch der nowcast eine echte Prognose ist. Die übrigen Prognosen werden für die Zeitpunkte $t + 1$, $t + 2$ und $t + 3$ gemacht.
2. Die Datenverfügbarkeitslücken sind bei den verschiedenen Zeitreihen unterschiedlich groß. Bei der Produktion und den Auftragseingängen sind es z.B. zwei Monate, bei den Stimmungsindikatoren der Europäischen Kommission ist es ein Monat und bei den ifo- und den ZEW-Indikatoren gibt es keinen publication lag.³ Das heißt, wenn man jeweils die größtmögliche Informationsmenge für die Erstellung der Prognose ausnutzen möchte, muss man für jeden der fünf Prognosehorizonte ($t - 1$, t , $t + 1$, $t + 2$, $t + 3$) eine eigene Gleichung schätzen.⁴ Die hier gewählte Bezeichnung „Einzelgleichungsansatz“ soll unterstreichen, dass jede der fünf Prognosegleichungen in Form einer Einzelgleichung geschätzt wird.

Kapitel 5 gliedert sich wie folgt: In Abschnitt 5.2 werden unsere Voruntersuchungen dargestellt, die das Ziel haben, aus der großen Anzahl der hier betrachteten Indikatoren diejeni-

²Das ist im vektorautoregressiven Ansatz anders. Dort werden alle aufgenommenen Variablen implizit prognostiziert.

³Einen Überblick über die Datenverfügbarkeitslücken aller verwendeten Variablen geben die Tabellen 2.1 und 2.2.

⁴In Tabelle 5.1 ist dargestellt, welche Verzögerungen einer Variable in den jeweiligen Prognosegleichungen verwendet werden können, wenn man die Datenverfügbarkeitslücke der Variable berücksichtigt.

gen auszuwählen, die besonders geeignet erscheinen, die Wachstumsrate der Outputgröße zu erklären. Für jeden Prognosehorizont werden deshalb Schätzungen für die Wachstumsrate der Produktion und jeweils einen Indikator gemacht. Auf Grundlage dieser bivariaten Voruntersuchungen werden vielversprechende Variablen für einen multivariaten General-to-Specific-Ansatz ausgewählt (Abschnitt 5.3). Da sich für jeden der fünf Prognosehorizonte mehrere mögliche Spezifikationen ergeben, werden diese Alternativen für verschiedenen lange Samples unter Echtzeitbedingungen hinsichtlich ihrer In-Sample-Prognosegüte verglichen. Am Ende dieses Auswahlprozesses steht für jeden der fünf Prognosehorizonte genau eine Prognosegleichung. Welche Indikatoren in den jeweiligen Gleichungen relevant sind, wird ebenfalls in Abschnitt 5.3 erläutert. Die fünf ausgewählten Gleichungen bilden die Grundlage für die Untersuchung der Prognosefähigkeit des Einzelgleichungsansatzes. In Abschnitt 5.4 wird die Schätzung der Rezessionswahrscheinlichkeit anhand von Dichte-Prognosen (density forecasts) erläutert und die Out-of-Sample-Evaluierung der Wendepunktprognosen unter Echtzeitbedingungen für den Einzelgleichungsansatz vorgenommen. In Abschnitt 5.5 werden Schlussfolgerungen hinsichtlich der Eignung des Einzelgleichungsansatzes für die Bestimmung konjunktureller Wendepunkte gezogen.

5.2 Voruntersuchungen für die Variablenauswahl

5.2.1 Bivariate Schätzungen

Wie in Abschnitt 2.1 dargestellt, wurden für die vorliegende Untersuchung 27 Konjunkturindikatoren ausgewählt, die sich den drei Bereichen realwirtschaftliche Indikatoren, Finanzmarktindikatoren sowie Umfragen (Surveys) zuordnen lassen. Diese Anzahl ist zu groß, als dass die Auswahl geeigneter Indikatoren über einen general-to-specific-Ansatz im Rahmen der Modellspezifikation erfolgen könnte. Vielmehr muss ein Weg gefunden werden, um bereits im Vorfeld potentiell geeignete von ungeeigneten Indikatoren zu unterscheiden.

Eine gängige Herangehensweise ist die Ermittlung der Vorlaufeigenschaften der Indikatoren anhand von Granger-Kausalitätstests (Granger 1969). Hier wird untersucht, ob die Ergänzung einer autoregressiven Schätzung der Variable x durch verzögerte Werte der Variable y (bis zu einer bestimmten, vordefinierten Verzögerung) die Schätzung verbessert. Ist dies der Fall, so ist y *Granger-kausal* für x . Gleichzeitig wird untersucht, ob eine autoregressive Schätzung der Variable y durch den Einsatz von verzögerten Werten der Variable

Tabelle 5.1: Mögliche Lagstrukturen in Abhängigkeit von der Datenverfügbarkeitslücke

Datenverfügbarkeits- lücke des Indikators	Schätzung bzw. Prognose zum Zeitpunkt t für				
	t-1 (backcast)	t (nowcast)	t+1	t+2	t+3
0 Monate	-1 bis k	0 bis k	1 bis k	2 bis k	3 bis k
1 Monat	0 bis k	1 bis k	2 bis k	3 bis k	4 bis k
2 Monate	1 bis k	2 bis k	3 bis k	4 bis k	5 bis k

Anmerkungen: Lag -1 steht hier für den Wert des Folgemonats. Positive Werte beziehen sich auf vorangegangene Monate. k: maximale Laglänge.

x verbessert wird. In diesem Fall wäre x Granger-kausal für y. Ist x Granger-kausal für y, aber y nicht für x, so nimmt man einen zeitlichen Vorlauf von x an. In unserem Fall ist der Granger-Kausalitätstest für eine Bestimmung der Vorlaufeigenschaften der Indikatoren jedoch ungeeignet, weil er bei den Variablen x und y die Verzögerungen ab Lag 1 betrachtet, unter Echtzeitbedingungen jedoch die ersten Verzögerungen für viele Variablen aufgrund der Datenverfügbarkeitslücke gar nicht zur Verfügung stehen. Wie sich an Tabelle 5.1 ablesen lässt, verschärft sich dieses Problem, je weiter man in die Zukunft prognostiziert.

Häufig werden auch Kreuzkorrelationen zwischen der abhängigen Variablen und Konjunkturindikatoren verwendet, um Variablen mit einem geeigneten Vorlauf zu identifizieren (vgl. z.B. de Bondt & Hahn 2010). Die Korrelationsanalyse kann sicher Anhaltspunkte liefern, weist aber auch Schwächen auf. So wird die Analyse durch hohe Autokorrelation erschwert und es ergeben sich Verzerrungen aufgrund von Überlagerungen von Schwingungen mit verschiedenen Phasenlängen (Fritsche & Stephan 2002, S.294).

Deswegen wird hier anders vorgegangen. Im Folgenden werden bivariate Gleichungen geschätzt. Die Wachstumsrate der Industrieproduktion wird durch eigene Verzögerungen und jeweils einen Indikator erklärt, wobei die Lagstruktur abhängig ist vom Prognosehorizont und von den Datenverfügbarkeitslücken der verwendeten Variablen (siehe Tabelle 5.1). Alle Schätzungen basieren auf stationären Zeitreihen. Das heißt, dass die Variablen gegebenenfalls transformiert wurden, um Stationarität zu gewährleisten. Die Transformationen der Daten können den Tabellen 5.2 bis 5.6 entnommen werden. Die Spezifikation der bivariaten Gleichungen erfolgt weitgehend automatisiert.⁵ Zunächst wird eine Startgleichung spezifiziert, in die beide Variablen mit einer maximalen Laglänge von 12 ein-

⁵Alle Schätz- und Prognoseprogramme laufen unter Eviews 7.

gehen. Dann werden alle Regressoren aus der Gleichung entfernt, deren p-Werte über der Signifikanzschwelle von 0,05 liegen. Aufgrund der Tatsache, dass unter Echtzeitbedingungen die ersten Verzögerungen der abhängigen Variable für Schätzungen ab dem Prognosehorizont t nicht zur Verfügung stehen, kann es zu Autokorrelation in den Residuen kommen. Deshalb werden die robusten Standardfehler von Newey & West (1987) verwendet, die diesem Umstand Rechnung tragen. Die bivariaten Schätzungen werden für zwei verschiedenen lange Beobachtungszeiträume gemacht; zum einen für das gesamte Sample 1993m2 bis 2010m9, zum anderen für das Sample 1993m2 bis 2006m11. Das kürzere Sample ist der Stützzeitraum für die erste 5-Schritt-Prognose im Rahmen der Out-of-Sample-Evaluierung der Wendepunktprognosen. Durch die Variation des Stützzeitraums soll überprüft werden, wie stabil der Zusammenhang zwischen Indikator und Referenzreihe im Zeitverlauf ist.

Das Ranking der Indikatoren erfolgt nun auf Grundlage ihres Erklärungsbeitrags (gemessen am \bar{R}^2) in den bivariaten Gleichungen. Die Tabellen 5.2 bis 5.6 zeigen die Ergebnisse für den gesamten Schätzzeitraum von 1993m2 bis 2010m9. Für das gesamte Sample ergeben die bivariaten Analysen ein recht eindeutiges Bild. Demnach erklärt der Composite Leading Indicator (CLI) der OECD bei allen Prognosehorizonten einen relativ hohen Anteil der monatlichen Veränderung der Industrieproduktion.⁶ Das Gleiche gilt für alle drei ifo-Indikatoren, während die ZEW-Indikatoren nicht überzeugen können. Auch die Surveys der Europäischen Kommission (industrial confidence, production expectations) schneiden gut ab, wengleich ihre Erklärungskraft mit zunehmendem Prognosehorizont abnimmt. Demgegenüber gewinnen die Zinsspreads mit zunehmendem Prognosehorizont an Erklärungskraft. Die Auftragseingänge aus dem In- und Ausland liefern nur bei den Horizonten $t-1$ und t einen Erklärungsbeitrag. Einige Variable liefern generell keinen oder fast keinen Erklärungsbeitrag. Dazu zählen die Kreditvergabe, der reale Außenbeitrag, die Zahl der offenen Stellen und der Konsumentenpreisindex. Auch der Early Bird Indikator der Commerzbank gehört zu den Indikatoren, die in der Vorauswahl schlecht abschneiden. Mehrere Studien bescheinigen dem Early Bird hingegen gute Prognoseeigenschaften. Dies gilt in der Regel für einen Prognosehorizont von rund einem Jahr, der hier nicht untersucht wird (Robinsonov & Wohlrabe 2008, Hinze 2003).

Betrachtet man nun den kürzeren Stützzeitraum (1993m2-2006m11), so zeigt sich ein gänzlich anderes Bild (vgl. Tabellen 5.12 bis 5.16 im Anhang 5.A). Abgesehen von den

⁶Gute Vorlaufeigenschaften des CLI für die monatliche Veränderung der deutschen Industrieproduktion werden durch andere Studien bestätigt. Vgl. z.B. Fichtner et al. (2009).

Schätzungen für $t-1$ sinkt der Erklärungsbeitrag der einzelnen Variablen in den bivariaten Gleichungen für alle anderen Prognosehorizonte dramatisch. Bemerkenswert ist, dass jetzt der Composite Leading Indicator der OECD überhaupt keine Erklärungskraft besitzt, während die Zinsspreads als Indikatoren bei den längeren Horizonten relativ an Bedeutung gewinnen. Die Tatsache, dass der Erklärungsgehalt aller Indikatoren im kürzeren Sample deutlich geringer ist als im langen Sample, weist darauf hin, dass eine grundlegende Prämisse für die Verwendung von Indikatoren unter Echtzeitbedingungen offenbar nicht erfüllt ist, nämlich die, dass es einen engen und im Zeitablauf stabilen Zusammenhang zwischen dem Indikator und der Referenzreihe gibt. Das relativ bessere Abschneiden der Indikatoren im langen Sample lässt sich möglicherweise durch die tiefe Rezession 2008/2009 erklären, die sich in den meisten Konjunkturindikatoren widerspiegelt. Dadurch wurde der Zusammenhang zwischen den Indikatoren und der Referenzreihe deutlich enger.

Die Vorauswahl der Variablen stützt sich im Wesentlichen auf die Ergebnisse für den langen Stützzeitraum, weil hier das Entscheidungskriterium des Erklärungsgehalts trennschärfer ist. Dagegen kann man einwenden, dass wir in diesem Fall Informationen verwenden, die wir unter Echtzeitbedingungen nicht gehabt hätten. Oder anders gesagt: wir wählen u.U. Indikatoren aus, die wir unter Echtzeitbedingungen nicht ausgewählt hätten.⁷ Dieser Einwand ist prinzipiell gerechtfertigt, wenngleich er in unserem Fall nur für eine Variable tatsächlich relevant ist, nämlich für den CLI. Im Großen und Ganzen sind die Überschneidungen bei der Variablenauswahl auf Basis des kurzen und des langen Samples groß. In den Fällen, in denen noch weitere Variablen für den kurzen Stützzeitraum in Frage kommen, die aber für den langen Stützzeitraum keine Rolle spielen, werden diese noch zusätzliche in die weitere Betrachtung einbezogen. Des Weiteren wird das Problem auch dadurch entschärft, dass wir für die Auswahl der Indikatoren noch ein zweites Kriterium heranziehen, nämlich ihre Korrelation untereinander.

⁷Benner & Meier (2004) weisen darauf hin, dass die Varianz der Prognosefehler in der Out-of-Sample-Prognose unterschätzt wird, wenn bei der Modellauswahl die gesamte Stichprobe zugrundegelegt wird. Streng genommen müsste die Modellauswahl zur Optimierung der Prognose mit jeder neuen Beobachtung neu erfolgen. Dies wäre im Rahmen einer komprimierten nachträglichen Echtzeitanalyse über mehrere Jahre mit einem unvermeidbaren Aufwand verbunden.

5 Wendepunktbestimmung mit einem linearen multivariaten Einzelgleichungsansatz

Tabelle 5.2: Bivariate Schätzungen für t-1 (backcast), Sample: 1993m2-2010m9

Indikator	Transformation der Indikatorreihe	Lags der abh. Variablen	Lags der Indika- torzeitreihe	\bar{R}^2
AE_INL_BUBA	log difference	1;2;5	1-4;9-10	0,37
CLI	log difference	1;3;5	3;6;9	0,37
INDCONF	level	1;3;5	1;3;6;8	0,34
PRODEX	level	1;3;5	0;7;8	0,31
IFO_ERW	log level	1;3;5	1;7	0,31
ECSSENT	log level	1;3;5	0;3	0,31
EURIBOR 3M	difference	1;3;5	-1;0;2;10	0,31
IFO_LAGE	log level	1;3;5	0;3	0,30
IFO_R1	log level	1;3;5	0;7	0,29
AE_AUSL_BUBA	log difference	1;2;4;5	1-6	0,29
YC1Y	level	1;3;5	3;5	0,27
C_SPREAD	difference	1-3;5	2;3;12	0,24
YC10Y	level	1-3;5	-1;1-4	0,24
YC2Y	level	1;3;5	3;5	0,23
BRENT	log difference	1;3	0-4	0,21
YC3Y	level	1;3;5	3;5	0,20
REX	log difference	1;2;3;5	1;3;12	0,20
YC5Y	level	1;2;3;5	3;5	0,19
ZEW_ERW	level	2;3;5	1-3	0,17
ZEW_SIT	level	1;3;5	-1;4	0,16
CDAX	log difference	2;3;5	0	0,16
CPI	log difference	2;3;5	1	0,16
VAC	log difference	2;3;5	0	0,15
RAW	log difference	2;3;5	9	0,15
KREDIT	difference	2;3;5	keines	0,12
COM_EB	difference	2;3;5	keines	0,12

Anmerkung: Lag -1 steht hier für den Wert des Folgemonats. Positive Werte beziehen sich auf vorangegangene Monate.

Tabelle 5.3: Bivariate Schätzungen für t (nowcast), Sample: 1993m2-2010m9

Indikator	Transformation der Indikatorreihe	Lags der abh. Variablen	Lags der Indika- torzeitreihe	\bar{R}^2
CLI	log difference	3;5	3;6;9	0,28
IFO_LAGE	log level	4;5	0;2;5;6;12	0,25
IFO_R1	log level	3;5	0;2;5;6	0,24
EURIBOR 3M	difference	3;5	0;2;3;10	0,24
AE_INL_BUBA	log difference	4;5	2;3;4;9;10	0,24
PRODEX	level	3;5	1;4	0,23
YC1Y	level	2;3;5	3;5	0,23
INDCONF	level	3;5	1;3	0,23
IFO_ERW	log level	3;5	1;2	0,22
YC2Y	level	2;3;5	3;4	0,21
YC10Y	level	3;5	0;1;2;3;4	0,21
ECSSENT	log level	3;5	1;3	0,20
YC3Y	level	2;3;5	3;4	0,20
AE_AUSL_BUBA	log difference	4;5	2;3;4;6	0,19
REX	log difference	2;3;5	1;3;12	0,19
C_SPREAD	difference	3;5	3;12	0,18
YC5Y	level	2;3;5	3;4	0,18
ZEW_ERW	level	2;3;5	1;2;3	0,17
CPI	log difference	2;3;5	1	0,16
ZEW_SIT	level	3;5	1;4	0,16
BRENT	log difference	2;3;5	0	0,15
RAW	log difference	2;3;5	9	0,15
CDAX	log difference	3;5	0	0,14
VAC	log difference	3;5	0	0,13
COM_EB	difference	2;3;5	keines	0,12
KREDIT	difference	2;3;5	keines	0,12

5 Wendepunktbestimmung mit einem linearen multivariaten Einzelgleichungsansatz

Tabelle 5.4: Bivariate Schätzungen für t+1, Sample: 1993m2-2010m9

Indikator	Transformation der Indikatorreihe	Lags der abh. Variablen	Lags der Indikatorzeitreihe	\bar{R}^2
CLI	log difference	3;5	3;6;9	0,28
YC1Y	level	3;5	3;5	0,22
IFO_ERW	log level	3;5	1;2	0,22
IFO_R1	log level	4;5	1;5;6	0,21
YC2Y	level	3;5	3;4	0,20
IFO_LAGE	log level	3;5	1;2;4-6	0,20
PRODEX	level	3;5	2;4	0,19
INDCONF	level	3;5	2;6;8	0,19
YC3Y	level	3;5	3;4	0,18
ECSSENT	log level	3;5	2;3	0,17
REX	log difference	3;5	1;3;12	0,17
C_SPREAD	difference	3	3;12	0,16
ZEW_SIT	level	3;5	1;4	0,16
YC10Y	level	3;5	2;3;4	0,15
ZEW_ERW	level	3;5	1;2;3	0,15
YC5Y	level	3	3;4	0,15
AE_AUSL_BUBA	log difference	4;5	3;4;5;6	0,14
CDAX	log difference	3;5	2	0,12
BRENT	log difference	3	3	0,10
EURIBOR 3M	difference	3;5	keines	0,09
CPI	log difference	3;5	keines	0,09
AE_INL_BUBA	log difference	3;5	keines	0,09
COM_EB	difference	3;5	keines	0,09
RAW	log difference	3;5	keines	0,09
KREDIT	difference	3;5	keines	0,09
VAC	log difference	3;5	keines	0,09

Tabelle 5.5: Bivariate Schätzungen für t+2, Sample: 1993m2-2010m9

Indikator	Transformation der Indikatorreihe	Lags der abh. Variablen	Lags der Indikatorzeitreihe	\bar{R}^2
CLI	log difference	5	4;6;9	0,23
IFO_ERW	log level	4;5	2;7	0,19
YC1Y	level	5	3;4	0,17
IFO_R1	log level	4;5	2;6	0,17
YC2Y	level	5	3;4	0,14
IFO_LAGE	log level	4;5	2;6	0,13
PRODEX	level	4;5	3;7	0,13
INDCONF	level	4;5	3;6	0,13
YC3Y	level	keines	3;4	0,12
ZEW_SIT	level	5	2;4	0,11
ECSSENT	log level	4;5	3;7	0,10
C_SPREAD	difference	keines	3;12	0,09
YC5Y	level	keines	3;4	0,09
EURIBOR 3M	difference	5	2	0,08
ZEW_ERW	level	5	3	0,08
YC10Y	level	keines	2;3;4	0,08
BRENT	log difference	keines	3	0,06
AE_INL_BUBA	log difference	4;5	4-6;9	0,06
REX	log difference	keines	3;12	0,05
AE_AUSL_BUBA	log difference	5	6	0,04
COM_EB	difference	keines	3	0,02
CDAX	log difference	5	keines	0,01
CPI	log difference	keines	keines	0,00
RAW	log difference	keines	keines	0,00
KREDIT	difference	keines	keines	0,00
VAC	log difference	keines	keines	0,00

Tabelle 5.6: Bivariate Schätzungen für t+3, Sample: 1993m2-2010m9

Indikator	Transformation der Indikatorreihe	Lags der abh. Variablen	Lags der Indika- torzeitreihe	\bar{R}^2
CLI	log difference	5	5;7;9	0,22
YC1Y	level	5	3;4	0,17
IFO_ERW	log level	5	3;7	0,16
YC2Y	level	5	3;4	0,14
IFO_R1	log level	5	3;6	0,14
YC3Y	level	keines	3;4	0,12
ZEW_SIT	level	5	3;4	0,11
IFO_LAGE	log level	5	5;6	0,10
YC5Y	level	keines	3;4	0,09
C_SPREAD	difference	keines	3;12	0,09
INDCONF	level	5	4;6	0,09
ZEW_ERW	level	5	3	0,08
ECSSENT	log level	5	4;7	0,08
BRENT	log difference	keines	3;8	0,07
PRODEX	level	5	4;6	0,07
REX	log difference	keines	3;12	0,05
YC10Y	level	keines	3	0,05
AE_AUSL_BUBA	log difference	5	6	0,04
CDAX	log difference	5	keines	0,01
AE_INL_BUBA	log difference	5	keines	0,01
EURIBOR 3M	difference	keines	keines	0,00
CPI	log difference	keines	keines	0,00
COM_EB	difference	keines	keines	0,00
RAW	log difference	keines	keines	0,00
KREDIT	difference	keines	keines	0,00
VAC	log difference	keines	keines	0,00

5.2.2 Analyse der Korrelation zwischen den Indikatoren

Zwischen einzelnen Variablen besteht z.T. eine sehr hohe zeitgleiche Korrelation (vgl. dazu Tabellen 5.7 bis 5.9). Dies gilt insbesondere für die Umfragedaten. Hoch korreliert sind beispielsweise die Reihen der Europäischen Kommission (ECSSENT, INDCONF und z.T. PRODEX) mit den Indikatoren des ifo-Instituts, sowie der ZEW-Indikator für die Lage mit dem ifo-Indikator für die Geschäftslage und einigen Reihen der Europäischen Kommission (ECSSENT, INDCONF). Dies erstaunt nicht, da z.B. sowohl die Befragungen der Europäischen Kommission als auch die des ifo-Instituts unter einem ähnlichen Kreis von Unternehmen durchgeführt werden. Darüber hinaus werden einige Reihen durch die Kombination anderer Reihen gebildet. Z.B. setzt sich das Ifo-Geschäftsklima aus den Geschäftserwartungen und der Geschäftslage zusammen. Die aggregierte Reihe sollte daher nicht zusammen mit einer Unterreihe in eine Gleichung verwendet werden. Ähnliches gilt für den Economic Sentiment Indicator (ECSSENT), den Industrial Confidence Indicator (INDCONF) und die Produktionserwartungen in der Industrie (PRODEX). Letztere sind eine Unterreihe des Industrial Confidence Indicators, welcher wiederum mit einem Gewicht von 40% in den Economic Sentiment Indicator eingeht. Der Composite Leading Indicator der OECD beinhaltet mehrere Reihen des Ifo-Geschäftsklimas, Auftragseingänge und Lagerbestände sowie einen Spread (Details siehe Anhang). Es überrascht, dass die Korrelation zwischen dem Composite Leading Indicator und den hier verwendeten Teilreihen nicht so hoch ist, wie in den oben erwähnten Fällen. Unter den Finanzmarktindikatoren sind nur die Yieldspreads der Staatsanleihen gegenüber dem 3-Monatszins untereinander sehr hoch korreliert. Dabei weist die Korrelation zwischen dem Yieldspread von 10jährigen Staatsanleihen und dem von 1jährigen Staatsanleihen mit 0,60 den geringsten Wert auf.

Um bei der multivariaten Schätzung Multikollinearitätsprobleme zu vermeiden, werden Indikatoren, deren Korrelationen betragsmäßig nah bei 1 liegen (als Grenze wurde 0,75 gewählt), nicht in der selben Gleichung verwendet. Diese Werte sind in den Tabellen 5.7 bis 5.9 fett gedruckt. Anhand von Tabelle 5.2 wird nun beispielhaft erläutert, wie die Auswahl der Indikatoren getroffen wurde. Im Grunde genommen wird die Liste von oben nach unten abgearbeitet, wobei die Korrelation zwischen einzelnen Indikatoren beachtet wird. Im vorliegenden Fall werden zunächst die Auftragseingänge aus dem Inland, der CLI und der Industrial Confidence Indicator (INDCONF) ausgewählt. Die Produktionserwartungen in der Industrie (PRODEX) werden aufgrund ihrer Korrelation mit INDCONF

Tabelle 5.7: Ausgewählte Korrelationen zwischen Frühindikatoren

	IFO_R1	IFO_LAGE	IFO_ERW	ECSSENT	INDCONF	PRODEX	ZEW_ERW	ZEW_SIT
IFO_R1								
IFO_LAGE	0,92							
IFO_ERW	0,82	0,53						
ECSSENT	0,86	0,79	0,71					
INDCONF	0,93	0,89	0,72	0,88				
PRODEX	0,85	0,64	0,91	0,81	0,87			
ZEW_ERW	0,12	-0,21	0,56	0,18	0,03	0,37		
ZEW_SIT	0,81	0,90	0,43	0,81	0,80	0,55	-0,25	
CLI	0,56	-0,33	0,38	-0,16	-0,25	0,09	0,60	0,38

Tabelle 5.8: Ausgewählte Korrelationen zwischen dem CLI und weiteren Unterreihen

	CLI
AE_INL_BUBA	0,07
AE_AUSL_BUBA	0,37
YC1Y	0,45
YC2Y	0,45
YC3Y	0,45
YC5Y	0,48
YC10Y	0,42

übersprungen. Als nächstes werden die ifo-Erwartungen ausgewählt, während die Variable ECSSENT aufgrund ihrer Korrelation mit INDCONF wieder übersprungen wird (etc.). Auf diese Weise wird eine Liste von Indikatoren erstellt. Ein Problem ist, dass diese Liste pfa-dabhängig ist. D.h. mit der Entscheidung, INDCONF auszuwählen, entscheiden wir uns gegen PRODEX und ECSSENT, aber z.B. auch gegen den ifo-Geschäftsklimaindex und die ifo-Geschäftslage. Um dieses Problem zu vermeiden, werden verschiedene Listen erstellt, indem z.B. PRODEX statt INDCONF ausgewählt wird, was dann dazu führt, dass im weiteren Auswahlprozess andere Indikatoren zum Zuge kommen. Welche der verschiedenen Alternativen letztendlich als „finale“ Prognosegleichung ausgewählt wird, ergibt sich aus einem Vergleich der alternativen Spezifikationen, der im nachfolgenden Abschnitt erläutert wird.

5.2 Voruntersuchungen für die Variablenauswahl

Tabelle 5.9: Ausgewählte Korrelationen zwischen Finanzmarktindikatoren

	REX	C_ SPREAD	CDAX	EURI BOR 3M	YC1Y	YC2Y	YC3Y	YC5Y
REX								
C_SPREAD	0,23							
CDAX	-0,25	-0,24						
EURIBOR 3M	-0,25	0,07	0,07					
YC1Y	-0,29	-0,24	0,22	0,50				
YC2Y	-0,28	-0,24	0,20	0,40	0,95			
YC3Y	-0,25	-0,23	0,20	0,31	0,88	0,98		
YC5Y	-0,19	-0,68	0,19	0,19	0,77	0,91	0,97	
YC10Y	-0,12	-0,21	0,17	0,06	0,60	0,79	0,88	0,97

5.3 Multivariate Prognosegleichungen

Im Folgenden werden für alle Prognosehorizonte lineare multivariate Gleichungen geschätzt. Die abhängige Variable ist die Wachstumsrate der Industrieproduktion, die durch eigene Verzögerungen und eine Auswahl an Indikatoren erklärt wird, wobei die Lagstruktur jeweils abhängig ist vom Prognosehorizont und von den Datenverfügbarkeitslücken der verwendeten Variablen (siehe Tabelle 5.1). Alle Gleichungen werden mit OLS geschätzt und verwenden ausschließlich stationäre Zeitreihen. Das heißt, dass die Variablen gegebenenfalls transformiert wurden, um Stationarität zu gewährleisten. Die Transformationen der Daten können den Tabellen 5.2 bis 5.6 entnommen werden. Aufgrund der Tatsache, dass unter Echtzeitbedingungen die ersten Verzögerungen der abhängigen Variable für Schätzungen ab dem Prognosehorizont t nicht zur Verfügung stehen, kann es zu Autokorrelation in den Residuen kommen. Deshalb werden die robusten Standardfehler von Newey & West (1987) verwendet, die diesem Umstand Rechnung tragen. Alle Schätzungen beruhen auf dem Stützzeitraum 1993m2 bis 2006m11.

Die Spezifikation der multivariaten Gleichungen erfolgt nach dem general-to-specific-Ansatz und läuft weitgehend automatisiert ab.⁸ Da wir für alle Variablen eine maximale Laglänge von 12 in der Startgleichung zulassen, können wir nicht beliebig viele Indikatoren in diese Gleichung aufnehmen, da wir sonst ein Problem mit den Freiheitsgraden bekommen würden. Deswegen wird die Anzahl der Indikatoren auf acht begrenzt. Die Eliminierung insignifikanter Regressoren erfolgt wie in Abschnitt 5.2.1 anhand einer festgelegten Signifikanzschwelle für den p-Wert. Allerdings ist die Prozedur jetzt dreistufig, wobei die Signifikanzschwelle schrittweise von 0,4 auf 0,05 abgesenkt wird. Die Automatisierung hat einerseits den Vorteil, dass viele Gleichungen spezifiziert und evaluiert werden können, auf der anderen Seite ergeben sich dadurch auch Restriktionen. So muss bei dem Algorithmus für die Eliminierung der insignifikanten Variablen sequenziell vorgegangen werden. D.h. zunächst wird der Algorithmus nur auf eine Variable angewendet, während alle übrigen Variablen nicht betrachtet werden. Erst wenn für die erste Variable der Auswahlprozess abgeschlossen ist, geht das Programm zur nächsten Variable über. Dieses Vorgehen birgt die Gefahr, dass das Ergebnis der Variablenauswahl von der Reihenfolge der Indikatoren im Eliminierungsprozess abhängt. Wir tragen diesem Problem Rechnung, indem wir die Variablen so anordnen, dass zuerst die vermeintlich schwächsten Indikatoren das Eliminierungsverfahren durchlaufen, während die vermeintlich besseren

⁸ Alle Schätz- und Prognoseprogramme laufen unter Eviews 7.

Indikatoren später geprüft werden. Des Weiteren wurde testweise die Reihenfolge der Indikatoren im Auswahlprozess verändert. Erfreulicherweise fiel die Variablenauswahl sehr ähnlich aus.

Wie in Abschnitt 5.2.1 beschrieben, ist die Indikatoren-Liste pfadabhängig und für jeden Prognosehorizont gibt es mehrere alternative Listen. Welche der verschiedenen Alternativen letztendlich als „finale“ Prognosegleichung ausgewählt wird, ergibt sich aus einem Vergleich der alternativen Spezifikationen. Für den Vergleich werden die folgenden Kriterien herangezogen:

- die Anpassungsgüte der Schätzgleichung in verschiedenen Stützzeiträumen⁹ (gemessen am \bar{R}^2),
- die Evaluierung ihrer In-Sample-Prognoseeigenschaften (gemessen am root mean squared error, RMSE, und am Theilschen Ungleichheitskoeffizienten, insbesondere an seinem Varianz- und seinem Biasanteil) und
- die Frage, ob die Residuen der Schätzgleichung normalverteilt sind.

Für jeden Prognosehorizont wird die Spezifikation ausgewählt, die nach den o.g. Maßen die beste Performance aufweist. Tabelle 5.10 gibt einen Überblick über die Regressoren, die für den jeweiligen Prognosehorizont von Bedeutung sind.

Für jeden Prognosehorizont wurde neben dem autoregressiven Term immer eine Kombination aus „harten“ und „weichen“ Indikatoren ausgewählt. Beim backcast (t-1) spielt der autoregressive Term eine gewichtige Rolle, weil er die Autokorrelation auffängt. Seine Bedeutung nimmt mit zunehmendem Prognosehorizont ab, was daran liegt, dass dann die zeitnahen Verzögerungen aufgrund der Datenverfügbarkeitslücke nicht zur Verfügung stehen. Tabelle 5.10 bestätigt im Wesentlichen die Ergebnisse der bivariaten Voruntersuchung. Gleichwohl gibt es aber auch weitergehende Einsichten. Bemerkenswert ist, dass die Auftragseingänge aus dem Ausland für alle fünf Prognosehorizonte eine Rolle spielen, auch wenn die bivariate Untersuchung ihre Eignung im kurzfristigen Bereich ergeben hatte, was im Falle der Auftragseingänge aus dem Inland hingegen bestätigt wurde. Bei den Indikatoren auf Basis von Umfragen schneiden die ifo-Indikatoren erwartungsgemäß gut ab. Überraschend ist hingegen, dass die ZEW-Erwartungen, die in der Voruntersuchung durchgefallen waren, in Kombination mit anderen Indikatoren durchaus hilfreich sind. Der

⁹Da unter Echtzeitbedingungen verfahren wurde, ist es sinnvoll, die Datenstände anzugeben: Januar 2007, Januar 2008, Januar 2009 und Januar 2010.

5 Wendepunktbestimmung mit einem linearen multivariaten Einzelgleichungsansatz

Tabelle 5.10: Ausgewählte Regressoren für die jeweiligen Prognosehorizonte

	t-1	t	t+1	t+2	t+3
Abhängige Variable	IPROD_OB	IPROD_OB	IPROD_OB	IPROD_OB	IPROD_OB
Realwirtschaftliche Variable	AE_AUSL_ BUBA	AE_AUSL_ BUBA	AE_AUSL_ BUBA	AE_AUSL_ BUBA	AE_AUSL_ BUBA
	AE_INL_ BUBA	AE_INL_ BUBA			
Umfrageindikatoren		IFO_R1	IFO_R1	IFO_R1	
	IFO_ERW IFO_LAGE				IFO_LAGE PRODEX
		ZEW_ERW	ZEW_ERW	ZEW_ERW	
Zusammengesetzte Indikatoren	CLI	CLI	CLI	CLI	CLI
Finanzmarktindikatoren	EURIBOR 3M	EURIBOR 3M		EURIBOR 3M	
	YC1Y		YC1Y	YC1Y	YC1Y
		YC10Y	YC10Y		YC10Y
	REX	REX	REX	REX	REX
		C_SPREAD	C_SPREAD	C_SPREAD	

CLI erfüllt die Erwartungen, die wir nach der Voruntersuchung an ihn hatten. Sehr bemerkenswert ist hingegen die Performance der Finanzmarktindikatoren. Zum einen bestätigt sich der Eindruck, den wir aus der Voruntersuchung gewonnen hatten, dass der Zinsspread in Bezug auf längere Prognosehorizonte an Bedeutung gewinnt. Zum anderen überrascht, dass sowohl der corporate spread als auch der Rentenindex (REX) in Kombination mit anderen Variablen Erklärungspotential besitzen. Das hätten wir angesichts der Voruntersuchung so nicht erwartet.

5.4 Schätzung der Rezessionswahrscheinlichkeit

Die Out-of-Sample-Prognose beginnt mit dem Datenstand Januar 2007. Mit jedem neuen Datenstand werden die fünf Prognosegleichungen neu geschätzt. Das heißt, dass sich die Auswahl der Variablen und ihre Lagstruktur mit jedem neuen Datenstand ändern (können).

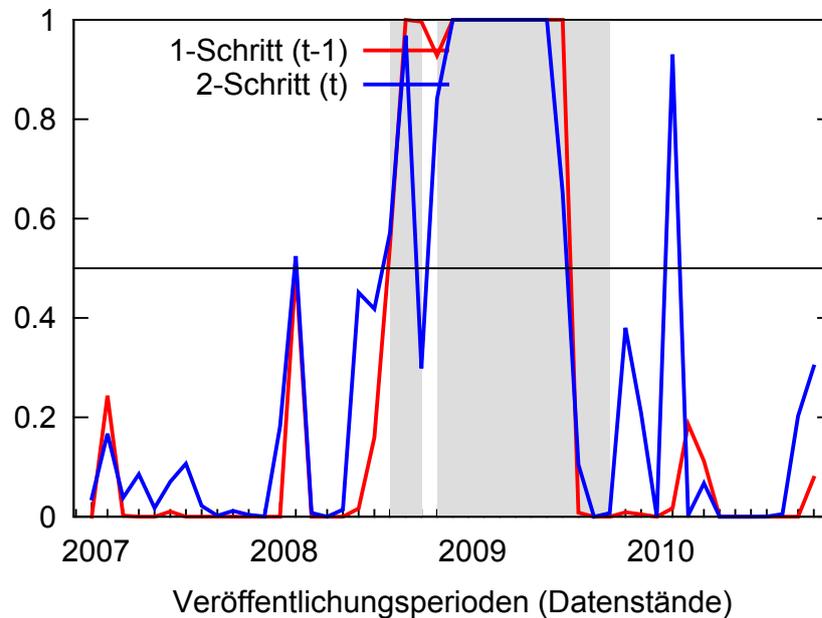
Die sich aus den fünf Einzelgleichungen ergebenden Punktprognosen für die Wachstumsrate der Industrieproduktion stellen lediglich ein Zwischenergebnis dar. Zur Ermittlung einer Rezessionswahrscheinlichkeit werden die Prognosen für die einzelnen Horizonte aufkumuliert. Für das Vorliegen einer Rezession wird nach dem Ansatz von Harding & Pagan (2002) wie in Abschnitt 4.2 beschrieben ein Schwellenwert von 1% des Ausgangsniveaus in einer Phase von fünf Monaten festgelegt. Für die Ableitung der Rezessionswahrscheinlichkeit zu jedem Veröffentlichungszeitpunkt wird folgendermaßen vorgegangen:

Aus der Normalverteilung der einzelnen Prognosefehler lässt sich die Normalverteilung der Summe der Prognosefehler über die Fünfmonatszeiträume ableiten. Für jeden Veröffentlichungszeitpunkt wird die Kovarianzmatrix der fünf Prognosefehler neu geschätzt. Es gilt: Die Varianz der Summe der Residuen ist gleich der Summe aller Elemente der Kovarianzmatrix. Mit dem Mittelwert und der Varianz ist die Normalverteilung eindeutig bestimmt und die Dichtefunktion kann abgeleitet werden. So kann die Wahrscheinlichkeit einer Realisation jenseits des Schwellenwertes ermittelt werden.

Die Rezessionswahrscheinlichkeit wird für jeden Veröffentlichungszeitpunkt anhand der kombinierten Prognosen für die 5 Horizonte ermittelt. Es wurde festgelegt, dass Wahrscheinlichkeiten oberhalb von 0,5 eine Rezession signalisieren. Dieser Wert ist gewissermaßen willkürlich, weil es kein objektives Kriterium für die „richtige“ Wahrscheinlichkeitsgrenze gibt.

Als Benchmark für die Rezessionsprognose wird eine Datierung nach dem Verfahren

Abbildung 5.1: Rezessionswahrscheinlichkeiten für die 1-Schritt- und die 2-Schritt-Prognose

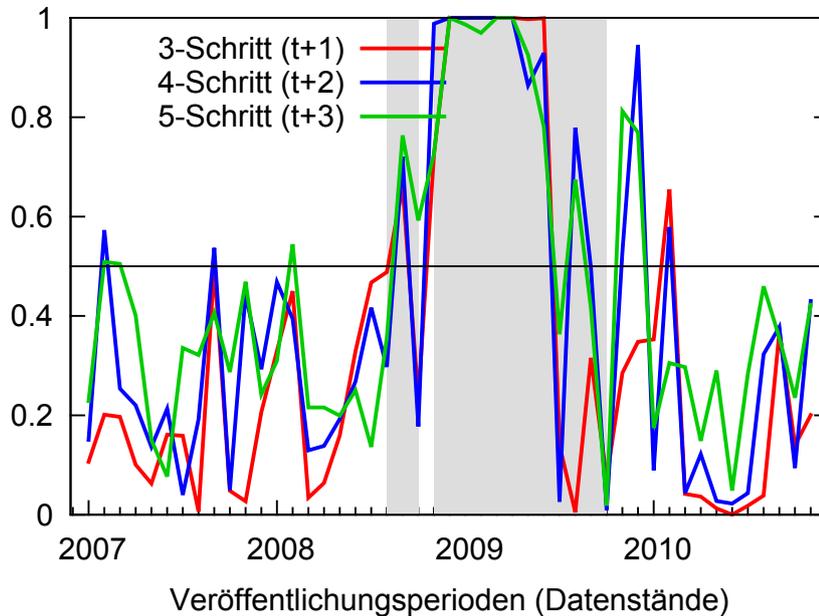


von Harding & Pagan (2002) auf der Basis der zu jedem Veröffentlichungszeitpunkt verfügbaren Daten verwendet. Aufgrund der Datenverfügbarkeitslücke von 2 Monaten zeigt der Benchmark daher an, ob im Zeitraum von 7 bis 2 Monaten vor dem Veröffentlichungszeitpunkt eine Rezession bestanden hat. Eine optimale 5-Schritt-Prognose würde die Rezession daher schon 5 Monate vor der amtlichen Statistik anzeigen. Vorläufe von 1 bis 4 Monaten wären jedoch auch ein Gewinn gegenüber der amtlichen Statistik.

Ergänzend zur vollen 5-Schritt-Prognose wurden auch 5-Monatszeiträume betrachtet, für die ein Teil der Daten veröffentlicht ist und ein Teil prognostiziert wird. Im ersten Fall („1-Schritt“) wurden vier ex post-Werte mit einer Einschritt-Prognose für den Monat nach dem letzten Monatsbericht kombiniert. Im zweiten Fall wurden drei ex post-Werte mit einer Zweischritt-Prognose („2-Schritt“, auf der Basis der Gleichungen für $t-1$ und t) kombiniert. Im dritten Fall („3-Schritt“) wurden zwei ex post-Werte mit einer Dreischritt-Prognose (auf der Basis der Gleichungen für $t-1$, t und $t+1$) miteinander verbunden und im vierten Fall („4-Schritt“) wurde ein ex post-Wert mit einer Vierschritt-Prognose (auf der Basis der Gleichungen für $t-1$, t , $t+1$ und $t+2$) kombiniert.

Abbildungen 5.1 und 5.3 zeigen die jeweils angezeigten Rezessionswahrscheinlichkeiten gemessen am grau unterlegten Benchmark. Gute Prognosen zeigen schon vorher eine

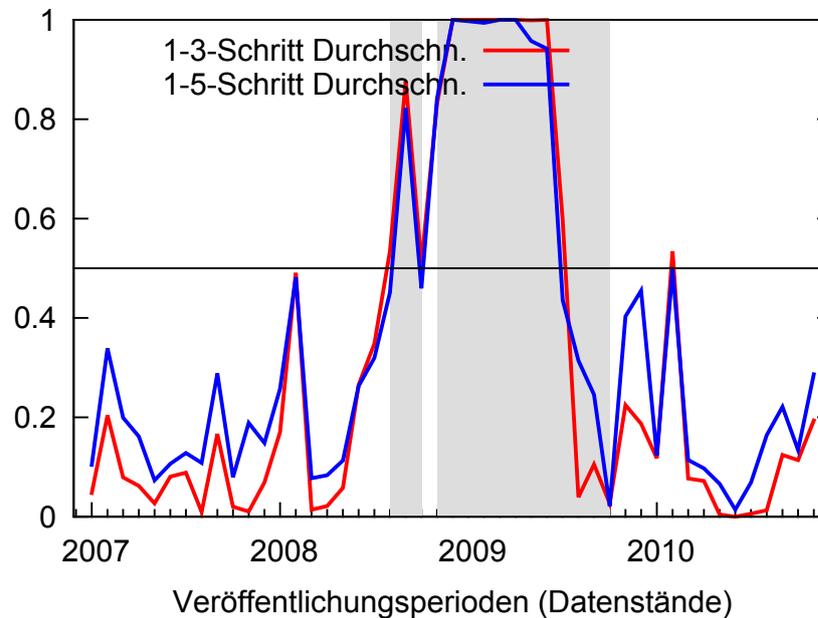
Abbildung 5.2: Rezessionswahrscheinlichkeiten für die 3-Schritt-, die 4-Schritt- und die 5-Schritt-Prognose



Rezession an, d.h. die Rezessionswahrscheinlichkeit sollte schon vor (also links von der) Benchmark über 50% steigen. Bei der 5-Schritt-Prognose müsste dies genau 5 Monate früher der Fall sein, bei der 4-Schritt-Prognose 4 Monate und so weiter. Es zeigt sich, dass die Prognosen mit dem langen Horizont insbesondere die Rezession 2008 eher zu spät angezeigt haben. Interessanterweise scheint ein Vorlauf bei einem kürzeren Prognosehorizont eher gegeben zu sein. Dieses paradoxe Ergebnis erklärt sich vielleicht dadurch, dass die Schätzgleichung für den kurzen Prognosehorizont unabhängig vom Stützzeitraum tendenziell besser ist. Für alle 5 Varianten ergibt sich das Problem, dass Rezessionen fälschlicherweise angezeigt werden. Diesem Problem ist am besten durch Bildung von Durchschnitts zu begegnen (vgl. Abbildung 5.3).

Der Durchschnitt aus den 1-3-Schrittprognosen weist einen – wenn auch nur geringen – Vorlauf aus. Anfang 2010 wird zwar immer noch fälschlicherweise eine Rezessionswahrscheinlichkeit von über 50% angezeigt. Generell sind die Wahrscheinlichkeiten bei den Durchschnitts weniger volatil, es ergibt sich ein klareres Bild.

Abbildung 5.3: Rezessionswahrscheinlichkeiten für Durchschnitte aus den Einzelprognosen



5.5 Zusammenfassung

Die Prognose der Rezessionswahrscheinlichkeit auf der Basis von fünf kombinierten Einzelgleichungen und einem Density-Forecast scheint noch verbesserungswürdig. Dies gilt insbesondere hinsichtlich der Vorlaufeigenschaften bei den längeren Prognosehorizonten. Hier wird die Rezession zu spät erkannt, so dass gegenüber dem Datenstand der amtlichen Statistik kein Informationsvorteil besteht. Besser schneiden paradoxerweise die Rezessionsprognosen für einen kürzeren Prognosezeitraum ab. Hier ergibt sich bei der Rezessionserkennung ein geringfügiger Vorlauf gegenüber der amtlichen Statistik. Problematisch bleiben Fehlanzeigen von Rezessionen. Die Bildung von Durchschnitten über mehrere Horizonte führt generell zu weniger volatilen Ergebnissen, kann die Probleme aber nicht beseitigen.

Anhang 5.A Weiteres Material

5.A.1 Zusammensetzung des Composite Leading Indicator der OECD für Deutschland

Tabelle 5.11 beschreibt die Variablen, die zur Bildung dieses Indikators verwendet werden.

Tabelle 5.11: Zusammensetzung des Composite Leading Indicator der OECD für Deutschland

Reihe	Quelle
Geschäftsklimaindikator (<i>Normallage</i> = 100)	Ifo Institut für Wirtschaftsforschung
Auftragseingang/Nachfrage: Tendenz (Verarbeitendes Gewerbe), (% Saldo)	Ifo Institut für Wirtschaftsforschung
Exportaufträge: Niveau (Verarbeitendes Gewerbe), (% Saldo)	Ifo Institut für Wirtschaftsforschung
Auftragseingang insgesamt (Verarbeitendes Gewerbe, 2000 = 100)	Statistisches Bundesamt
Bestand an Fertigerzeugnissen: Niveau (Verarbeitendes Gewerbe) (% Saldo, invertiert)	Ifo Institut für Wirtschaftsforschung
Zinsspread (% per annum)	ECB

Anmerkung: Quelle http://www.oecd.org/document/41/0,3746,en_2649_34349_1891113_1_1_1_1,00.html, eigene Übersetzung. Die OECD-Quelle spezifiziert den Zinsspread nicht näher.

5.A.2 Variablenauswahl

In den folgenden Tabellen 5.12 bis 5.16 werden einige Ergebnisse für den kurzen Stützzeitraum berichtet; siehe auch den Haupttext.

5 Wendepunktbestimmung mit einem linearen multivariaten Einzelgleichungsansatz

Tabelle 5.12: Bivariate Schätzungen für t-1 (backcast), Sample: 1993m2-2006m11

Indikator	Transformation der Indikatorreihe	Lags der abh. Variablen	Lags der Indika- torzeitreihe	\bar{R}^2
IFO_ERW	log level	1;2	1;3;6;7;10;11	0,33
EURIBOR 3M	difference	1;2	0;2;11	0,32
ECSSENT	log level	1;2	0;3;6;7	0,30
IFO_LAGE	log level	1;2;11;12	0;4;11;12	0,30
AE_AUSL_BUBA	log difference	1;2;3	1-6	0,30
IFO_R1	log level	1;2	1;2;11	0,29
YC10Y	level	1;2	0-4	0,29
CDAX	log difference	1;2;11;12	-1;0;4;5;7	0,29
CLI	log difference	1;2	2	0,28
INDCONF	level	1;2	0;7;8	0,28
REX	log difference	1;2	3;9	0,28
PRODEX	level	1;2	0;7	0,27
YC1Y	level	1;2	3;4	0,27
YC2Y	level	1;2	3;4	0,27
ZEW_ERW	level	1;2	-1;4	0,27
YC3Y	level	1;2	3;4	0,26
ZEW_SIT	level	1;2	-1;5	0,26
YC5Y	level	1;2	3;4	0,24
VAC	log difference	1;2	0	0,22
RAW	log difference	1;2	3;9	0,21
CPI	log difference	1;2	7	0,20
C_SPREAD	difference	1;2	3;12	0,19
AE_INL_BUBA	log difference	1;2	keines	0,18
BRENT	log difference	1;2	keines	0,18
KREDIT	difference	1;2	keines	0,18
COM_EB	difference	1;2	keines	0,18

Anmerkung: Lag -1 steht hier für den Wert des Folgemonats. Positive Werte beziehen sich auf vorangegangene Monate

Tabelle 5.13: Bivariate Schätzungen für t (nowcast), Sample: 1993m2-2006m11

Indikator	Transformation der Indikatorreihe	Lags der abh. Variablen	Lags der Indika- torzeitreihe	\bar{R}^2
YC10Y	level	3	0-4	0,17
YC5Y	level	keines	2-4	0,11
YC1Y	level	keines	3;4	0,09
YC2Y	level	keines	3;4	0,09
YC3Y	level	keines	3;4	0,09
REX	log difference	keines	3;9	0,09
IFO_LAGE	log level	keines	0;2;5;6	0,07
IFO_R1	log level	keines	0;2;5;6	0,07
AE_AUSL_BUBA	log difference	keines	3;6	0,07
EURIBOR 3M	difference	keines	2;3	0,06
AE_INL_BUBA	log difference	3;11	6;9	0,06
CDAX	log difference	3	0	0,05
CPI	log difference	3	7	0,04
CLI	log difference	keines	2	0,03
C_SPREAD	difference	keines	6	0,03
RAW	log difference	keines	3	0,03
COM_EB	difference	keines	10	0,02
VAC	log difference	keines	0	0,01
PRODEX	level	keines	keines	0,00
INDCONF	level	keines	keines	0,00
IFO_ERW	log level	keines	keines	0,00
ECSSENT	log level	keines	keines	0,00
ZEW_ERW	level	keines	keines	0,00
ZEW_SIT	level	keines	keines	0,00
BRENT	log difference	keines	keines	0,00
KREDIT	difference	keines	keines	0,00

5 Wendepunktbestimmung mit einem linearen multivariaten Einzelgleichungsansatz

Tabelle 5.14: Bivariate Schätzungen für t+1, Sample: 1993m2-2006m11

Indikator	Transformation der Indikator- reihe	Lags der abh. Variablen	Lags der Indika- torzeitreihe	\bar{R}^2
YC10Y	level	3	2-4	0,13
YC5Y	level	keines	2-4	0,11
YC1Y	level	keines	3;4	0,09
YC2Y	level	keines	3;4	0,09
YC3Y	level	keines	3;4	0,09
REX	log difference	keines	3;9	0,09
IFO_R1	log level	11	1;2;3;5;6	0,07
IFO_LAGE	log level	11	5;6;10;12	0,07
AE_AUSL_BUBA	log difference	keines	3;6	0,07
EURIBOR 3M	difference	keines	2;3	0,06
CPI	log difference	3	7	0,04
C_SPREAD	difference	keines	6	0,03
RAW	log difference	keines	3	0,03
ECSSENT	log level	keines	2;3	0,02
COM_EB	difference	keines	10	0,02
CLI	log difference	keines	keines	0,00
IFO_ERW	log level	keines	keines	0,00
PRODEX	level	keines	keines	0,00
INDCONF	level	keines	keines	0,00
ZEW_SIT	level	keines	keines	0,00
ZEW_ERW	level	keines	keines	0,00
CDAX	log difference	keines	keines	0,00
BRENT	log difference	keines	keines	0,00
AE_INL_BUBA	log difference	keines	keines	0,00
KREDIT	difference	keines	keines	0,00
VAC	log difference	keines	keines	0,00

Tabelle 5.15: Bivariate Schätzungen für t+2, Sample: 1993m2-2006m11

Indikator	Transformation der Indikator- reihe	Lags der abh. Variablen	Lags der Indika- torzeitreihe	\bar{R}^2
YC10Y	level	keines	2-4	0,12
YC5Y	level	keines	2-4	0,11
YC1Y	level	keines	3;4	0,09
YC2Y	level	keines	3;4	0,09
YC3Y	level	keines	3;4	0,09
REX	log difference	keines	3;9	0,09
IFO_LAGE	log level	11	5;6;10;12	0,07
PRODEX	level	keines	7;8;10-12	0,07
EURIBOR 3M	difference	keines	2	0,04
AE_AUSL_BUBA	log difference	keines	6	0,04
ECSSENT	log level	keines	6	0,03
IFO_R1	log level	keines	5;7	0,02
COM_EB	difference	keines	10	0,02
CPI	log difference	keines	7	0,02
CLI	log difference	keines	keines	0,00
IFO_ERW	log level	keines	keines	0,00
INDCONF	level	keines	keines	0,00
ZEW_SIT	level	keines	keines	0,00
C_SPREAD	difference	keines	keines	0,00
ZEW_ERW	level	keines	keines	0,00
BRENT	log difference	keines	keines	0,00
AE_INL_BUBA	log difference	keines	keines	0,00
CDAX	log difference	keines	keines	0,00
RAW	log difference	keines	keines	0,00
KREDIT	difference	keines	keines	0,00
VAC	log difference	keines	keines	0,00

5 Wendepunktbestimmung mit einem linearen multivariaten Einzelgleichungsansatz

Tabelle 5.16: Bivariate Schätzungen für t+3, Sample: 1993m2-2006m11

Indikator	Transformation der Indikator- reihe	Lags der abh. Variablen	Lags der Indika- torzeitreihe	\bar{R}^2
YC1Y	level	keines	3;4	0,09
YC2Y	level	keines	3;4	0,09
YC3Y	level	keines	3;4	0,09
REX	log difference	keines	3;9	0,09
IFO_LAGE	log level	11	5;6;10;12	0,07
YC5Y	level	keines	3;4	0,07
PRODEX	level	keines	7;8;10-12	0,07
YC10Y	level	keines	3;4	0,06
AE_AUSL_BUBA	log difference	keines	6	0,04
C_SPREAD	difference	keines	6	0,03
INDCONF	level	keines	keines	0,03
IFO_R1	log level	keines	5;7	0,02
CPI	log difference	keines	7	0,02
COM_EB	difference	keines	10	0,02
CLI	log difference	keines	keines	0,00
IFO_ERW	log level	keines	keines	0,00
ZEW_SIT	level	keines	keines	0,00
ZEW_ERW	level	keines	keines	0,00
ECSSENT	log level	keines	keines	0,00
BRENT	log difference	keines	keines	0,00
CDAX	log difference	keines	keines	0,00
AE_INL_BUBA	log difference	keines	keines	0,00
EURIBOR 3M	difference	keines	keines	0,00
RAW	log difference	keines	keines	0,00
KREDIT	difference	keines	keines	0,00
VAC	log difference	keines	keines	0,00

6 Wendepunktbestimmung mit Subset-VAR-Modellen

6.1 Einleitung

Dieses Kapitel wendet einen konsequenten Systemansatz auf das Problem der Auf- und Abschwungprognosen an, indem relativ große vektorautoregressive Modelle (VARs) zur Prognose aller verwendeten Variablen spezifiziert werden. Wie im Kapitel 5 dieser Studie erhalten wir durch diese linearen Modelle Prognosen der stetigen Output-Variable. Ähnlich wie dort berechnen wir dann mit Hilfe der geschätzten Prognose-Dichtefunktionen die Wahrscheinlichkeit davon, dass sich ein Wert unterhalb der definierten Rezessionsschwelle realisieren wird.¹ Der hauptsächliche Unterschied zum Verfahren des Kapitels 5 liegt darin, dass wir die Mehrschritt-Prognosen iterativ aus den Ein-Schritt-Prognosen generieren, indem prognostizierte Werte rekursiv eingesetzt werden, wie bei VAR-Prognosen üblich. Außerdem müssen wir wegen der großen Zahl zu spezifizierender Gleichungen die Variablenselektion mit geeigneten Algorithmen automatisieren.

Um die Echtzeitproblematik in dieser Modellklasse erfassen zu können, wählen wir in diesem Kapitel eine etwas andere Datendarstellung, in der sich die gewählte Zeitachse für die Datenvektoren komplett auf Perioden der Datenveröffentlichungen (Rechenstände) bezieht (eine detailliertere Darstellung erfolgt weiter unten). Die implizit betroffenen Berichtsmonate müssen aus den Ergebnissen abgeleitet werden, indem am Schluss die bekannten (aber teils variablenspezifischen) Publikationslags berücksichtigt werden.

Die in diesem Kapitel verwendete Modellklasse hat folgende Vorteile: Erstens sind wegen der Linearität der Schätzer analytische Lösungen für jeden Schätzschritt des Algorithmus gegeben, wodurch die Existenz der Wendepunktprognose bei jeder Datenkonstellati-

¹Wir berücksichtigen nur die Residuenvarianz, nicht zusätzlich die Schätzunsicherheit der Parameter. Dies sollte allerdings hauptsächlich in den Randbereichen der Verteilung bedeutsam sein, wobei diese Bereiche für unsere Fragestellung eher nachrangig sind.

on garantiert ist. Bei nichtlinearen Schätzern wie in den Kapiteln 7 und 8 besteht dagegen prinzipiell immer die Möglichkeit, dass die Maximierung der Likelihood-Funktion fehlschlägt. In diesem Kapitel besteht der einzige nichtlineare Schritt darin, eine Art p-Wert einer Normalverteilung zu finden, was offensichtlich keine numerischen Probleme bereiten kann. Zweitens besteht wie im Kapitel 5 die Hoffnung, dass die direkte Prognose der metrisch skalierten Variable Industrieproduktion besser in der Lage ist, die vorhandene Dateninformation auszunutzen, als dies bei einer vorherigen Reduktion auf eine binäre Regime-Variable wie im Probit-Modell der Fall ist. Ein frühes Beispiel dafür, dass linearstetige Prognosemodelle auf das Problem der Wendepunktbestimmung angewendet werden können, lieferten Stock & Watson (1993). Schließlich ist es durch die automatisierte Variablenselektion in unserem Verfahren prinzipiell leicht möglich, relevante und irrelevante Prognose-Regressoren zu erkennen. Allerdings zeigt sich in unserem konkreten Fall, dass im Gesamtsystem keine der insgesamt aufgenommenen Variablen als völlig irrelevant ausgeschlossen wird. Insofern können wir hier die Sinnhaftigkeit verschiedener ähnlicher neuerer Prognoseverfahren bestätigen, die auf zunehmend größere Datensätze und Informationsmengen rekurren.

Allerdings ist die hier gewählte Modellklasse nicht frei von Nachteilen: VARs leiden bekanntlich am „Fluch der großen Dimension“, d.h. die Kombination von zu vielen Variablen und zugelassenen Verzögerungen stößt leicht an die Obergrenze der verfügbaren Freiheitsgrade. Dies trifft auch in unserem Fall zu; unsere Modelle umfassen 21 Variablen, wodurch die anfänglich mögliche maximale Laglänge bei der gegebenen Stützstichprobe auf sechs begrenzt ist. In der Tat stellt auch schon die Dimension von 21 einen Kompromiss dar: Wir verwenden von allen Variablen mit bekannter Revisionshistorie nur die ersten beiden Veröffentlichungen pro Berichtsmonat, um den Rahmen der Freiheitsgrade nicht zu sprengen. In unserem Fall erscheint dies zwar nicht sehr problematisch, da bereits die Zweitveröffentlichung (d.h. die erste Revision) den Großteil der Endinformation enthält (siehe Kapitel 3). Falls jedoch zukünftig von weiteren Variablen die Revisionshistorien zur Verfügung stehen würden, würde sich dieses Problem verschärfen.

Ein weiterer theoretischer Nachteil des hier gewählten Ansatzes besteht u.U. darin, dass zur Entscheidung, ob eine Rezession vorliegt, analog zum im Kapitel 4 vorgestellten Dreiecksansatz mehrere Berichtsmonate zusammengefasst werden müssen. Im Gegensatz dazu verwenden die Probit- und Markov-Switching-Ansätze ein mehr oder weniger direkt fassbares Regime-Konstrukt. In diesem Kapitel jedoch wird ein gleitender Durchschnitt (von 5 Monaten) daraufhin untersucht, ob er die Schwelle zur Rezession durchbricht. Daher leidet

die VAR-Prognosemethode explizit unter der Rezessionserkennungsverzögerung, während die nichtlinearen Regimeschätzungen nur implizit davon betroffen sind.

Das Problem der Knappheit der Freiheitsgrade in großen dynamischen Modellen kann auf verschiedene Weisen gelöst werden, die alle eine Art von „shrinkage“-Schätzmethoden darstellen. Zum einen kann die Dimension des Datenraums reduziert werden durch die Extraktion von gemeinsamen Faktoren oder Hauptkomponenten. Diese Faktormodelle sind ein aktives Forschungs- und Anwendungsfeld, allerdings ist es nicht ganz einfach, die Relevanz bestimmter Prediktoren zu bestimmen. Alternativ kann zum zweiten der Parameterraum reduziert werden, wobei es wiederum mehrere Möglichkeiten gibt. Falls es a priori-Information zur Ausprägung der Koeffizienten gibt, können bayesianische Methoden angewendet werden. Kürzlich zeigten Banbura et al. (2010), dass große BVARs sehr kompetitive Prognosequalitäten haben können. Bei geeigneter Wahl der priors existiert dort auch kein anfängliches Freiheitsgrad-Problem. Abgesehen von praktischen Problemen bei der Implementierung der gewünschten Parameter-priors ist es allerdings oft zweifelhaft, ob tatsächlich a priori-Information zu den Koeffizienten vorliegt, oder ob BVARs einfach als beliebiges shrinkage-Werkzeug verwendet werden. In diesem Fall ist auch ein eher klassischer Ansatz im Geiste der *general-to-specific*-Methodologie möglich, bei der insignifikante Prediktoren sukzessive entfernt werden, um die Effizienz der Prognose zu erhöhen. Diesen „Subset-VAR“-Ansatz mit automatisierten Nullrestriktionen verfolgen wir in diesem Kapitel.

Wir betrachten in unserer Anwendung zwei grundsätzliche Modellvarianten: In der ersten werden alle Variablen in ihrer ursprünglichen Ausprägung in (ggf. logarithmierten) Niveaus belassen, in der zweiten Variante werden die nicht-stationären Variablen differenziert und so in eine Wachstumsraten-Darstellung überführt. Die log-Niveau-Variante berücksichtigt implizit das Vorliegen von vielfältigen Kointegrationsbeziehungen zwischen den Variablen. Z.B. sollten die Erst- und Zweit-Veröffentlichungen desselben Variablenkonzepts kointegriert sein, und auch zwischen der Industrieproduktion und Variablen wie Auftragseingängen ist Kointegration zu erwarten. Die zweite Modellvariante enthält per Konstruktion nur stationäre Variablen; solche Spezifikationen werden in der Literatur auch oft bei Kurzfristprognosen der Inflationsrate angewendet. Zwar werden hier die Niveauintformationen vernachlässigt, d.h. die impliziten Fehlerkorrekturterme des Niveau-Modells sind nicht enthalten, dies könnte aber für relativ kurze Prognosehorizonte vernachlässigbar sein. Im Gegenzug erkaufte man sich durch die Differenzierung der Variablen u.U. eine gewisse Prognosestabilität, falls Verschiebungen in den Niveaubeziehungen auftreten könn-

ten (Clements & Hendry 1999). Letztlich ist dies jedoch eine empirisch zu beantwortende Frage.

Analysiert wurden ursprünglich 1- bis 5-Schritt Prognosen der Erstveröffentlichungen und der ersten Revisionen (Zweitveröffentlichungen). Bei der Terminologie ist zu beachten, dass die gewählte Modellklasse und die gewählte Zeitachse, bei der sich alles auf das Veröffentlichungsdatum bezieht, die „Prognoseschritte“ teilweise etwas anders definiert als dies in anderen Teilen dieser Studie der Fall ist: Wegen eines Verfügbarkeitslags von zwei Monaten bei der Erstveröffentlichung der Industrieproduktion bedeutet eine 1-Schritt-Prognose die Vorhersage des Werts des Berichtsmonats, der vor dem Prognosezeitpunkt liegt (backcast, $t-1$). Entsprechend betrifft eine 2-Schritt-Prognose der Erstveröffentlichung den laufenden Berichtsmonat (nowcast, t). Bei der Prognose der ersten Revisionen verschiebt sich dies entsprechend um einen Monat (1-Schritt: Berichtsmonat $t-2$, 2-Schritt: backcast $t-1$, 3-Schritt: nowcast t). Es ergab sich generell, dass die 4- und 5-Schritt-Prognosen für die Wendepunktbestimmung unbrauchbar waren, da die Unsicherheit hier deutlich zu groß wurde. Daher stellen wir in diesem Kapitel nur die Ergebnisse für 1- bis 3-Schritt-Prognosen dar. Diese kürzeren Prognosehorizonte sind zum Teil deutlich weniger ambitioniert als weiter voraus schauende Prognosen, die sich auf die Berichtsperioden nach dem Zeitpunkt der Prognoseerstellung beziehen. Allerdings wird sich zeigen, dass in unserem Fall das Auftreten der Rezessionssignale in Echtzeit trotzdem recht früh erfolgt.

Der gesamte Schätz- und Evaluationsalgorithmus wurde als *gretl*-Batchprogramm implementiert, siehe Cottrell & Lucchetti (2010).

6.2 VAR-Modellrahmen und Modellreduktion

Das zu Grunde liegende VAR hat die folgende übliche Form:

$$\mathbf{x}_t = \sum_{k=1}^K A_k \mathbf{x}_{t-k} + \boldsymbol{\mu} + \mathbf{u}_t, \quad t = 1 \dots T_\tau, \tau = 2007m1 \dots 2010m11 \quad (6.1)$$

Der Vektor \mathbf{x}_t enthält 21 Elemente: Erst- und Zweitveröffentlichungen der Variablen Industrieproduktion, inländische Auftragseingänge, ausländische Auftragseingänge und (Konsumenten-) Preisindex, sowie die Variablen ohne Revisionshistorie Ölpreis, CDAX, REX, offene Stellen, Euribor, Renditeaufschlag für Unternehmensanleihen, fünf verschie-

dene Zinsspreads, Ifo-Geschäftsklima- und ZEW-Erwartungsindizes; jeweils ggf. in transformierter Form wie an anderer Stelle der Studie dargestellt. Je nach erwähnter Modellvariante werden die nicht-stationären Variablen entweder alle in log-Niveaus aufgenommen oder aber durch Differenzierung in (approximative) Wachstumsraten transformiert. Die in jedem Fall als stationär behandelten Variablen sind die Zinsspreads sowie die Ifo- und ZEW-Indikatoren.

Nicht ganz so üblich ist die Anwendung solcher VAR-Modelle auf Echtzeitdaten mit Publikationslags: Es ist zu beachten, dass \mathbf{x}_t aber bei der hier gewählten Darstellung trotzdem eine „normale“ multivariate Zeitreihe ist in dem Sinne, dass die Elemente selbst nicht revidiert werden. Die (bzw. ein Teil der) Revisionsinformation ist dadurch enthalten, dass die verschiedenen veröffentlichten Revisionen jeweils in eigene Zeitreihen eingetragen werden.² Der Zeitindex t bezieht sich dabei immer auf den Publikationszeitpunkt des jeweiligen Vektorelements. Wenn z.B. $x_{1,t}$ die Erstveröffentlichung der (log) Industrieproduktion darstellt und $x_{2,t}$ die Zweitveröffentlichung, dann enthält $x_{1,2000m3}$ wegen des Publikationslags von zwei Monaten den Wert der Industrieproduktion für den Berichtsmonat zwei Monate vorher, d.h. für 2000m1. Aber auch $x_{2,2000m4}$ enthält einen Wert für denselben Berichtsmonat 2000m1, nämlich den der ersten Revision. Das Element $x_{1,2000m3}$ selbst ist aber fix. Die vollständige Revisionsinformation erhielte man, indem noch beliebig viele weitere Zeitreihen mit der Dritt-, Viert-Veröffentlichung usw. aufgenommen würden. Dies jedoch würde den Rahmen der anfänglich verfügbaren Freiheitsgrade sprengen.

Die verfügbare Stichprobe ist in diesem Modellrahmen dadurch begrenzt, dass tatsächlich auch am Anfang der Stichprobe die Erst- und Zweitveröffentlichungen der entsprechenden Variablen noch vorliegen müssen, um Fehlwerte in \mathbf{x}_t zu vermeiden, und nicht erst spätere Rechenstände. Daher kann unsere Stützstichprobe nicht bereits 1993 beginnen, sondern erst 1995. Das jeweilige Ende der Stichprobe T_τ ist variabel, da für jeden Rechenstand in der Evaluationsperiode neu geschätzt wird.

Unser Schätzansatz für ein gegebenes τ beginnt mit üblichen OLS-Schätzungen jeder einzelnen Gleichung dieses Systems, bei der das bekannte Problem der begrenzten Freiheitsgrade auftaucht.³ In unserem Fall wäre die maximal zulässige Zahl der Verzögerungen $K = 6$. Da große Systeme zum Abbilden der ökonomischen Prozesse erfahrungsgemäß

²Mit „Revision“ ist hier nicht die Differenz zur vorherigen Erstveröffentlichung gemeint, sondern der geänderte (log-) Indexwert.

³Durch die VAR-Form ist dies auch gleichzeitig der effiziente Systemschätzer, dies ist aber hier durch die folgende Modellreduktion irrelevant.

weniger Verzögerungen benötigen als kleine Systeme, sollte diese Halbjahresspanne ausreichen. Für die Wahl von K gibt es wiederum verschiedene Möglichkeiten; auf Grund des besten \bar{R}^2 der Niveau-Gleichung für die Erstveröffentlichung der Industrieproduktion (für die anfängliche Stützstichprobe) entschieden wir uns für $K = 4$. Für die Spezifikation mit Wachstumsraten bedeutet dies entsprechend $K = 3$, da dort die vierte Niveau-Verzögerung implizit in den Differenzen enthalten ist.

Die nächsten Schritte sind durch eine automatische sequenzielle Elimination von insignifikanten Regressoren gegeben. Um übermäßigen numerischen Aufwand zu vermeiden, führen wir diese Elimination separat für jede Gleichung durch, statt jeweils das Gesamtsystem effizient zu schätzen. Der Regressor mit dem höchsten p-Wert der üblichen zweiseitigen t-Statistik wird jeweils aus der betreffenden Gleichung entfernt und die Gleichung wird (wieder mit OLS) neu geschätzt. Dieses Verfahren wird so lange wiederholt, bis kein Regressor einen p-Wert über der gewählten Schwelle p_{cutoff} mehr aufweist. Wie in anderen Teilen dieser Studie auch wurde hier eine Schwelle von $p_{cutoff} = 0,05$ gewählt. Insgesamt führt grob gesprochen dieses Verfahren im vorliegenden Fall dazu, dass mindestens die Hälfte der Regressoren entfernt wird.

Sobald alle Gleichungen auf diese Art reduziert wurden, wird das Gesamtsystem mit der SUR-Methode neu geschätzt, d.h. in effizienter Art und Weise unter Berücksichtigung der gleichungsübergreifenden Kovarianzmatrix der Residuen. Die (nicht iterierte) SUR-Methode ist bekanntlich ein Spezialfall der *feasible*-GLS-Methode und daher ebenfalls mit analytischen Lösungen berechenbar. Dieses so parametrisierte System wird dann für die Prognosen ausgehend vom Echtzeit-Prognosepunkt τ verwendet.

Für jedes τ in der Evaluationsperiode wird der beschriebene Algorithmus mit der neuen vorliegenden Stichprobe wiederholt, d.h. die Parameterschätzungen werden nicht fixiert.

6.3 In-Sample-Eigenschaften

Durch die große Dimension des zu Grunde liegenden Systems ist es kaum möglich, die Eigenschaften aller geschätzten Gleichungen detailliert zu charakterisieren. Wir beschränken uns hier daher v.a. auf die beiden Gleichungen der Industrieproduktion (Erst- und Zweitveröffentlichung). Außerdem stellen wir nur die Ergebnisse für die kürzeste Stützstichprobe dar, d.h. für den Datenstand 2007m1.

Zunächst zeigen die Abbildungen 6.1 und 6.2 die In-sample-Prognosefehler (Ein-Schritt)

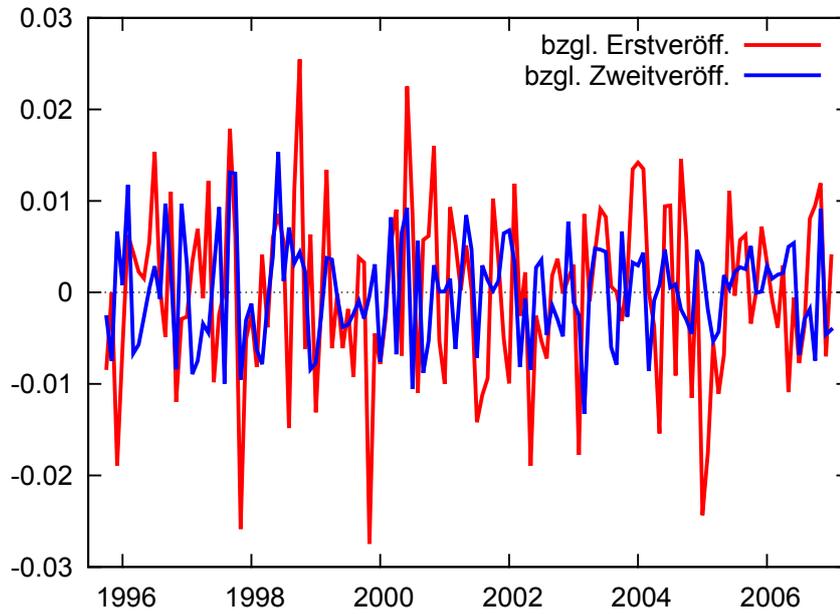


Abbildung 6.1: Ein-Schritt-Prognosefehler (Residuen in-sample) für log-Niveau-Spezifikation

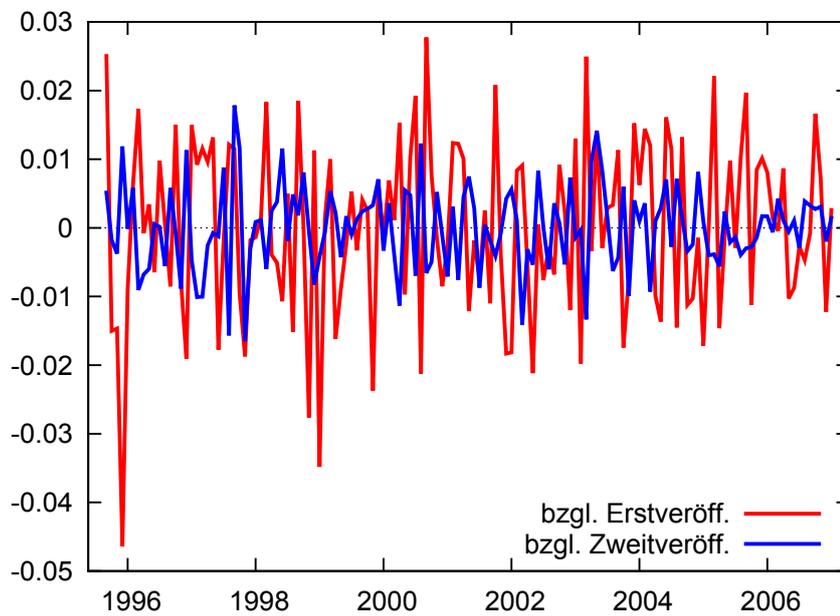


Abbildung 6.2: Ein-Schritt-Prognosefehler (Residuen in-sample) für Wachstumsraten-Spezifikation

Tabelle 6.1: Test auf Normalverteilung der Prognosefehler

	bzgl. Erstveröff.	bzgl. Zweitveröff.
log-Niveau-Spezifikation	2,60 (0,273)	0,70 (0,706)
Wachstumsraten-Spezifikation	4,26 (0,119)	1,52 (0,467)

Anmerkungen: Jarque-Bera-artiger Test mit $\chi^2(2)$ -Verteilung, p-Werte in Klammern. Die Prognosefehler sind in den Abbildungen 6.1 und 6.2 dargestellt.

der Gleichungen für die Industrieproduktion der verschiedenen Modellvarianten. Die Varianz der Prognosefehler bei der Zweitveröffentlichung ist deutlich geringer; dies liegt natürlich daran, dass im jeweiligen Monat vorher bereits die Erstveröffentlichung für denselben Monatsmonat vorliegt und zur Prognose der Zweitveröffentlichung verwendet werden kann, was einen beträchtlichen Teil der Unsicherheit eliminiert.

Für die verwendete Methode ist es zum Teil bedeutsam, ob die Residuen (d.h. Prognosefehler) als normalverteilt betrachtet werden können. In Tabelle 6.1 sind die formalen Tests dieser Nullhypothese aufgeführt, mit dem für unseren Modellansatz erfreulichen Ergebnis, dass diese Hypothese nicht abgelehnt werden kann. Der Ablehnung am nächsten kommt die Wachstumsraten-Spezifikation bzgl. der Erstveröffentlichung, was offensichtlich am Ausreißer Ende 1995 liegt, der auch in Abbildung 6.2 zu erkennen ist.

6.4 Relevante Prediktoren

Es ist natürlich von erheblichem Interesse, welche Indikatoren für die Prognosegleichungen relevant sind. Die Tabellen 6.2 und 6.3 geben eine Übersicht über die nicht ausgeschlossenen Regressoren in den Gleichungen für die Industrieproduktion. Dabei unterscheiden wir nicht zwischen den verschiedenen möglichen Verzögerungen. Konstruktionsbedingt durch die automatische Variablenselektion sind alle verbleibenden Terme hoch signifikant, weswegen wir auf die Angabe von p-Werten o.ä. verzichten. Bei der Interpretation dieser Tabelle ist zu beachten, dass ein wegfallender Prediktor in einer Gleichung zunächst nur für die Ein-Schritt-Prognose irrelevant ist. Für Mehr-Schritt-Prognosen einer Variable y kann auf Grund der Querverbindungen im System der Prediktor z trotzdem relevant sein, auch wenn er in der Gleichung von y nicht direkt enthalten ist.

In der für die log-Niveau-Spezifikation maßgeblichen Tabelle 6.2 ergibt sich vor allem, dass grundsätzlich alle Variablen zur Gesamtinformation etwas beisteuern und kei-

ne gänzlich irrelevant ist. Dies betrifft auch die verschiedenen Revisionen der entsprechenden Variablen, insofern scheint im Zusammenspiel der Revisionen durchaus auch systematische Information enthalten zu sein. Abgesehen davon fällt auf, dass der Ifo-Geschäftsklimaindex nur in der Gesamtstichprobe und für die Erstveröffentlichung signifikant erscheint.

In der folgenden Tabelle 6.3 bezüglich der Spezifikation in Wachstumsraten ist dann allerdings ersichtlich, dass einige Variablen ganz aus den Gleichungen für die Industrieproduktion herausfallen. Die Inflationsraten z.B. erweisen sich hier als offensichtlich nachlaufende Variablen, die keinen Prognosegehalt für die Produktion aufweisen. Gleiches gilt für den Ölpreis, den Rentenindex REX und den Renditeaufschlag für Unternehmensanleihen (*corp. spread*). Aber als interessantes Ergebnis ist insbesondere zu nennen, dass sowohl der Ifo- wie auch der ZEW-Index als augenscheinlich irrelevant für die Erstveröffentlichung der Produktion eingestuft werden.

Tabelle 6.2: Regressoren in den Produktionsgleichungen, log-Niveau-Spezifikationen

	$\tau = 2007m1$ Erstveröff. (Rev. 0)	$\tau = 2010m11$ Erstveröff. (Rev. 0)	$\tau = 2007m1$ Zweitveröff. (Rev. 1)	$\tau = 2010m11$ Zweitveröff. (Rev. 1)
Ind'prod. Rev. 0	x	x	x	x
Ind'prod. Rev. 1	x	x	x	x
CPI Rev. 0		x	x	
CPI Rev. 1	x	x	x	
AA Rev. 0	x	x	x	x
AA Rev. 1	x	x	x	
AI Rev. 0	x	x		x
AI Rev. 1		x	x	x
Spread 10Y		x	x	
Spread 1Y		x		x
Spread 2Y	x	x		x
Spread 3Y		x	x	x
Spread 5Y		x		x
Ifo-G'klima		x		
ZEW-Erwart.	x	x		x
Ölpreis	x			x
CDAX	x	x		x
REX			x	x
offene Stellen	x	x		x
Euribor	x		x	
Corp. spread	x		x	

Anmerkungen: AA = ausländische Auftragseingänge, AI = inländische Auftragseingänge

Tabelle 6.3: Regressoren in den Produktionsgleichungen, Wachstumsraten-Spezifikationen

	$\tau = 2007m1$ Erstveröff. (Rev. 0)	$\tau = 2010m11$ Erstveröff. (Rev. 0)	$\tau = 2007m1$ Zweitveröff. (Rev. 1)	$\tau = 2010m11$ Zweitveröff. (Rev. 1)
Ind'prod. Rev. 0	x	x	x	x
Ind'prod. Rev. 1	x	x	x	x
CPI Rev. 0				
CPI Rev. 1				
AA Rev. 0	x			
AA Rev. 1	x			
AI Rev. 0	x			
AI Rev. 1		x	x	x
Spread 10Y			x	x
Spread 1Y		x	x	x
Spread 2Y	x	x		x
Spread 3Y		x	x	
Spread 5Y		x		
Ifo-G'klima				
ZEW-Erwart.			x	
Ölpreis				
CDAX	x	x		
REX				
offene Stellen	x	x	x	x
Euribor		x	x	
Corp. spread				

Anmerkungen: AA = ausländische Auftragseingänge, AI = inländische Auftragseingänge

6.5 Schätzung der Rezessionswahrscheinlichkeit

Mittels des geschätzten Subset-VARs lassen sich prognostizierte Rezessionswahrscheinlichkeiten auf die folgende Weise ableiten, wobei zwischen dem VAR in (Log-) Niveaus und in Wachstumsraten unterschieden werden muss.

6.5.1 Niveau-Spezifikation

Beginnen wir mit dem etwas einfacheren Fall der log-Niveau-Spezifikation. Hierbei genügt die Verteilung der h -Schritt-Prognose (der Industrieproduktion). Unter der Annahme von normalverteilten Residuen und unter Vernachlässigung der Parameterunsicherheit ist die Dichteverteilung der h -Schritt-Prognose ebenfalls normal und hängt nur von der h -Schritt-Punktprognose des Systems sowie von der Kovarianzmatrix der Systemresiduen ab. Die Standardabweichung der Prognoseverteilung (bzw. äquivalent des Prognosefehlers) wird von einschlägigen Softwarepaketen automatisch ausgegeben.

Nummehr subtrahieren wir von dieser Prognoseverteilung den bekannten (log-) Wert der entsprechenden Variable fünf Berichtsmonate vorher, um die kumulierte Wachstumsrate über diesen 5-Monatszeitraum zu berechnen. Dabei verwenden wir immer die aktuellste in Echtzeit verfügbare Revision für diese frühere Berichtsperiode. Als Zwischenergebnis erhalten wir die Verteilung der prognostizierten kumulierten Wachstumsrate zu diesem h -Schritt-Horizont, deren Varianz nach wie vor durch die schon geschätzte Prognosefehlervarianz gegeben ist. Schließlich wird ermittelt, wieviel Wahrscheinlichkeitsmasse unterhalb des gewählten Schwellwerts der Rezessionsdefinition liegt. Diese Masse gibt die Wahrscheinlichkeit dafür an, dass für die Publikation in $t + h$ die Rezessionsdefinition erfüllt sein wird.⁴ Der hier verwendete 5-Monats-Schwellwert beträgt $-0,025 * 2/5 = -1\%$, vergleiche auch Kapitel 4.

Diese Methode ist offensichtlich vom Rezessionserkennungslag betroffen. Wenn man in der Publikations-Periode t aus der h -Schritt-Prognose der Erstveröffentlichung ein Rezessionssignal erhält (d.h. eine prognostizierte Wahrscheinlichkeit von z.B. über 50%), so bedeutet dies im Grunde, dass der Wendepunkt vor ca. $-(h - 2 - 5)$ Monaten lag und seitdem die Rezession bereits im Gange ist. Dies ist einerseits zwar bedauerlich, andererseits ist dies einfach die logische Konsequenz dessen, dass ein Rückgang der Wirtschaftsaktivi-

⁴Auf welchen Berichtsmonat sich diese Publikation bezieht, hängt von der gerade betrachteten Revision ab. Für die Erstveröffentlichung mit Publikationslag 2 lautet der Berichtsmonat $t + h - 2$, für die Zweitveröffentlichung entsprechend $t + h - 3$.

tät laut Definition eine gewisse Mindestdauer aufweisen muss, um ihn auch als Rezession bezeichnen zu können.

6.5.2 Spezifikation in Wachstumsraten

Wenn das System in Wachstumsraten spezifiziert ist, muss eine explizite Kumulation der Prognosefehler von Schritt 1 bis Schritt h erfolgen, da konstruktionsgemäß keine Niveau-Information vorliegt. Daher reicht in diesem Fall die Verteilung des h -ten Schritts allein nicht mehr aus. Die Ermittlung der Verteilung der kumulierten Wachstumsrate wird auch dadurch erschwert, dass die Prognosefehler nicht unabhängig sind.

Aus diesem Grund verwenden wir in diesem Fall ein einfaches parametrisches Monte-Carlo-Simulationsverfahren. Hierbei werden ausgehend von der geschätzten Verteilung der Systemresiduen (d.h. v.a. mit Hilfe ihrer geschätzten Kovarianzmatrix und wieder unter Annahme einer Normalverteilung) wiederholt Zufallsrealisationen aus dieser multivariaten Verteilung gezogen und damit die möglichen Pfade der endogenen Variablen, insbesondere der Industrieproduktion, bis zum Horizont h berechnet. Für jeden Pfad werden die Wachstumsraten summiert zur Approximation des kumulierten Wachstums. Für Horizonte $h < 5$ werden zusätzlich die bekannten $5 - h$ Wachstumsraten der Vergangenheit hinzugerechnet, um immer einen einheitlichen 5-Monats-Zeitraum zu betrachten. Mit dieser simulierten Häufigkeitsverteilung der kumulierten Wachstumsraten lässt sich dann wie vorher weiterarbeiten. Da in unserem Anwendungsfall das Zentrum der Verteilung relevanter ist als die Ränder, reicht eine relativ kleine Anzahl an Simulationsläufen aus; konkret arbeiten wir hier mit 500 Wiederholungen. Natürlich muss diese Simulation aber für jede neue Stichprobe im Evaluationszeitraum neu durchgeführt werden.

6.6 Out-of-Sample-Evaluation der Wendepunktprognosen

Wir wenden uns nun der Prognose-Performance unseres Ansatzes im zur Verfügung stehenden Evaluationszeitraum zu. Wie bereits erwähnt bezeichnet die Zeitachse in den folgenden Darstellungen ausschließlich den Echtzeit-Informationsstand, d.h. die Datenverfügbarkeit; z.B. sind alle Einträge bei 2007m1 mit den verfügbaren Daten in eben diesem Monat geschätzt. Das bedeutet z.B. für die 2-Schritt-Prognose, dass bis zur Veröffentli-

chung in zwei Monaten prognostiziert wird (und damit dann implizit die Rezessionswahrscheinlichkeit geschätzt wird). Wegen des Verfügbarkeitslags von zwei Monaten bedeutet dies also eine Prognose der Erstveröffentlichung für den laufenden Berichtsmonat (nowcast). Entsprechendes gilt für die anderen Prognoseschritte. Entscheidend ist, dass die abgebildeten Rezessionssignale (oder deren Ausbleiben) als Echtzeitsignale interpretiert werden können. Die betroffenen Berichtsmonate müssen implizit mit dem Wissen über die Verfügbarkeitslags hergeleitet werden.

In Abbildung 6.3 sind zunächst die prognostizierten Rezessionswahrscheinlichkeiten auf Grund der Prognosen der Erstveröffentlichungen dargestellt, in der oberen Grafik aus der Spezifikation in Wachstumsraten, in der unteren Grafik bzgl. der Niveau-Spezifikation. Die schraffierten Flächen dienen zur Abgrenzung von „trivialen“ Signalen: in diesem Bereich ist die verwendete Rezessionsdefinition bereits mit den veröffentlichten Daten erfüllt (bezogen auf den aktuellsten verfügbaren 5-Monats-Zeitraum), ohne dass es noch einer Prognose bedürfte. Allerdings ist natürlich die Definition der schraffierten Fläche rein vergangenheitsbezogen, während die Signale der Prognosen einen anderen Zeitraum betreffen. Dieses einfache ex post-Rezessionssignal schlägt in 2008m8 an. Es ist auch ersichtlich, dass es gegen Ende 2008 einen Publikationsmonat gibt, bei dem das kumulierte 5-Monats-Wachstum über der Trennschwelle liegt, so dass die schraffierte Fläche eine Lücke aufweist.

In der Tat gelingt es den prognostizierten Rezessionswahrscheinlichkeiten durchaus, früher ein Rezessionssignal zu liefern als die veröffentlichten Daten. Gleichwohl leiden einige der Prognosen darunter, fälschlicherweise Rezessionen anzuzeigen, die nicht stattfanden. Da dies an einer gewissen vorhandenen Volatilität der Wahrscheinlichkeitsprognosen liegt, gibt es noch die Möglichkeit, die verschiedenen Prognosen zu kombinieren, indem wir die durchschnittliche Wahrscheinlichkeit über alle Prognosehorizonte betrachten. Das Ergebnis ist in Abbildung 6.4 dargestellt und ist durchaus ermutigend. Insbesondere die Niveau-Spezifikation zeigt nun keine falsche Rezession mehr an. Gleichzeitig ergibt sich das erste Rezessionssignal in 2008m6, zwei Monate vor dem Signal der veröffentlichten Daten. Auch das Ende des Vorliegens der Rezessionsdefinition wird von dieser Reihe zwei Monate vor der entsprechenden Datenveröffentlichung angezeigt.

Die entsprechenden Ergebnisse für die Prognose bezüglich der Zweitveröffentlichung sind in den Abbildungen 6.5 und 6.6 dargestellt. Auffallend ist hierbei, dass generell eine geringere Neigung zu falschen Rezessionssignalen zu bestehen scheint. Natürlich ist aber auch zu beachten, dass die Zweitveröffentlichung sich immer auf einen Berichtsmonat

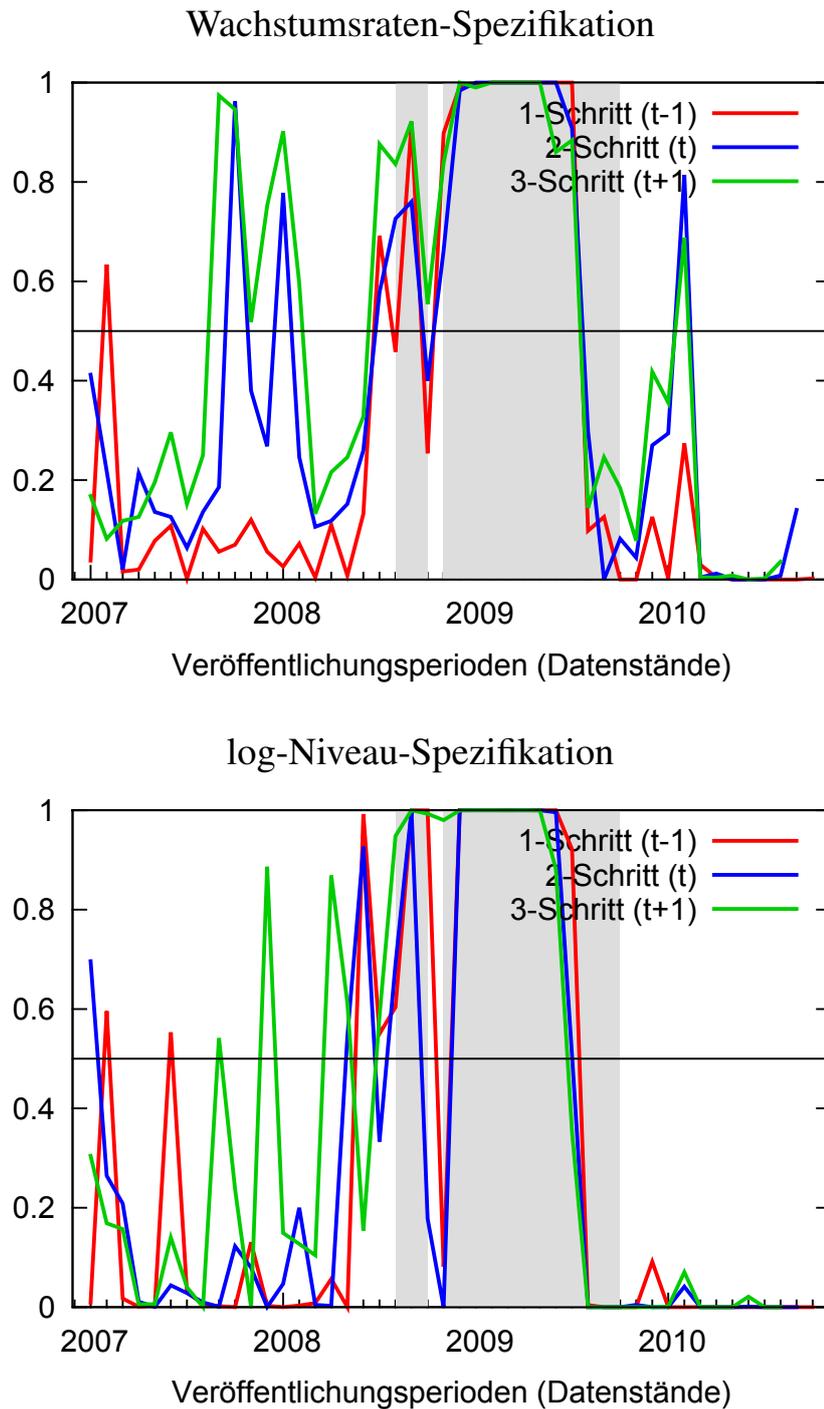


Abbildung 6.3: Out-of-sample-Rezessionswahrscheinlichkeiten (Rev. 0) – d.h. gemessen an Erstveröffentlichung. Obere Grafik VAR-Spezifikation in Wachstumsraten (für die $I(1)$ -Variablen) mit max. 3 Verzögerungen, untere Grafik VAR-Spezifikation in log-Niveaus mit max. 4 Verzögerungen. — In den schattierten Bereichen ist die zu Grunde gelegte Rezessionsdefinition mit den in Echtzeit bekannten Daten schon erfüllt, ein Wendepunktprognosesignal müsste also jeweils links der Schattierungsgrenzen auftreten, um nicht einfach den Rezessionserkennungslag widerzuspiegeln.

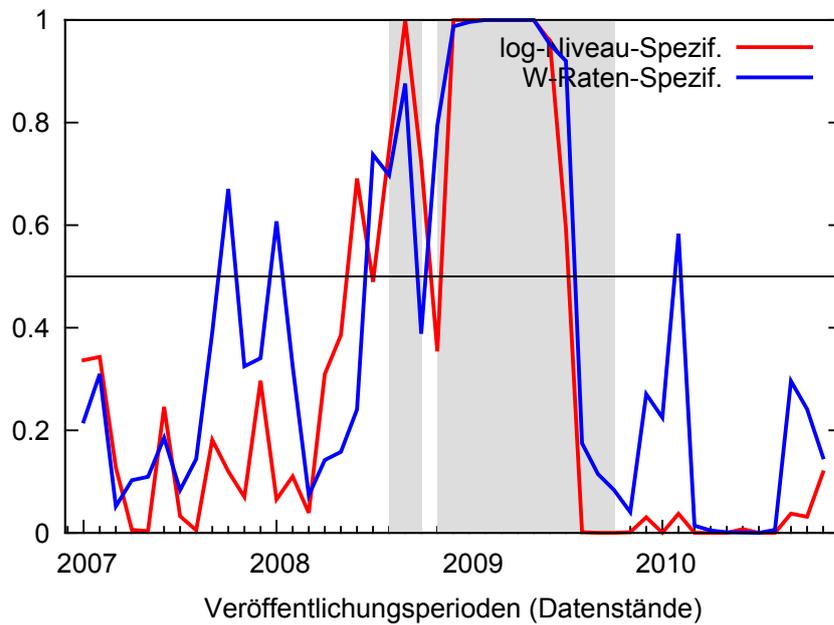


Abbildung 6.4: Durchschnittliche Out-of-sample-Rezessionswahrscheinlichkeiten (Rev. 0) – Durchschnitte der vorher gezeigten Prognosehorizonte, immer gemessen an der Erstveröffentlichung. Siehe auch die Erläuterungen zu Abbildung 6.3.

weiter zurück bezieht, so dass die Prognoseunsicherheit offensichtlich geringer sein muss. Entscheidend ist hierbei also, ob sich diese zusätzliche Zeitverzögerung auch im Timing des Rezessionssignals niederschlägt. Erstaunlicherweise ist dies hier nicht unbedingt zu beobachten; insbesondere bei der Niveau-Spezifikation ergibt sich ein Signal, das wieder zwei Monate vor dem Signal der veröffentlichten Daten liegt (mit Ausnahme des 1-Schritt-Horizonts).

Obwohl dies bei der Analyse der Zweitveröffentlichungen weniger dringlich zu sein scheint, ist eine einfache Durchschnittsbildung über die Horizonte auch hier möglich, siehe Abbildung 6.6. Für die Niveauspezifikation ist das Ergebnis recht ähnlich wie vorher bei der Erstveröffentlichung, bzgl. der Wachstumsraten-Spezifikation ist eine deutliche Verbesserung sichtbar, es werden keine falschen Rezessionssignale gesendet.

Schließlich ist es neben der Betrachtung des Timings der Rezessionssignale bei Prognosen allgemein üblich und sinnvoll, quantitative Evaluationsmaße heranzuziehen. Diese sind für die hier verwendeten Modellvarianten in Tabelle 6.4 zu finden. Bei der Interpretation der Evaluationsstatistiken in Tabelle 6.4 sind allerdings die Problematiken des Informationsstands und des Rezessionserkennungslags zu beachten: Zum Vergleich mit der ex post-Benchmarkreihe, die sich auf Berichtsmonate bezieht, verschieben wir die Reihen der prognostizierten Wahrscheinlichkeiten um den Publikationslag (rückwärts) und um den Prognosehorizont (vorwärts). Im Ergebnis vergleichen wir also die Rezessionssignale für jeweils gleiche Berichtsmonate. Allerdings ist hierbei zu beachten, dass trotzdem die unterschiedlichen Informationsstände der Prognosen teilweise unberücksichtigt bleiben. Z.B. bezieht sich die 2-Schritt-Prognose der Erstveröffentlichung zur Prognosezeit t auf denselben Berichtsmonat (nämlich gerade t) wie die 1-Schritt-Prognose der Erstveröffentlichung zur Prognosezeit $t + 1$. Natürlich hat aber die Prognose zur Zeit $t + 1$ einen Informationsvorteil. Dies ist sowohl bei Vergleichen verschiedener Spezifikationen dieses Kapitels zu beachten, wie auch bei Vergleichen mit anderen Spezifikationen und Modellen dieser Studie. Daher lässt sich die insgesamt am besten geeignete Spezifikation nicht allein aus Tabelle 6.4 ersehen.

Des Weiteren bedeutet in diesem Kapitel ein Rezessionssignal, dass der implizierte Berichtsmonat ca. fünf Monate nach dem eigentlichen Wendepunkt liegt, da ja die kumulierte Wachstumsrate über diesen Zeitraum betrachtet wird. Im Gegensatz dazu wird die Benchmarkreihe vollständig ex post generiert. Wir verschieben aber unsere prognostizierten Reihen *nicht* auch noch um diesen Rezessionserkennungslag, um die Vergleichbarkeit mit den anderen Prognosen nicht völlig auszuhebeln.

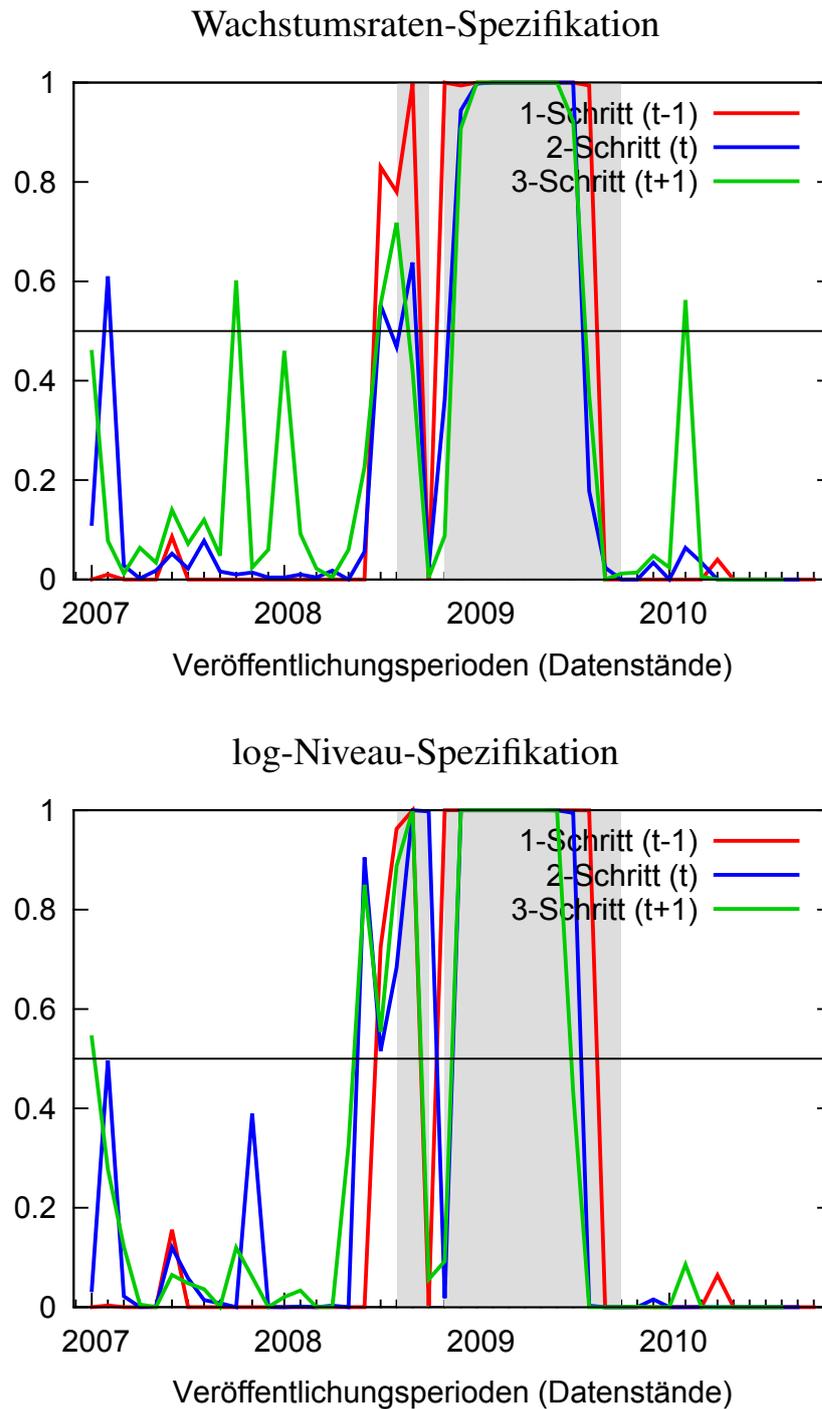


Abbildung 6.5: Out-of-sample-Rezessionswahrscheinlichkeiten (Rev. 1) – d.h. gemessen an der Zweitveröffentlichung (erste Revision). Obere Grafik VAR-Spezifikation in Wachstumsraten (für die $I(1)$ -Variablen) mit max. 3 Verzögerungen, untere Grafik VAR-Spezifikation in log-Niveaus mit max. 4 Verzögerungen. Siehe auch die Erläuterungen zu Abbildung 6.3.

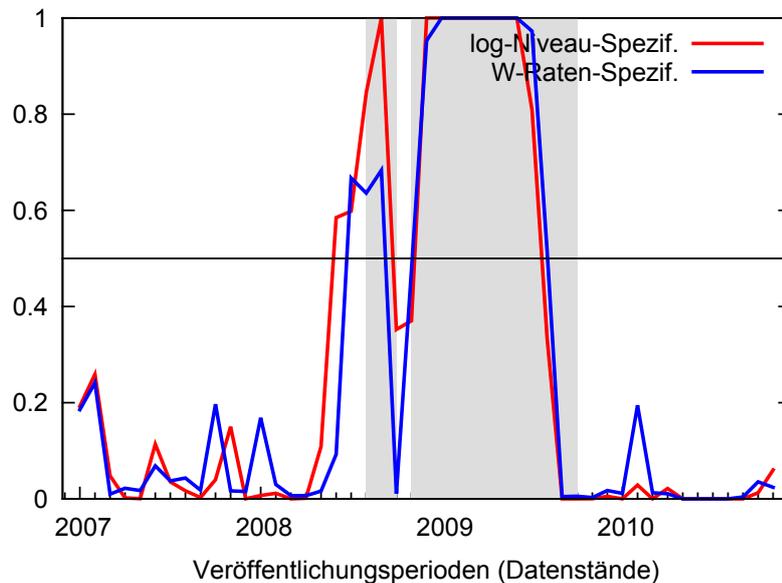


Abbildung 6.6: Durchschnittliche Out-of-sample-Rezessionswahrscheinlichkeiten (Rev. 1) – Durchschnitte der vorher gezeigten Prognosehorizonte, immer gemessen an der Zweitveröffentlichung. Siehe auch die Erläuterungen zu Abbildung 6.3.

Abgesehen von diesen Komplikationen der Interpretation ergibt sich aus den Evaluationsmaßen relativ deutlich, dass die Prognosen für die Zweitveröffentlichung am besten abschneiden. Je nach präferiertem Maß „gewinnt“ dabei entweder die 1-Schritt-Prognose in Wachstumsraten oder der Durchschnitt der Horizonte 1-3 in der Niveauspezifikation.

6.7 Schlussfolgerungen

In der Gesamtschau, die die Evaluationsergebnisse bezüglich der Benchmarkreihe ebenso berücksichtigt wie das Timing des Rezessionssignals, scheint die log-Niveau-Spezifikation beruhend auf der Prognose für die Zweitveröffentlichung die empfehlenswerte Modellvariante zu sein, insbesondere wenn die drei Prognosehorizonte zu einer durchschnittlichen Wahrscheinlichkeit zusammengefasst werden. Das Rezessionssignal wäre damit in Echtzeit zum Datenstand 2008m6 erstmals erfolgt. Trotz der Abhängigkeit dieser hier verwendeten Methode vom Rezessionserkennungslag ist ein relativ frühes Signal anscheinend möglich.

Bezüglich der Frage, welche Indikatoren am nützlichsten für die Wendepunkterkennung

Tabelle 6.4: Out-of-sample-Evaluation (Vergleich mit ex-post-Benchmark) der VAR-Spezifikationen

Spezifikation	MAE	RMSE	Theil
<i>log-Niveau-Spezif.</i>			
1-Schritt, Rev. 0	0,1543	0,3498	0,3280
1-Schritt, Rev. 1	0,1203	0,3329	0,3104
2-Schritt, Rev. 0	0,2129	0,4035	0,3861
2-Schritt, Rev. 1	0,1563	0,3562	0,3319
3-Schritt, Rev. 0	0,2180	0,3947	0,3524
3-Schritt, Rev. 1	0,2064	0,3991	0,3788
Durchschnitt, Rev. 0	0,1972	0,3475	0,3308
Durchschnitt, Rev. 1	0,1569	0,3256	0,3156
<i>Wachstumsraten-Spezif.</i>			
1-Schritt, Rev. 0	0,1851	0,3636	0,3489
1-Schritt, Rev. 1	0,1187	0,3318	0,3103
2-Schritt, Rev. 0	0,3134	0,4594	0,4186
2-Schritt, Rev. 1	0,2024	0,4012	0,3958
3-Schritt, Rev. 0	0,3586	0,4999	0,4270
3-Schritt, Rev. 1	0,3012	0,4794	0,4603
Durchschnitt, Rev. 0	0,2991	0,4310	0,3970
Durchschnitt, Rev. 1	0,2019	0,3844	0,3787

Anmerkungen: Der Vergleich erfolgt mit der ex-post generierten Benchmarkreihe, wobei jeweils auf den gleichen Berichtsmonat abgestellt wird. Für Einzelheiten siehe die Erläuterungen im Haupttext. Rev. 0: Erstveröffentlichung, Rev. 1: Zweitveröffentlichung.

sind, ergaben sich bei der Niveau- und der Wachstumsratenspezifikation unterschiedliche Ergebnisse. Bei der Niveau-Spezifikation erscheinen grundsätzlich alle berücksichtigten Variablen relevant, während bei der Wachstumsratenspezifikation einige Variablen herausfallen, wie Preise, Ölpreis, REX und *corporate spread*, aber auch die Ifo-Geschäftsklima- und ZEW-Erwartungsindizes. Insofern scheinen die Dynamik der realwirtschaftlichen revisionsanfälligen Variablen (wie Lags und Revisionen der Produktion und der Auftragseingänge) sowie die zeitnah verfügbaren Finanz- und Arbeitsmarktvariablen (Spreads, CDAX, offene gemeldete Stellen) in diesem Sinne wichtiger zu sein.

Abschließend sei generell noch einmal darauf hingewiesen, dass längere Prognosehorizonte so unerfreuliche Resultate lieferten im Sinne einer deutlich übertriebenen Volatilität, dass die Ergebnisse für $h = 4$ und $h = 5$ hier nicht berichtet wurden.

7 Wendepunktbestimmung mit Markov-Switching-Modellen

7.1 Einleitung

Dieses Kapitel behandelt das Problem der Erkennung von Wendepunkten mit Hilfe eines Markov-Switching-Modells (MS). Dieses stellt implizit mehr Information zur Verfügung als eine reine Vorhersage einer Rezessionswahrscheinlichkeit. Und da sich der überwiegende Teil der MS-Literatur bezüglich des Konjunkturzyklus auf autoregressive Schätzungen konzentriert, verspricht die in diesem Kapitel hergestellte Verbindung eines MS-Modells und verschiedener vorlaufender Indikatoren zusätzliche Hilfestellung für die Konjunkturdiagnose. Nichtsdestotrotz steht die frühzeitige und sichere Erkennung von Rezessionsregimen im Mittelpunkt des hier vorgestellten Modells.

Im Allgemeinen basiert ein Datierungs- oder Vorhersagealgorithmus für den Konjunkturzyklus auf zwei Fragen: Erstens, welche Technik wird benutzt und zweitens, welche Frühindikatoren bezüglich der Referenz-Zeitreihe werden integriert. Während eine Wendepunktbestimmung im Nachhinein nicht zwangsläufig auf solche Indikatoren angewiesen ist, erhalten Schätzungen zukünftiger Werte ihre Vorhersagekraft hauptsächlich aus der Tatsache, dass vorlaufende Indikatoren mit der passenden Technik verbunden werden. Seit dem fundamentalen Beitrag von Bry & Boschan (1971), in dem sie einen nicht-parametrischen Algorithmus zur Aufdeckung konjunktureller Wendepunkte entwickelten, wurden viele neue Verfahren und Frühindikatoren erforscht. An dieser Stelle seien nur Chauvet & Potter (2005), Stock & Watson (1989) sowie Hamilton (1989) erwähnt. Erstere nutzten verschiedene Spezifikationen des Probit-Modells. Stock & Watson (1989) entwickelten ein dynamisches Faktormodell für den Konjunkturzyklus. Hamilton (1989) schließlich führte das Markov-Switching-Modell ein, das wir in diesem Kapitel des Forschungsberichts auf deutsche monatliche Echtzeitdaten anwenden.

Gleichzeitig mit der Erforschung neuer Datierungsmethoden wurde die Menge vorläufiger realwirtschaftlicher Indikatoren (vgl. Kapitel 2) um Finanzmarktindikatoren erweitert. Zu nennen ist hier etwa das Zinsdifferenzial zwischen den Stützstellen der Laufzeitkurve (*term spread*), z.B. verwendet von Estrella & Hardouvelis (1991), und das Zinsdifferenzial zwischen privaten und öffentlichen Emittenten (*public-corporate* oder einfach *corporate spread*), z.B. verwendet von Friedman & Kuttner (1992). Kürzlich wurde der Zusammenhang zwischen diesem Differenzial und der Wirtschaftsentwicklung von Meeks (2011) analysiert. Da die Ausweitung der Volumina von Krediten, die mit einschlägigen Derivaten verknüpft wurden, als ein Grund für das Ausmaß der letzten Krise angesehen wird, liegt es nahe, auch das Kreditwachstum auf seine Vorhersagekraft für den Konjunkturzyklus hin zu untersuchen. Man beachte für die allgemeinen Zusammenhänge Biggs et al. (2009). Demgegenüber konzentriert sich dieses Kapitel auf das Kreditwachstum, wie es als Bilanzbestandteil monatlich von Finanzdienstleistungsinstituten an das Europäische System der Zentralbanken (ESZB) gemeldet wird.

Die Unterschiede zwischen realwirtschaftlichen und Finanzmarktindikatoren werden besonders evident, wenn man Vorhersagen in Echtzeit trifft. Während Finanzdaten, zumindest bei monatlicher Frequenz, unmittelbar zur Verfügung stehen und keinen Revisionen unterliegen, ist das Gegenteil für die meisten makroökonomischen Daten der Fall. Wie Diebold & Rudebusch (1991) für den US-amerikanischen *Composite Index of Leading Indicators* zeigen, beeinflussen Revisionen und tatsächlich noch mehr die verzögerte Datenverfügbarkeit die Vorhersagekraft von Frühindikatoren auf substanzielle Weise. Deshalb werden hier nicht nur Echtzeitdaten benutzt, sondern das verwendete Modell wird zusätzlich auf dem Echtzeitpfad angepasst. Die Berücksichtigung verschiedener Spezifikationen stellt eine konkrete Ausarbeitung eines Vorschlags von Hamilton (2010) dar. Während dies für andere Schätzverfahren schon ein teilweise üblicher Ansatz ist, etwa in Kapitel 8, versucht sich sehr wenig MS-Literatur an dieser Vorgehensweise. Tatsächlich können wir zeigen, wie eine von der Veröffentlichung abhängige Veränderung der Anzahl eingebetteter Regime Echtzeitvorhersagen stabilisiert.

Dieses Kapitel ist wie folgt aufgebaut. Abschnitt 7.2 wiederholt kurz die Hauptschritte des von Hamilton (1989, 1990) eingeführten Filters. Dabei wird die Tatsache betont, dass exogene Variablen sehr wohl einen Einfluss auf die durch das Markov-Switching erzeugten Zustandswahrscheinlichkeiten haben. Die Anwendung erfolgt durch eine von Perlin (2009) implementierte numerische Optimierung mittels des Expectation-Maximization (EM) Algorithmus.

Abschnitt 7.3 beschreibt den Input und die In-sample-Ergebnisse des Modells. Zunächst werden die Daten und damit sowohl die Auswahl als auch die Echtzeiteigenschaften der Referenz-Zeitreihe sowie der Frühindikatoren vorgestellt. Danach wird die Modellspezifizierung im Detail diskutiert. Dabei kommt zum Ausdruck, wie die Modellstruktur aus einem Kompromiss von Informationsbedarf und Echtzeitdatenlage entsteht. Während erstere eher die Verwendung eines hoch parametrisierten Modells empfiehlt, schränkt letztere den Parameterraum klar ein. Der nächste Teil des Abschnitt 7.3 beschreibt die In-sample-Ergebnisse der Veröffentlichung vom November 2010. Diese bestätigen die gewählte Modellspezifikation. Im letzten Teil des Abschnitts werden Standardmaße im Fall von mehr als zwei Regimen diskutiert. Basierend auf diesen Maßen wird analysiert, ob im Allgemeinen jene Frühindikatoren besser abschneiden, die keiner verzögerten Datenveröffentlichung unterliegen.

Abschnitt 7.4 behandelt die Vorhersageergebnisse des Modells. Zuerst werden die Merkmale der Echtzeitvorhersagen betrachtet, darunter insbesondere die Modellanpassung. Danach werden Out-of-Sample-Ergebnisse mit einmonatigem Prognosehorizont angegeben, bei denen die Industrieproduktion wieder als Referenz-Zeitreihe der gesamtwirtschaftlichen Aktivität fungiert. Abschnitt 7.5 fasst die Erkenntnisse des Kapitels zusammen.

7.2 Markov-Switching und exogene Variablen

Wenn man sich mit Konjunkturmodellen beschäftigt, ist es sinnvoll, eine klare und allgemein anerkannte Abgrenzung zwischen Rezession und Expansion zu suchen. Eine solche ist durch die schon im Abschnitt 4.1 erwähnte Definition des amerikanischen *National Bureau of Economic Research* gegeben (Business Cycle Dating Committee 2008, S.1).

Obwohl diese Definition auf Grund der darin genannten makroökonomischen Variablen die Verwendung eines fünf-dimensionalen MS-VAR-Systems suggeriert (vgl. Krolzig 1997), wird hier ein anderer Ansatz verfolgt. Das Hauptmotiv dafür ist die Tatsache, dass ein MS-VAR zusätzlich den Parameterraum vergrößern würde. Dies ist schwer in Einklang zu bringen mit der vorhandenen Echtzeitdatenlage, vgl. Tabelle 2.3. Aus demselben Grund erlauben auch eindimensionale Gleichungen ausschließlich die Verwendung zweier qualitativ unterschiedlicher Regressoren. Diese sind einerseits Lags der abhängigen Variable und andererseits eine exogene Variable einschließlich ihrer Lags. Insgesamt führt dies zu

der folgenden Form:

$$\begin{aligned}
 y_t &= \beta_0^{S_t} + \beta_{1,y}^{S_t} y_{t-f-D_y} + \sum_{j=2}^p \beta_{j,y}^{S_t} y_{t-j+1-f-D_y} \\
 &\quad + \beta_{1,x}^{S_t} x_{t-f-D_x} + \sum_{j=2}^q \beta_{j,x}^{S_t} x_{t-j+1-f-D_x} + u_t, \\
 u_t &\sim N\left(0, \sigma^{S_t}\right), \quad t = 1, \dots, T,
 \end{aligned} \tag{7.1}$$

wobei f den Vorhersagehorizont und D_y, D_x die Datenverfügbarkeit der abhängigen und unabhängigen Variable repräsentieren. S_t steht für die verschiedenen Zustände, die den Gesamtprozess der beobachtbaren Zeitreihe y_t generieren. Wenn im Folgenden die Schritte des von Hamilton (1989, 1990) entworfenen Filters wiedergegeben werden, wird zur Vereinfachung die Datenverfügbarkeitslücke auf 0 und der Vorhersagehorizont auf 1 gesetzt. Des Weiteren wird Vektornotation verwendet. So lässt sich (7.1) wie folgt schreiben:

$$\begin{aligned}
 y_t &= z_t' \beta^{S_t} + u_t, \quad u_t \sim N\left(0, \sigma^{S_t}\right) \quad \text{mit} \\
 z_t' &= (1, y_{t-1}, y_{t-2}, \dots, y_{t-p}, x_{t-1}, x_{t-2}, \dots, x_{t-q}) \\
 \beta^{S_t'} &= (\beta_0^{S_t}, \beta_{1,y}^{S_t}, \beta_{2,y}, \dots, \beta_{p,y}, \beta_{1,x}^{S_t}, \beta_{2,x}, \dots, \beta_{q,x}) \\
 \theta' &= (\beta^{S_t'}, \sigma^{S_t})
 \end{aligned} \tag{7.2}$$

Im Zentrum des Markov-Switching-Modells steht die Annahme, dass die unbeobachtbaren Zustände der abhängigen Variable einer Markov-Kette erster Ordnung folgen, deren Übergangsmatrix bei zwei Regimen wie folgt aussieht:

$$P = (p_{ij}) = \begin{pmatrix} \mathcal{P}(S_t = 1 | S_{t-1} = 1) & \mathcal{P}(S_t = 1 | S_{t-1} = 2) \\ \mathcal{P}(S_t = 2 | S_{t-1} = 1) & \mathcal{P}(S_t = 2 | S_{t-1} = 2) \end{pmatrix} \tag{7.3}$$

Später wird das Modell auf vier Regime erweitert, so dass allein für die Übergangsmatrix 16 Parameter zu schätzen sind. Dies ist der Punkt, an dem die vorhandene Echtzeitdatenlage eine Rolle spielt. Um allerdings die Notation in diesem Abschnitt überschaubar zu halten, wird hier nur der Fall zweier Regime betrachtet. Das zweite Element, mit dem eine

Markov-Kette dann eindeutig bestimmt ist, ist ihre Startverteilung, d.h.

$$\hat{\xi}_{1|0} = \begin{pmatrix} \mathcal{P}(S_1 = 1|y_0) \\ \mathcal{P}(S_1 = 2|y_0) \end{pmatrix}. \quad (7.4)$$

Beachtet man die Beobachtungshistorie in (7.1), so existiert kein beobachtbarer Wert y_0 . Daher werden sowohl die Einträge der Übergangsmatrix als auch die Startverteilung als gleichverteilt gewählt, wenn der Maximierungsalgorithmus initiiert wird. Dies wird später relevant, wenn man entscheiden muss, welche Regime Rezessionen und welche Expansionen zuzuordnen sind. Dasselbe gilt für die Einträge von θ . Die Annahme gegebener Normalverteilung liefert den folgenden Dichtevektor:

$$\eta_t = \begin{pmatrix} f(y_t|S_t = 1, z_t, \theta) = \frac{1}{\sqrt{2\pi\sigma_1}} \exp\left(-\left(\frac{y_t - z_t'\beta^1}{2\sigma_1}\right)^2\right) \\ f(y_t|S_t = 2, z_t, \theta) = \frac{1}{\sqrt{2\pi\sigma_2}} \exp\left(-\left(\frac{y_t - z_t'\beta^2}{2\sigma_2}\right)^2\right) \end{pmatrix}. \quad (7.5)$$

Wendet man den Satz totaler Wahrscheinlichkeit an und bezeichnet mit \odot eine elementweise Multiplikation, so erhält man das folgende Resultat von Hamilton (1994):

$$\hat{\xi}_{t|t} = \frac{\hat{\xi}_{t|t-1} \odot \eta_t}{\mathbf{1}'(\hat{\xi}_{t|t-1} \odot \eta_t)}, \quad (7.6)$$

wobei dies für den Beginn der Zeitreihe nichts anderes bedeutet als dass gefilterte Wahrscheinlichkeiten

$$\hat{\xi}_{1|1} = \begin{pmatrix} \mathcal{P}(S_1 = 1|y_1, \theta) \\ \mathcal{P}(S_1 = 2|y_1, \theta) \end{pmatrix} \quad (7.7)$$

berechnet wurden. Wichtig für dieses Kapitel ist, dass immer wenn θ in den verschiedenen Spezifikationen neu geschätzt wird, so werden sich auch $\hat{\xi}_{1|1}$ und allgemein alle gefilterten Wahrscheinlichkeiten verändern. Durch die Äquivalenz von $S_t = i \Leftrightarrow \xi_t = e_i$, wobei letzteres den i -ten Einheitsvektor darstellt, lässt sich eine vorläufige Zustandswahrscheinlichkeit für die zweite Beobachtung berechnen durch

$$\hat{\xi}_{2|1} = \hat{P}\hat{\xi}_{1|1}, \quad (7.8)$$

wobei P die Übergangsmatrix bezeichnet. Wiederholt man das Verfahren rekursiv und wei-

tet man es auf die folgenden von Kim (1994) eingeführten Wahrscheinlichkeiten aus,

$$\hat{\xi}_{t|T} = \hat{\xi}_{t|t} \odot \left(P' \left(\hat{\xi}_{t+1|T} \odot \hat{\xi}_{t+1|t} \right) \right), \quad (7.9)$$

so erhält man schließlich (geglättete) Zustandswahrscheinlichkeiten für alle Beobachtungen der Zeitreihe. Was verbleibt ist eine Konkretisierung des Maximierungsprozesses. Durch Umformulierung des üblichen Maximum-Likelihood-Ansatzes entwickelte Hamilton (1994) für Gleichungen der Form (7.2) die folgende Zielfunktion:

$$\sum_{t=1}^T \sum_{i=1}^{r-2} \mathcal{P}(S_t = i | Z_T) \log f(Y_t | S_t = i, Z_t, \theta) \quad (7.10)$$

Schreibt man die notwendige Bedingung der Maximierung in Vektornotation, so erhält man

$$\sum_{t=1}^T \left(\frac{\partial \log \eta_t}{\partial \theta'} \right)' \hat{\xi}_{t|T} = 0. \quad (7.11)$$

Zur Vereinfachung beschränke man θ' auf $(\beta_0^{S_t}, \beta_{1,y}^{S_t}, \beta_{1,x}^{S_t}, \sigma^{S_t})$. Dann ist (7.11) äquivalent zu folgendem nicht-linearen Gleichungssystem:

$$\begin{aligned} \sum_{t=1}^T \mathcal{P}(S_t = 1 | z_T) \frac{\partial \log \eta_t^1}{\partial \beta_0^1} + \mathcal{P}(S_t = 2 | z_T) \frac{\partial \log \eta_t^2}{\partial \beta_0^2} &= 0 \\ \sum_{t=1}^T \mathcal{P}(S_t = 1 | z_T) \frac{\partial \log \eta_t^1}{\partial \beta_{1,y}^1} + \mathcal{P}(S_t = 2 | z_T) \frac{\partial \log \eta_t^2}{\partial \beta_{1,y}^2} &= 0 \\ \sum_{t=1}^T \mathcal{P}(S_t = 1 | z_T) \frac{\partial \log \eta_t^1}{\partial \beta_{1,x}^1} + \mathcal{P}(S_t = 2 | z_T) \frac{\partial \log \eta_t^2}{\partial \beta_{1,x}^2} &= 0 \\ \sum_{t=1}^T \mathcal{P}(S_t = 1 | z_T) \frac{\partial \log \eta_t^1}{\partial \sigma^1} + \mathcal{P}(S_t = 2 | z_T) \frac{\partial \log \eta_t^2}{\partial \sigma^2} &= 0 \end{aligned} \quad (7.12)$$

Die Nichtlinearität entsteht aus der Ableitung nach σ^{S_t} . Perlin (2009) nutzt das Gradientenverfahren, um $\hat{\theta}$ als Lösung von (7.12) zu erhalten. Außerdem leitet Hamilton (1990) Schätzer für die Übergangswahrscheinlichkeiten her. Diese sind

$$\hat{p}_{ij} = \frac{\sum_{t=2}^T \mathcal{P}(S_t = j, S_{t-1} | z_T, \hat{\theta})}{\sum_{t=2}^T \mathcal{P}(S_{t-1} = i | z_T, \hat{\theta})}. \quad (7.13)$$

Mit $\hat{\theta}$, $\hat{\xi}_{1,T}$ und \hat{p}_{ij} , welche die zufällig gewählten Startwerte ersetzen, erreicht der Algorithmus seine zweite Runde. Hier werden erneut Zustandswahrscheinlichkeiten berechnet und der gesamte Prozess wird iteriert bis zur Konvergenz der Likelihood-Funktion. Das Verfahren wird hier so detailliert beschrieben, um die Tatsache zu betonen, dass sich im Fall von mehreren Schätzungen mit unterschiedlichen erklärenden Variablen auch das nicht-lineare Gleichungssystem (7.12) verändert. So ist es der Doppeleffekt aus unterschiedlichen Einträgen in z'_t und in β^{S_t} , am besten sichtbar in der bedingten Erwartung in (7.5), der während des gesamten Maximierungsprozesses auftritt und schließlich zu unterschiedlichen Zustandswahrscheinlichkeiten führt.

Auf den ersten Blick mag eine solche Feststellung trivial erscheinen, aber die Auswirkungen werden klar, wenn man beachtet, dass das hier vorgestellte Modell nicht nur aus einer Gleichung vom Typ (7.1) besteht, sondern aus mehreren. Dabei ist die Konstruktion in Anlehnung an Timmermann (2006) gewählt, so dass jede der Gleichungen einer möglichen „Fehlspezifizierung unbekannter Form“ unterliegen mag. Wie zuvor erwähnt, kann der Grund für eine solche Verzerrung die Begrenzung des Parameterraums bei gegebener Echtzeitdatenlage sein. Im Folgenden wird jede Gleichung durch den darin enthaltenen exogenen Regressor repräsentiert. Innerhalb dieses Aufbaus kann dann die Vorteilhaftigkeit einer Durchschnittsbildung über die Einzelvorhersagen gezeigt werden. Da wir möglichst gute Vorhersagen treffen wollen, wird die exogene Variable ein (potenzieller) Frühindikator des Konjunkturzyklus sein.

7.3 Modellspezifizierung und In-sample-Ergebnisse

7.3.1 Datenauswahl

Die Datenbanken der Deutschen Bundesbank (Deutsche Bundesbank 2011, 2010a,b) bilden zusammen mit dem Geschäftsklimaindex des ifo-Instituts (ifo Institut für Wirtschaftsforschung 2011) die Datengrundlage dieses Kapitels. Bei der Bearbeitung des Forschungsauftrags werden Daten monatlicher Frequenz genutzt, um eine Aussage über den Konjunkturzyklus zwischen den Quartalsveröffentlichungen des Bruttoinlandsprodukts treffen zu können. Als Referenzzeitreihe dient die schon in den Abschnitten 2.2 und 4.3.1 vorgestellte Produktion im Produzierenden Gewerbe ohne Bau.

Betrachtet man die Wachstumsraten der monatlichen Referenzzeitreihe in Abbildung 7.1, so führt die hohe Volatilität sogar in Rezessionen zu mehreren (nicht hintereinander

liegenden) Monaten stark positiver Werte; genauso lassen sich in Expansionen mehrere (nicht hintereinander liegende) Monate stark negativer Werte finden, bspw. $\pm 2\%$. Wie sich herausstellt, kann diese Art kurzfristiger gegenläufiger Dynamik nicht durch die autoregressiven Terme aufgefangen werden. Das führt dazu, dass sich die Regime des MS-Modells inkonsistent zu den Konjunkturphasen verhalten. In diesem Fall scheinen die gefilterten Regime vielmehr die Asymmetrie zwischen positiven und negativen Wachstumsraten gewisser Größenordnung darzustellen. Aber wie in der NBER-Definition erwähnt entspricht dies nicht automatisch Rezessionen und Expansionen, da die Dynamik in der Umgebung, insbesondere die Dauer von mehr als ein paar Monaten und die Signifikanz der kumulativen Bewegung während dieser Zeit beachtet werden muss.

Eine Möglichkeit, mit diesem Problem umzugehen, ist das Rückwärtsglätten der Referenzzeitreihe durch einen laufenden Durchschnitt vom Grad 3. Verfährt man in dieser Weise, so erhält man dann tatsächlich wieder die Koinzidenz zwischen iterierten Regimen und Konjunkturphasen, wie sie schon von Hamilton (1989) gefunden wurde. (Sein Modell beruhte auf Quartalsdaten.) Trotz der notwendigen Glättung ist ein solches Vorgehen nicht nutzlos, da es Aussagen zwischen den Quartalsveröffentlichungen ermöglicht.

Wie beispielsweise Krolzig (1997) diskutiert, gehört das MS eindeutig zur Klasse nichtlinearer Modelle. Daher wurde die Referenz-Zeitreihe zunächst auf Nichtlinearität getestet, bevor MS-Modelle geschätzt wurden. Ein oft verwendeter Test auf Nichtlinearität ist derjenige von Brock et al. (1996). Dabei wird analysiert, ob die Residuen einer ARMA-Schätzung die Hypothese

$$H_0 : \text{Die Zeitreihe ist i.i.d.}$$

einhalten. Wie Tabelle 7.1 zeigt, kann H_0 auf jedem regulären Signifikanzniveau abgelehnt werden. Dies bestätigt die Anwendung einer nichtlinearen Schätzmethode.

Als nächstes stellen wir die verwendeten Frühindikatoren vor. Diese sind aus- und inländische Auftragseingänge, Baugenehmigungen im gewerblichen Hochbau, der CDAX-Kursindex, ein Spread zwischen der Umlaufrendite privater und öffentlicher Emissionen (*Corporate spread*), der 3-Monats-Euribor-Zinssatz, der ifo-Geschäftsklimaindex, Kreditwachstumsraten aus dem ESZB-System, der Spread zwischen 10jährigen Bundesanleihen und dem 3-Monats-Euribor (*term spread*) sowie offene Stellen. Abgesehen vom Corporate spread, den Kreditwachstumsraten und offenen Stellen, für die die Einleitung relevante Literatur enthält, werden ähnliche Zeitreihen zur Aggregation im amerikanischen *Composite Index of Leading Indicators* (CLI) herangezogen mit dem Ziel, konjunkturelle Wende-

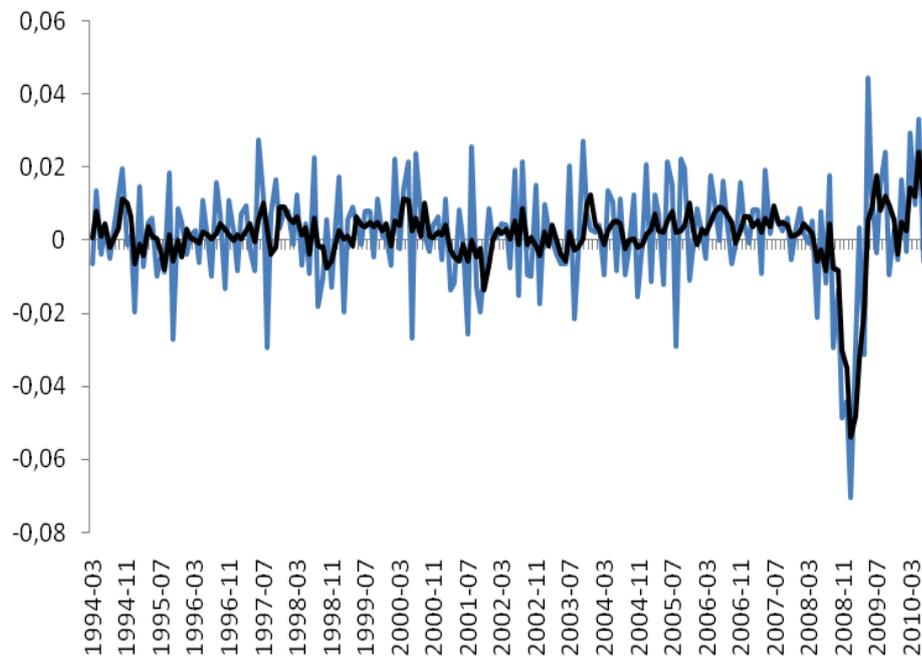


Abbildung 7.1: Wachstumsraten rückwärts geglätteter Industrieproduktion. Die blaue Linie zeigt die Wachstumsraten der monatlichen Referenz-Zeitreihe für die Veröffentlichung November 2010 und die Beobachtungen (Berichtsmonate) von März 1994 bis September 2010. Diese führt zur Inkonsistenz zwischen MS-Regimen und Konjunkturphasen, was sich bei Rückwärtsglättung ändert (schwarze Linie).

punkte zu antizipieren.

Der Vorlauf der meisten Indikatoren ist offensichtlich, da sie Vorstufen des Produktionsprozesses, wie etwa Auftragseingänge oder Baugenehmigungen, oder Erwartungen in Bezug auf die künftige wirtschaftliche Lage widerspiegeln, wie Geschäftsklimaindex oder offene Stellen. Wie sich herausgestellt hat, liefern letztere allerdings abgeschwächte Ergebnisse in Bezug auf die Signifikanz ihrer Koeffizienten. Nichtsdestotrotz erscheint es sinnvoll, mindestens eine Variable des Arbeitsmarkts ins Modell zu integrieren.

Wann immer eine Rezession heraufzieht, sollte dies zu höheren Ausfallraten von Unternehmen führen, während Bundesanleihen als vergleichsweise sichere Anlage stärker nachgefragt würden. Daher kann ein Anstieg des Unternehmensrenditeaufschlags als Rezessionssignal gedeutet werden, während dieser Aufschlag im Prinzip bis zum Ende der Rezession wieder auf sein Ausgangsniveau zurückkehren sollte. Kurzfristige Zinsen reagieren auf die aktuelle wirtschaftliche Lage wesentlich sensibler als langfristige, so dass

Tabelle 7.1: BDS-Test für die Wachstumsraten der geglätteten Industrieproduktion

Dimension	BDS Statistic	Std. Error	z-Statistic	Normal Prob.	Bootstrap Prob.
2	0.027095	0.005469	4.954080	0.0000	0.0000
3	0.056127	0.008721	6.435694	0.0000	0.0000
4	0.067283	0.010419	6.457709	0.0000	0.0000
5	0.069270	0.010894	6.358438	0.0000	0.0000
6	0.062452	0.010539	5.925663	0.0000	0.0000
Raw epsilon		0.006371			
Pairs within epsilon		27118.00	V-Statistic	0.705904	
Triples within epsilon		4039498.	V-Statistic	0.536487	
Dimension	C(m,n)	c(m,n)	C(1,n-(m-1))	c(1,n-(m-1))	c(1,n-(m-1))^k
2	9877.000	0.522178	13309.00	0.703621	0.495083
3	7530.000	0.402222	13144.00	0.702099	0.346095
4	5726.000	0.309046	12992.00	0.701209	0.241763
5	4417.000	0.240892	12889.00	0.702934	0.171622
6	3420.000	0.188482	12848.00	0.708074	0.126029

Anmerkungen: Das Testresultat mit einer klaren Ablehnung der i.i.d-Hypothese bestätigt die Verwendung eines nichtlinearen Modells. Berechnet man ε/σ , so liegt dies innerhalb von $[0,5;2]$. Dieses Intervall sowie eine maximale eingebettete Dimension von $m = 5$ repräsentieren die relevanten Bereiche von i.i.d-Zeitreihen, wie Brock et al. (1996) darlegen. Da N/m keine für die Approximation mit Sicherheit ausreichende Größenordnung erreicht, werden zusätzlich p-Werte per Bootstrap berechnet.

die Zinsdifferenziale zu Rezessionsbeginn rückläufig sein sollten. In Extremfall führt dies sogar zu einer inversen Zinskurve. Die dem Kreditwachstum zu Grunde liegenden Zusammenhänge zur Rezession werden an späterer Stelle beschrieben.

Im Allgemeinen unterliegen Finanz- und Umfragevariablen keiner Revision oder verzögerten Datenverfügbarkeit, im Gegensatz zu realwirtschaftlichen makroökonomischen Variablen. Für alle makroökonomischen Variablen, die für dieses Kapitel genutzt wurden, beträgt die Datenverfügbarkeitslücke zwei Monate (vgl. Kapitel 2) – mit Ausnahme der offenen Stellen, die unmittelbar zur Verfügung gestellt werden. Die einzige Finanzvariable, für die es etwas Zeit dauert, bis sie in der ESZB-Statistik enthalten ist, ist das Kreditwachstum.

7.3.2 Modellspezifizierung

Der gewöhnliche Trade-off zwischen einer Verbesserung des Modells durch Einbeziehung zusätzlicher signifikanter Regressoren und Verschlechterung durch Überspezifizierung wird besonders deutlich an MS-Modellen. Während in linearen Standardmodellen eine Gleichung mit nur zwei unterschiedlichen Regressoren sowie einer beschränkten Lagauswahl oft auf ein verzerrtes Ergebnis aufgrund ausgelassener Variablen deutet, ist dies für MS-Modelle nur dann klar, wenn jede der Zustandsgleichungen der Verzerrung unterliegt. Erweitert man aber die Gleichungen durch zusätzliche Variablen oder switchende Parameter zu stark, so läuft man Gefahr, dass der Optimierungsalgorithmus aus Abschnitt 7.2 nur lokale Maxima findet. Wie sich aus den in Tabelle 7.3 angegebenen Freiheitsgraden entnehmen lässt, versucht das hier vorgestellte Modell, die einzelnen Gleichungen nicht zu hoch parametrisiert zu gestalten. So kommen maximal vier switchende Koeffizienten pro Gleichung vor. Trotzdem stellt die limitierte Echtzeitdatenhistorie ein zentrales Problem dar.

Tabelle 7.2 zeigt die Gestalt jeder Gleichung, wobei die Lags anhand des Schwarzischen Informationskriteriums in einem *general-to-specific*-Ansatz gewählt wurden. Demnach können maximal 5 Lags einer Variable gewählt werden, sofern eine gewisse Mindestübereinstimmung mit den Zustandswahrscheinlichkeiten der in Abschnitt 4.3 beschriebenen Benchmark-Methode besteht. In einer früheren Version wurde die Lagauswahl zudem bei jeder Neuveröffentlichung der Daten durchgeführt. Da dies zu einer ausufernden Rechenzeit bei den Echtzeitschätzungen führte, wurde die in Tabelle 7.2 enthaltene Lagstruktur

schließlich fixiert.¹

Als natürlicher Vergleichsmaßstab für die Anwendung des MS auf den Konjunkturzyklus gilt die folgende Spezifizierung von Hamilton (1989):

$$y_t - \beta_{S_t} = \sum_{j=1}^4 \phi_j (y_{t-j} - \beta_{S_{t-j}}) + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2), \quad S_t = 1, 2.$$

Obwohl diese nicht benutzt wurde, um die Resultate des hier vorgestellten MS-Modells auszuwerten, ist sie hilfreich, um zu entscheiden, ob die Konstante regimeabhängig zu wählen ist oder nicht. Perlin (2009) zeigt, dass bei Schätzung der Gleichungen

$$y_t = \beta_{S_t} + \varepsilon_{a,t}, \quad \varepsilon_{a,t} \sim N(0, \sigma_a^2), \quad S_t = 1, 2$$

$$\varepsilon_{a,t} = \sum_{j=1}^4 \phi_j \varepsilon_{a,t-j} + \varepsilon_{b,t}, \quad \varepsilon_{b,t} \sim N(0, \sigma_b^2)$$

ähnliche Werte für die mit den Zustandswahrscheinlichkeiten verbundenen Schätzgrößen auftreten. Dies identifiziert die switchende Konstante als den wichtigsten Term für die Iteration der Zustandswahrscheinlichkeiten. Daher enthält jede der Gleichungen eine regimeabhängige Konstante. Wie schon zuvor erwähnt, treten allerdings in jeder Gleichung mehr switchende Koeffizienten auf. Gründe dafür sind wie folgt: Erstens erlaubt eine switchende Varianz des Fehlerterms die Anwendung des Welch-Tests, um verschiedene Normalverteilungen zu identifizieren, wenn man zu einem Design mit 4 Regimen wechselt. Zweitens ermöglicht zumindest ein switchender Koeffizient, welcher der Regressorvariable des vorlaufenden Indikators zugeordnet ist, die Veränderung zwischen Regimen zu messen, welche für unterschiedliche Intensitäten derselben Konjunkturphase stehen (starke und schwache Rezession bzw. starke und schwache Expansion). Siehe hierzu Tabelle 7.6 in Abschnitt 7.4.1.

Um die Ergebnisse des MS-Modells auswerten zu können, wurde der in Abschnitt 4.3 vorgestellte Datierungsalgorithmus als Benchmark-Methode verwendet, welcher auf der Arbeit von Bry & Boschan (1971) sowie Harding & Pagan (2002) basiert. Aufgrund des hierbei notwendigen Aufdeckens lokaler Extrema kann dieser Datierungsalgorithmus aus-

¹Vorab basiert die Lagauswahl auf der Veröffentlichung Januar 2007, d.h. der letzten Datenlage vor Beginn der Echtzeitvorhersagen in Abschnitt 7.4.2. Allerdings ist zu diesem Zeitpunkt eine höhere Anzahl von Regimen nur hypothetisch. Daher überprüfen die in Abschnitt 7.3.3 dargestellten Ergebnisse, ob die Lagstruktur für eine Veröffentlichung auch dann standhält, nachdem die Echtzeitanpassung des Modells bezüglich der Anzahl der Regime tatsächlich stattgefunden hat.

schließlich zur Expostdatierung genutzt werden; genauer gesagt müssen fünf Monate sowie die Datenverfügbarkeitslücke abgelaufen sein. Dies muss beachtet werden, wenn in den Abbildungen 7.2 und 7.4 MS-Ergebnisse mit der Benchmark verglichen werden.

7.3.3 In-sample-Evaluation

Tabelle 7.2 und 7.4 sowie Abbildung 7.2 präsentieren die In-sample-Ergebnisse für die verschiedenen MS-Spezifikationen. Die hierfür notwendige Ausgestaltung der Gleichungen wurde in Abschnitt 7.3.2 entwickelt. Tabelle 7.2 bezieht sich auf das 4-Regime-Modell und enthält p-Werte für jeden der Parameter in Klammern. Standardfehler wurden mit der von Perlin (2009) implementierten Methode so berechnet, dass sie robust sind gegen Heteroskedastizität und Autokorrelation gemäß dem Vorgehen von Newey & West (1987). Die einzige systematische Insignifikanz besteht in der Ablehnung der Parameter des zweiten Regimes innerhalb der Spezifikation der unbesetzten Stellen. Das bedeutet, dass in diesem Fall nicht ausgeschlossen werden kann, dass die dem Regime zugeordneten Koeffizienten und damit auch der entsprechende Erklärungsbeitrag identisch Null sind. Um das gesamte System konsistent zu gestalten und eine Arbeitsmarktvariable im Modell zu behalten, wurde dieses Ergebnis jedoch toleriert. Wie immer bei Echtzeitschätzungen ist zu beachten, dass die angegebenen Werte bezogen sind auf eine Veröffentlichung, nämlich hier diejenige von November 2010. Mit jeder davor und danach gelegenen Neuschätzung mag sich zwar auch das Signifikanzergebnis leicht verändern, aber es würde den Rahmen einer solchen Untersuchung sprengen, jede der Publikationen zu berücksichtigen.

Im Allgemeinen würde man einen Anstieg des Kreditflusses (definiert z.B. als Neugeschäft abzüglich Tilgung) mit einer Verbesserung der wirtschaftlichen Lage verbinden, da dieser einen besseren Zugang für die Investitionstätigkeit impliziert. Demgegenüber beschreiben Biggs et al. (2009), „dass die Erholung nach der Finanzmarktkrise ohne eine entsprechende Wiederbelebung der Kreditmenge“ ablief. Allerdings schließen sie diesen Zusammenhang nicht für den Kreditfluss aus. Betrachtet man jedoch die Kreditwachstumsraten, wie sie anhand der Bilanzangaben monetärer Finanzinstitute in die ESZB-Statistiken eingepflegt werden, so ergibt sich ein völlig anderes Bild. So scheinen die Mitten der letzten Rezessionen, wie sie durch das Benchmark-Modell berichtet werden, in etwa mit den Hochpunkten des Kreditwachstums zusammen zu fallen. Diese Kohärenz wird evident, wenn man die Kredite an andere monetäre Finanzinstitute miteinbezieht. Man könnte in diesem Zusammenhang auch von einem ungesunden Kreditwachstum oder einer Kredit-

blase sprechen.

Angesichts der Nichtstationarität der Referenz-Zeitreihe, vgl. Levanon (2010), basiert das hier vorgestellte MS-Modell auf Wachstumsraten, was bedeutet, dass die Wachstumsrate des Kreditwachstums als Regressorvariable genutzt wird. Lässt man eine solche Spezifikation laufen, so erhält man nur dann überwiegende Signifikanz (Tabelle 7.2), wenn man auch die Interbankengeschäfte berücksichtigt. Dennoch soll der hier vorgestellte Zusammenhang nicht überbewertet werden. Zum einen reicht die Beobachtungshistorie für die Kreditwachstumsraten nur bis zum Beginn des ESZB-Systems, d.h. bis Ende 1998. Zum anderen muss erwähnt werden, dass die meisten der Kreditderivate, die als mögliche Erklärung dienen könnten, in der Zeit vor der Finanzmarktkrise zumeist außerbilanziell geführt wurden.

Betrachtet man die Zustandswahrscheinlichkeiten in der Abbildung (7.2), so wird offenkundig, dass sich diese mit den verschiedenen MS-Spezifikationen ändern. Dies wurde auf theoretischer Basis auch in Abschnitt 7.2 diskutiert. Diesbezüglich entsteht auch der Eindruck, dass sich die Vorhersagekraft eines einzelnen Indikators je nach Rezession ändert. Jede der Spezifikationen identifiziert Teile der Rezession nach der Finanzmarktkrise als zusätzliches Regime. Daher muss im Folgenden erläutert werden, wie die verschiedenen Regime Rezessionen und Expansionen zugeordnet werden können.

Zunächst ist eine solche Zuordnung deswegen nahe liegend, weil einerseits Rezession und Expansion den Konjunkturzyklus bestimmen und andererseits im MS-Modell die Regime den Referenzzyklus (die Referenzzeitreihe) generieren. Technisch gesehen erfolgt die Zuordnung über die switchende Konstante, dies wird im Abschnitt über die Vorhersageergebnisse noch einmal ausgeführt werden. Da zu Beginn jeder Iteration gleichverteilte Startwerte stehen, ist es nicht notwendig, dass Rezessions- oder Expansionsregime das gleiche Label tragen, z.B. Regime 1. Daher ist es zulässig, das Regime als Krisenregime zu deklarieren, das zu der am deutlichsten negativen Wachstumsrate führt.

Eine interessante Frage, die im Zuge der verschiedenen MS-Spezifikationen auftritt, ist, ob vorlaufende Indikatoren ihre Vorhersagekraft an die zugehörige Spezifikation vererben können. Hinweise auf ein solches Verhalten finden sich. So kann die vermeintlich zusammenhängende Rezession zwischen Februar 2001 und September 2003 in den Spezifikationen des CDAX und des ifo-Geschäftsklimaindex genannt werden. Diese korrespondiert mit dem Verlauf der genannten Indikatoren. Ein anderes Beispiel wäre der frühere Beginn

7 Wendepunktbestimmung mit Markov-Switching-Modellen

Tabelle 7.2: Zusammenfassung der MS-Regressionen mit 4 Regimen – Konstante und switchender autoregressiver Term

Stichprobe: 1994:03-2010:09 Veröffentlichung: 2010:11	Switchende				Switchendes			
	Konstante				Endogenes Lag			
Autoregressiv	0,0101 (0,00)	0,0035 (0,00)	-0,0022 (0,05)	-0,0395 (0,00)	- (-)	- (-)	- (-)	- (-)
Ausländische Auftragseingänge	0,0054 (0,00)	-0,0001 (0,59)	-0,0203 (0,00)	-0,0341 (0,00)	-0,0919 (0,00)	-0,1320 (0,00)	-1,2808 (0,00)	0,3485 (0,00)
Inländische Auftragseingänge	0,0134 (0,00)	0,0032 (0,00)	-0,0018 (0,00)	-0,0326 (0,00)	- (-)	- (-)	- (-)	- (-)
Baugenehmigungen (Stichprobe 1994:07-2010:09)	0,0049 (0,00)	0,0048 (0,00)	-0,0015 (0,00)	-0,0397 (0,00)	-0,0099 (0,00)	0,4554 (0,00)	-0,1212 (0,00)	-0,0690 (0,00)
CDAX	0,0059 (0,00)	-0,0005 (0,03)	-0,0097 (0,00)	-0,0311 (0,00)	-0,1006 (0,00)	-0,0944 (0,00)	0,2203 (0,00)	0,9772 (0,00)
Corporate Spread	0,0097 (0,00)	0,0034 (0,00)	-0,0034 (0,12)	-0,0538 (0,00)	-0,0574 (0,00)	-0,0894 (0,00)	-0,1222 (0,00)	-0,4204 (0,00)
Euribor - 3M	0,0117 (0,00)	0,0035 (0,00)	-0,0025 (0,00)	-0,0427 (0,00)	- (-)	- (-)	- (-)	- (-)
ifo-Geschäftsklimaindex	0,0097 (0,00)	0,0085 (0,00)	0,0055 (0,00)	-0,0001 (0,00)	0,2738 (0,00)	1,0877 (0,00)	-0,0854 (0,04)	-0,1269 (0,00)
Kreditwachstum (Stichprobe 1999:03-2010:09)	0,0103 (0,00)	0,0047 (0,85)	-0,0025 (0,16)	-0,0431 (0,00)	-0,1169 (0,08)	-0,1221 (0,00)	-0,1279 (0,05)	-0,2321 (0,25)
Spread 10Y	0,0099 (0,00)	0,0030 (0,00)	-0,0020 (0,00)	-0,0270 (0,00)	- (-)	- (-)	- (-)	- (-)
Offene Stellen	0,0115 (0,01)	0,0028 (0,03)	-0,0024 (0,71)	-0,0162 (0,00)	-0,1226 (0,01)	-0,0953 (0,02)	-0,1751 (0,30)	-0,3057 (0,00)

Anmerkungen: Jede Gleichung wird durch den eingebetteten Frühindikator repräsentiert. Der erste Tabellenteil behandelt die Konstante und den switchenden autoregressiven Teil der Gleichungen, der zweite (siehe Tabelle 7.3) die weiteren Terme. Die Reihenfolge zwischen den Spalten, welche switchende Koeffizienten enthalten, richtet sich nach der Größenordnung der entsprechenden Konstante. Koeffizienten des Regimes mit der größten Konstante stehen links, diejenigen mit der kleinsten stehen rechts. So erklärt sich die Struktur der Spalten *Switchende Konstante*, *Switchendes Endogenes Lag* und *Switchendes Exogenes Lag*. P-Werte für jeden der Koeffizienten finden sich in Klammern.

7.3 Modellspezifizierung und In-sample-Ergebnisse

Tabelle 7.3: Zusammenfassung der MS-Regressionen mit 4 Regimen – Restterme

Stichprobe: 1994:03-2010:09 Veröffentlichung: 2010:11	Endogen ohne Switch		Switchendes Exogenes Lag			Exogen ohne Switch		Freiheits- grade
Autoregressiv	-0,1128 (0,00)	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)	162
Ausländische Auftragseingänge	- (-)	0,0152 (0,00)	0,0446 (0,00)	0,7366 (0,00)	-0,0778 (0,00)	0,0320 (0,00)	- (-)	154
Inländische Auftragseingänge	- (-)	0,1592 (0,00)	0,0520 (0,00)	0,0508 (0,04)	0,0531 (0,00)	0,0360 (0,01)	- (-)	158
Baugenehmigungen (Stichprobe 1994:07-2010:09)	-0,0171 (0,18)	-0,0025 (0,45)	0,1594 (0,00)	-0,0028 (0,10)	-0,2093 (0,00)	-0,0023 (0,45)	-0,0038 (0,08)	148
CDAX	- (-)	-0,0071 (0,00)	-0,0047 (0,00)	0,1356 (0,00)	0,2885 (0,00)	-0,0045 (0,00)	- (-)	154
Corporate Spread	- (-)	0,0414 (0,00)	0,0176 (0,04)	-0,1345 (0,00)	0,1974 (0,00)	- (-)	- (-)	155
Euribor - 3M	- (-)	0,6540 (0,00)	0,0429 (0,00)	-0,1018 (0,00)	1,3878 (0,00)	- (-)	- (-)	159
ifo-Geschäftsklimaindex	- (-)	1,5692 (0,00)	0,2059 (0,00)	0,0883 (0,00)	0,1032 (0,00)	- (-)	- (-)	155
Kreditwachstum (Stichprobe 1999:03-2010:09)	- (-)	0,0253 (0,02)	0,0085 (0,07)	0,0116 (0,28)	0,0587 (0,34)	0,0011 (0,70)	-0,0097 (0,00)	93
Spread 10Y	- (-)	-0,0432 (0,16)	0,0178 (0,17)	0,0061 (0,24)	-0,2036 (0,00)	- (-)	- (-)	159
Offene Stellen	- (-)	0,0682 (0,00)	0,0201 (0,19)	0,0399 (0,70)	0,8619 (0,00)	0,0481 (0,03)	0,0188 (0,12)	153

Anmerkungen: Fortsetzung von Tabelle 7.2, siehe die dortigen Erläuterungen. Da nur das aktuellste Lag jeder Art von Regressor switchen kann, werden die verbleibenden Lags der Auswahl in Abschnitt 7.3.2 in den Spalten *Endogen ohne Switch* und *Exogen ohne Switch* der Aktualität nach von links nach rechts aufgeführt. Die Modellspezifizierung aus Abschnitt 7.3.2 führt zu leicht unterschiedlichen Lagstrukturen. So schwanken die Freiheitsgrade zwischen 153 und 162 – mit der Ausnahme von Baugenehmigungen und Kreditwachstum, für die nur kürzere Beobachtungshistorien zur Verfügung stehen. P-Werte für jeden der Koeffizienten finden sich in Klammern.

der letzten Rezession, welcher in der Spezifikation ausländischer Auftragseingänge den Verlauf des Indikators reflektiert.

Trotzdem sind auch Gegenbeispiele zu nennen. So passt der Verlauf des Renditeaufschlags der Unternehmensanleihen und der offenen Stellen sehr gut zur Abwärtsbewegung von August 2002 zu September 2003, während die entsprechenden Spezifikationen keine Rezession vermelden. Aber selbst wenn eine solche intuitive Beziehung zwischen Indikator und Spezifikation nicht gegeben wäre, so dass Verbesserungseffekte aus der Durchschnittsbildung nur aufgrund von weißem Rauschen erzielt würden, gäbe es nichts Näherliegenderes, als vorlaufende Indikatoren für diesen Zweck zu verwenden.

Tabelle 7.4 fasst die Anzahl korrekt erkannter Rezessions- und Expansionsphasen für jede Spezifikation zusammen. Wie in Abschnitt 7.3.2 erläutert, kann die autoregressive Schätzung größtenteils mit dem Hamilton-Modell verglichen werden. Es finden sich jedoch sowohl Spezifikationen mit einer höheren Anzahl korrekt erkannter Rezessionen (z.B. CDAX) als auch mit einer höheren Anzahl korrekt erkannter Expansionen (z.B. corporate spread). Schließlich existiert auch eine Spezifikation mit höherer Gesamtkorrektheit (inländische Auftragseingänge). Natürlich besteht dabei ein Trade-off zwischen der korrekten Anzahl von Rezessionen und von Expansionen. Anhand der Ergebnisse von Tabelle 7.4 erscheint es reizvoll, für die Vorhersage nach Verbesserungseffekten aus der Durchschnittsbildung zu suchen.

Zusätzlich zur Korrektheit der Regimewahrscheinlichkeiten enthält Tabelle 7.4 Standardmaße, welche in Verbindung zur Anpassung an die Wachstumsraten der Referenz-Zeitreihe stehen. Diese sind das bereinigte R^2 und der (die Wurzel des) mittlere quadratischen Fehler. Nachdem Regimewahrscheinlichkeiten gefiltert und die regimeabhängigen Koeffizienten geschätzt sind, können die Anpassungswerte aus den einzelnen Zustandsgleichungen berechnet werden. Dafür wird jede Zustandsgleichung mit der entsprechenden Regimewahrscheinlichkeit gewichtet.

Mit Ausnahme der das ifo-Geschäftsklima enthaltenden Spezifikation erreicht jede der übrigen Gleichungen für das bereinigte R^2 Werte von über 70% und übertrifft in dieser Hinsicht eine einfache AR(1)-Benchmark. Der niedrige Wert in einer einzelnen Spezifikation mag darauf zurück zu führen sein, dass der in Abschnitt 7.2 beschriebene Algorithmus hier nur ein lokales Maximum findet. Dies könnte für die Veröffentlichung November 2010 bedeuten, dass eine der elf Regressionen einer Fehlspezifizierung unterliegt. Eine Möglichkeit damit umzugehen, besteht gerade in der Durchschnittsbildung über die Einzelvorhersagen. Timmermann (2006) folgert für solche Fälle, in denen die Ursache der Verzerrung

7.3 Modellspezifizierung und In-sample-Ergebnisse

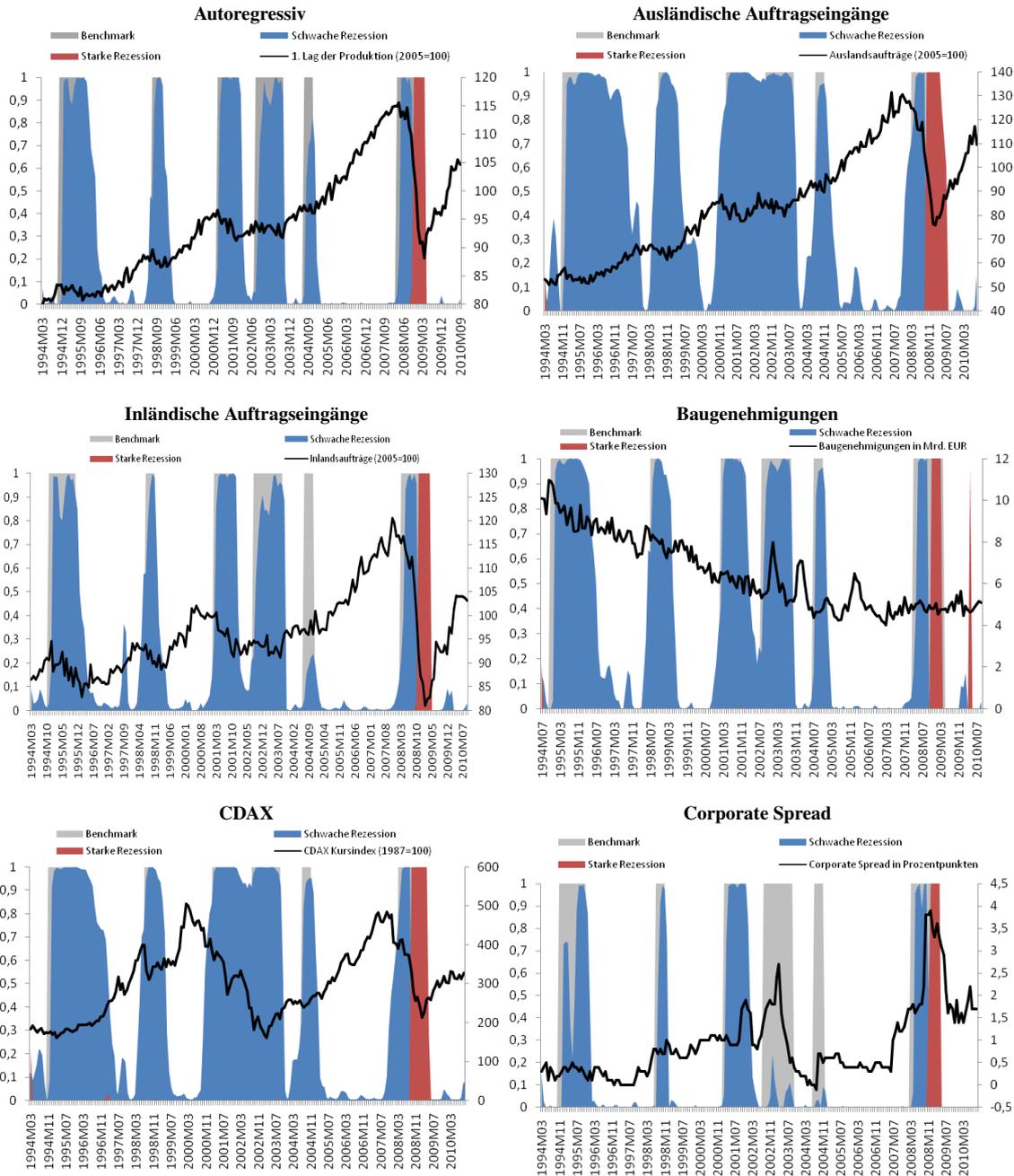


Abbildung 7.2: Rezessionswahrscheinlichkeiten der MS-Spezifikationen 1-12 (linke Achse) und Verlauf der eingebetteten Frühindikatoren (rechte Achse); erster Teil, Fortsetzung siehe Abbildung 7.3. Während aus Illustrationszwecken letztere in Absolutwerten angegeben sind, wird jeweils die Wachstumsrate der Referenzzeitreihe auf die Wachstumsrate der Frühindikatoren regressiert. Alle verwendeten Daten sind kalender- und saisonbereinigt. Schwache und starke Expansion sind hier weiß gefärbt, aber in den meisten Spezifikationen fällt letztere mit der Erholung nach der Finanzmarktkrise zusammen.

7 Wendepunktbestimmung mit Markov-Switching-Modellen

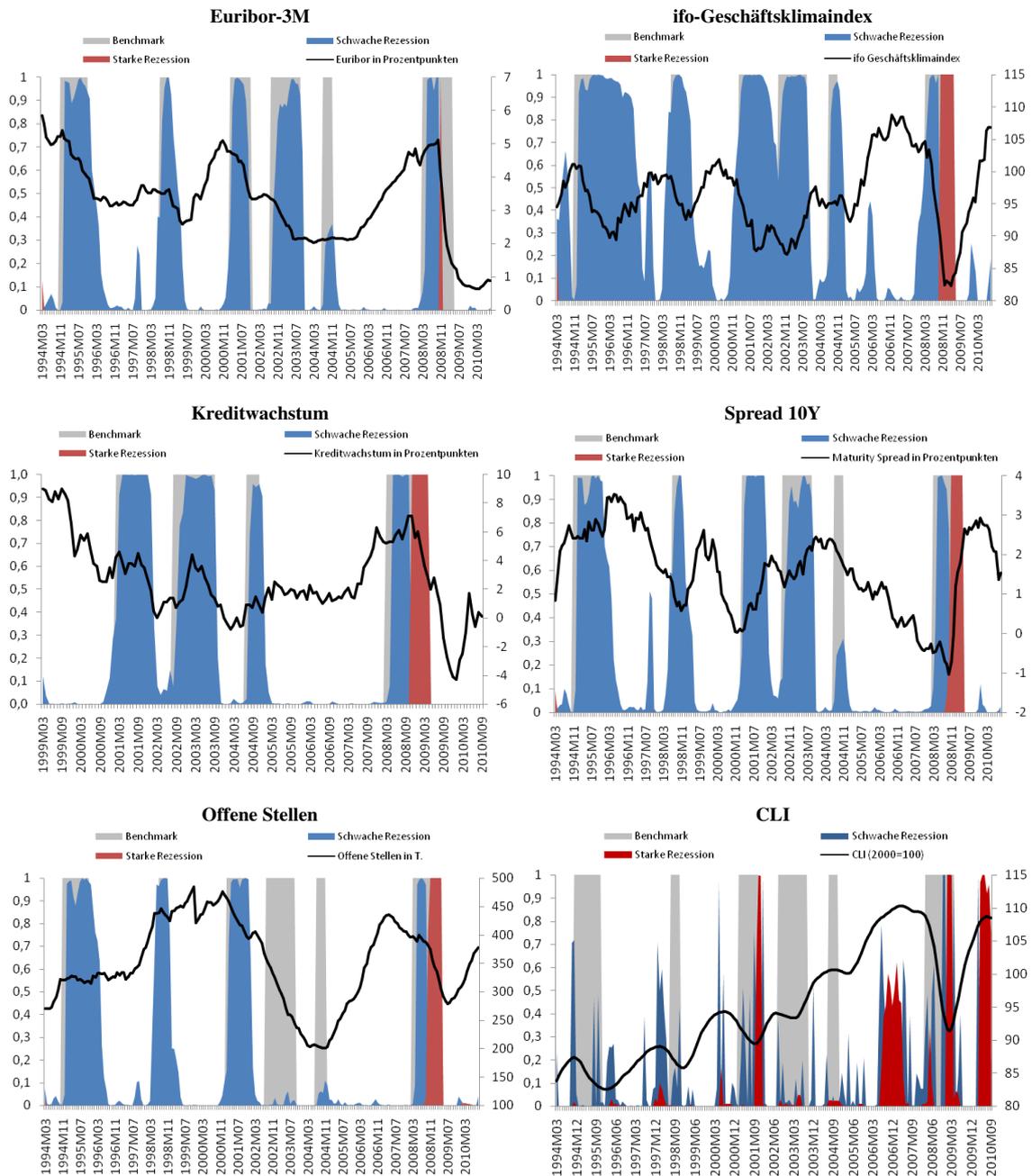


Abbildung 7.3: Fortsetzung von Abbildung 7.2, siehe auch Erläuterungen dort. Eine Spezifikation mit deutschen CLI-Daten der OECD (2010) wurde aufgrund der hier nicht persistenten Regime verworfen (letzte Teilabbildung).

Tabelle 7.4: Güte der Anpassung verschiedener MS Spezifikationen

Stichprobe: 1994:03 – 2010:09	Rezession		Expansion		Total		\bar{R}^2	RMSE	SIC	LR_T
	#	%	#	%	#	%				
Autoregressiv	52	85,25	127	92,03	179	89,95	0,8234	0,0033	-1467,39	260,54
Ausländische Auftragseingänge	59	96,72	89	44,72	148	74,37	0,7664	0,0037	-1440,76	228,32
Inländische Auftragseingänge	50	81,97	133	96,38	183	91,96	0,7435	0,0039	-1437,91	187,09
Baugenehmigungen	52	85,25	115	85,82	167	85,64	0,8131	0,0033	-1369,53	228,73
CDAX	58	95,08	98	71,01	156	78,39	0,7825	0,0035	-1440,37	250,77
Corporate Spread	33	54,10	131	94,93	164	82,41	0,7803	0,0036	-1438,46	265,84
Euribor - 3M	42	68,85	127	92,03	169	84,92	0,7528	0,0038	-1446,56	192,57
ifo-Geschäftsklimaindex	58	95,08	91	65,94	149	74,87	0,0980	0,0072	-1447,86	247,49
Kreditwachstum	37	86,05	90	93,75	127	91,37	0,8023	0,0037	-957,64	219,76
Spread 10Y	49	80,33	124	89,86	173	86,93	0,8409	0,0031	-1432,48	231,48
Offene Stellen	37	60,66	125	90,58	162	81,41	0,7944	0,0034	-1452,31	268,24
AR(1)	-	-	-	-	-	-	0,5916	0,0055	-	-

Anmerkungen: Die ersten sechs Spalten stehen für den Vergleich der iterierten MS-Regimewahrscheinlichkeiten mit den Ergebnissen der Benchmark-Methode. Das bereinigte R^2 sowie der mittlere quadratische Fehler (RMSE) werden mit Bezug auf die Anpassung an die Wachstumsraten berechnet. Hierfür werden ebenfalls AR(1) Vergleichswerte angegeben. Das Informationskriterium und die Likelihood-Ratio-Teststatistik stehen in Verbindung zur Loglikelihood-Funktion in Abschnitt 7.2, Gleichung (7.10). Um letztere zu berechnen, wurde für jede MS-Spezifikation eine lineare Version ohne switchende Koeffizienten geschätzt, vgl. (7.14).

nicht eindeutig zu Tage tritt, dass einfache Kombinationen oft bessere Resultate bringen; dies rechtfertigt insbesondere die benutzte Gleichgewichtung.

Ein ähnlicher Eindruck entsteht nach Berechnung des quadratischen Fehlers, für die man in Gleichung (7.20) einfach die Wahrscheinlichkeitsvorhersage durch die Anpassung an die Wachstumsrate ersetze. Erneut ist es die Regression mit dem ifo-Geschäftsklimaindex, welche einen schlechteren Wert als die übrigen liefert. Schließlich werden in Tabelle 7.4 auch Maße aufgeführt, die in Verbindung zur Loglikelihood-Funktion der Einzelschätzungen stehen. Hierbei sticht keines der Ergebnisse hervor. Weitere Details für die Interpretation der Resultate werden in den nächsten Abschnitten angegeben.

7.3.4 Likelihood-Ratio-Test in einem MS-Modell mit 4 Regimen

Die letzte Spalte aus Tabelle 7.4 enthält die Teststatistik für den Likelihood-Ratio-Test zwischen dem unbeschränkten MS-Modell und den linear beschränkten Versionen für jede der Gleichungen, d.h.

$$LR_T = -2 \left(Q_T^{\text{unbeschränkt}}(\hat{\theta}, (\hat{p}_{ij})) - Q_T^{\text{beschränkt}}(\tilde{\theta}) \right). \quad (7.14)$$

Unbeschränktheit bedeutet also hier die Möglichkeit der Parameter, in Abhängigkeit der Datenlage zwischen den Regimen zu switchen. Des weiteren bezeichnet Q_T die Loglikelihoodfunktion und T steht für den Zeitindex der letzten Beobachtung der Veröffentlichung November 2010. Betrachtet man die Notation, so fällt auf, dass ein Teil der Parameter, nämlich die Einträge der Übergangsmatrix, beliebig unter der Nullhypothese eines wahren linearen Modells sind. Dies bezeichnet Hansen (1996) als "inference when a nuisance parameter is not identified under the null". Wählt man eine lineare und eindeutige Zerlegung switchender Parameter in Gleichung (7.1), d.h.

$$\begin{aligned} y_t = & (\alpha_0^0 + (\alpha_0^1 + \alpha_0^2 + \alpha_0^3) S_t) + (\alpha_{1,y}^0 + (\alpha_{1,y}^1 + \alpha_{1,y}^2 + \alpha_{1,y}^3) S_t) y_{t-f-D_y} \\ & + \sum_{j=2}^p \beta_{j,y} y_{t-j+1-f-D_y} + (\alpha_{1,x}^0 + (\alpha_{1,x}^1 + \alpha_{1,x}^2 + \alpha_{1,x}^3) S_t) x_{t-f-D_x} \\ & + \sum_{j=2}^q \beta_{j,x} x_{t-j+1-f-D_x} + (\sigma^0 + (\sigma^1 + \sigma^2 + \sigma^3) S_t) \varepsilon_t, \quad \varepsilon_t \sim N(0, 1), \quad t = 1, \dots, T, \end{aligned}$$

so kann die Nullhypothese wie folgt geschrieben werden als

$$H_0 : (\alpha_0^1, \alpha_0^2, \alpha_0^3, \alpha_{1,y}^1, \alpha_{1,y}^2, \alpha_{1,y}^3, \alpha_{1,x}^1, \alpha_{1,x}^2, \alpha_{1,x}^3, \sigma^1, \sigma^2, \sigma^3)' = 0, \quad (7.15)$$

um ein ineinander verschachteltes Design zu erhalten. Auf Grund der Parameter, die ausschließlich unter der Alternativhypothese identifizierbar sind, wird die asymptotische Verteilung der Teststatistik nicht standardmäßig der χ_{d-c}^2 -Verteilung entsprechen. Dabei geben die Freiheitsgrade gewöhnlich die Differenz in der Anzahl zu schätzender Parameter an.

Obwohl Garcia (1998) und auch Carrasco et al. (2005) asymptotische Verteilungen für ein solches Testproblem entwickeln, werden deren Ergebnisse nicht ganz den hier gestellten Anforderungen gerecht, da beide nur den Spezialfall von 2 Regimen und einem leicht geringer parametrisiertem Design betrachten.²³ Für die vorliegende Studie lassen sich jedoch auch Aussagen treffen, ohne dass die Verteilung der Teststatistik explizit simuliert wird: Erstens kann die Teststatistik rein komparativ verwendet werden, wenn im nächsten Abschnitt eine Rangfolge der Frühindikatoren erstellt wird, da der Unterschied in den Verteilungen für die einzelnen Spezifikation als vernachlässigbar eingestuft wird. Zweitens deutet eine Analyse der kritischen Werte für die von Garcia (1998) verfolgten Spezifikationen darauf hin, dass eine Ablehnung der Nullhypothese (7.15) für alle hier aufgeführten MS-Spezifikationen wahrscheinlich ist.

7.3.5 Rangfolge der Frühindikatoren

An sich ist das hier vorgestellte MS-Modell so konstruiert, dass die Information aller Frühindikatoren aus Abschnitt 7.3.1 gleichgewichtet Eingang findet. Trotzdem kann Bedarf bestehen, eine Rangfolge der Indikatoren zu erstellen, bspw. um die Frage zu beantworten, ob im Allgemeinen Indikatoren ohne verzögerte Datenveröffentlichung besser abschneiden. Hierfür werden die in den vorangegangenen Abschnitten eingeführten Maße bezogen auf die Veröffentlichung November 2010 benutzt. Implizit bedeutet das, die Frühindikatoren als Stellvertreter ihrer Spezifikationen zu betrachten.

²Das ist der komfortable Fall, in dem $E(S_t = i) = \frac{1 - \mathcal{P}(S_t = j | S_{t-1} = j)}{2 - \mathcal{P}(S_t = i | S_{t-1} = i) - \mathcal{P}(S_t = j | S_{t-1} = j)}$, $i, j = 1, 2$ gilt. Dies wird als Eintrag in allen Kovarianzmatrizen benutzt, mit deren Hilfe Garcia (1998) dann den asymptotischen χ^2 -Prozess bestimmt.

³Das Problem mit dem Ansatz von Carrasco et al. (2005) ist, dass das Testresultat ohne Schätzung des Modells unter der Alternativhypothese erhalten wird. Auf diese Weise ist es schwer, die Unterschiede

Tabelle 7.5: Rangfolge der Frühindikatoren gemäß der Standardmaße von Tabelle 7.4

Stichprobe: 1994m03-2010m09	1. Position	2. Position	3. Position
Gesamtkorrektheit der Phasen	Inländische Aufträge	Spread 10Y	Euribor
Bereinigtes R^2	Spread 10Y	Baugenehmigungen	Kreditwachstum
Quadratischer Fehler	Spread 10Y	Baugenehmigungen	Offene Stellen
LR Test Statistik	Offene Stellen	Corporate Spread	CDAX

Anmerkungen: Die Frühindikatoren werden als Stellvertreter der zugeordneten Spezifikation betrachtet. Es werden jeweils drei Indikatoren aufgeführt, die für das jeweilige Maß die besten Ergebnisse erreichten. Ergebnisse der autoregressiven Schätzung werden hier vernachlässigt.

Tabelle 7.5 führt die drei besten Frühindikatoren bezüglich jedes der Maße aus Tabelle 7.4 auf. Die weitgehende Konsistenz zwischen bereinigtem R^2 und quadratischem Fehler kann dadurch erklärt werden, dass beide auf der Summe quadrierter Residuen beruhen. Ersteres berücksichtigt allerdings zusätzlich die Anzahl geschätzter Parameter. Auf Grundlage von Tabelle 7.5 schneidet der Spread auf der Zinskurve mit 3 Nennungen am Besten ab. Dieser unterliegt keiner Veröffentlichungslücke, wie auch die Mehrzahl der genannten Indikatoren. Trotzdem ist das Verhältnis zwischen unmittelbar und verzögert zur Verfügung stehenden Indikatoren mit 8 zu 4 Nennungen nicht eindeutig genug, um im Allgemeinen auf eine größere Leistungsfähigkeit in Konjunkturanalysen zu schließen.

7.4 Out-of-sample-Evaluation

7.4.1 Modellanpassung in Echtzeit

Vorhersagen mit MS-Modellen werden ähnlich zu Gleichung (7.8) geleistet, d.h.

$$\hat{\xi}_{T+f|T} = \hat{P}^f \hat{\xi}_{T|T}, \quad (7.16)$$

wobei der Exponent den Vorhersagehorizont bezeichnet. Man erhält also die zukünftige Zustandswahrscheinlichkeit eines bestimmten Regimes aus den Zustandswahrscheinlichkeiten der letzten Beobachtung, wobei diese gewichtet werden mit den Übergangswahrscheinlichkeiten für den Eintritt in das bestimmte Regime. Man beachte, dass das hier vorgestellte Modell insoweit einem Standardansatz folgt, dass die Einträge der Übergangs-

zwischen den einzelnen MS-Spezifikationen aufzufangen.

matrix nicht abhängig sind von Amplitude oder Dauer des letzten Regimes. Ein solches Modell wurde etwa von Durland & McCurdy (1994) entwickelt.

Es ist offensichtlich, dass die Idee, ein drittes und viertes Regime einzuführen, unter dem Eindruck des wirtschaftlichen Abschwungs der Finanzmarktkrise bzw. der hierauf folgenden Erholung entstand. Allerdings dürfen Echtzeitschätzungen nicht auf Informationen beruhen, die erst später gegeben sind. Daher wurde ein Kriterium zur Entscheidung entwickelt, ob bei Verfügbarkeit neuer Daten in Echtzeit neue Regime im Modell erlaubt werden sollen.

Betrachtet man die Abbildung 7.2, so werden den zusätzlichen Regimen nennenswerte Wahrscheinlichkeiten erst seit dem Auftreten der Finanzmarktkrise zugeordnet. Allerdings sind dies In-sample-Ergebnisse für die Veröffentlichung im November 2010. Im Fall von Echtzeitvorhersagen ist es klar, dass wann immer neue Regime eingeführt werden, diesen auch Wahrscheinlichkeitsmasse zugeteilt wird. Die zentrale Frage lautet dabei jedoch, wieviel Wahrscheinlichkeitsmasse. Genau dies berücksichtigt das folgende Kriterium. Zunächst einmal gilt es sich zu verdeutlichen, dass mit steigender Anzahl von Regimen auch extremere Ereignisse abgebildet werden können. Beachtet man, dass das Verhalten in einem einzigen Monat oft einen Ausreißer darstellen kann, dann ist es sinnvoll, die Anzahl der Regime dann zu verändern, wenn die Wahrscheinlichkeit eines Extremregimes diejenige des gleichgerichteten Regimes für zwei aufeinander folgende Monate übertrifft, d.h.

$$\begin{aligned} & \mathcal{P}(S_{T+t+1} = \text{stark} | y_{T+t}) > \mathcal{P}(S_{T+t+1} = \text{schwach} | y_{T+t}) \\ \wedge & \mathcal{P}(S_{T+t+2} = \text{stark} | y_{T+t+1}) > \mathcal{P}(S_{T+t+2} = \text{schwach} | y_{T+t+1}), \end{aligned} \quad (7.17)$$

wobei $T + 1$ für den Anfang der Out-of-Sample Vorhersage steht.

Von einem operationellen Blickwinkel gesehen erfordert dieses Kriterium beide Modelleinstellungen parallel laufen zu lassen, also diejenige mit weniger und diejenige mit mehr Regimen, um dann anhand von (7.17) die relevante Einstellung auswählen zu können.

Eine weitere Frage, die es in diesem Zusammenhang zu erläutern gilt, ist, ob die Regimeanzahl in einem Schritt um ein oder zwei Regime wachsen kann. In diesem Artikel wird ein symmetrisches Vorgehen gewählt, d.h. dass sich die Anzahl aller Regime im Echtzeitverlauf ausschließlich von zwei auf vier verändern kann. Der Grund für dieses Vorgehen liegt in dem mit der Regimezuordnung verbundenen Aufwand. Führt man ein drittes Regime ein, ist ohne vollständige Berechnung zu jeder Veröffentlichung nicht klar, ob dieses Regime eine Rezession oder eine Expansion repräsentiert.

7 Wendepunktbestimmung mit Markov-Switching-Modellen

Tabelle 7.6: Änderung der Koeffizienten und Welch-Testresultate der Veröffentlichung August 2008 – Teil 1

Stichprobe: 1994m3-2008m6 Veröff.: 2008m8	Kons- tante	Exogener Koeffizient	σ Fehlerterm	Änderung Konstante	Änderung Koeffizient	Welch-Test- resultat
Autoregressive	0,0060	-	0,0026	45,80%	-	T=8,44
	0,0032	-	0,0010			abgelehnt
	-0,0001	-	0,0012	97,50%	-	T=24,73
	-0,0054	-	0,0019			abgelehnt
Ausländische Auftragseingänge	0,0065	0,0041	0,0023	64,89%	-701,24%	T=18,57
	0,0023	0,0327	0,0018			abgelehnt
	-0,0009	0,0801	0,0002	70,17%	-230,28%	T=1,97
	-0,0029	0,0242	0,0028			abgelehnt
Inländische Auftragseingänge	0,0099	0,0423	0,0013	69,07%	73,26%	T=21,76
	0,0031	0,0113	0,0027			abgelehnt
	0,0001	0,1650	0,0015	101,58%	239,02%	T=21,13
	-0,0041	-0,1187	0,0024			abgelehnt
Baugenehmigungen	0,0113	-0,0889	0,0000	54,33%	94,34%	T=17,92
	0,0052	-0,0050	0,0025			abgelehnt
	-0,0015	-0,0317	0,0011	45,07%	259,60%	T=2,20
	-0,0027	0,0198	0,0033			nicht abgelehnt
CDAX	0,0064	-0,0118	0,0025	38,44%	16,06%	T=9,63
	0,0039	-0,0099	0,0010			abgelehnt
	0,0002	-0,0042	0,0013	103,26%	30,05%	T=19,19
	-0,0052	-0,0061	0,0021			abgelehnt
Corporate Spread	0,0050	0,0263	0,0028	71,20%	-61,62%	T=12,73
	0,0014	0,0425	0,0017			abgelehnt
	-0,0019	0,0487	0,0024	67,75%	187,37%	T=14,72
	-0,0059	-0,0558	0,0012			abgelehnt

Anmerkungen: Zum gewählten Zeitpunkt wechselt das Modell von 2 zu 4 Regimen, so dass die relative Änderung der Koeffizienten bezüglich der Regime für jede der MS-Spezifikationen gemessen werden kann. Die Zeilenreihenfolge ist dabei wie folgt: Koeffizienten des Regimes mit der größten Konstante stehen oben und diejenigen mit der kleinsten Konstante unten. Welchtest-Resultate beziehen sich auf die Unterscheidung der Verteilungen zwischen schwacher und starker Rezession bzw. Expansion. Fortsetzung in Tabelle 7.7.

Tabelle 7.7: Änderung der Koeffizienten und Welch-Testresultate der Veröffentlichung Aug. 2008 – Fortsetzung von Tabelle 7.6

Stichprobe: 1994m3-2008m6 Veröff.: 2008m8	Kons- tante	Exogener Koeffizient	σ Fehlerterm	Änderung Konstante	Änderung Koeffizient	Welch-Test- resultat
Euribor - 3M	0,0065	0,2096	0,0034	47,22%	93,89%	T=7,51
	0,0034	0,0128	0,0027			abgelehnt
	0,0026	0,4194	0,0011	199,79%	619,07%	T=14,49
	-0,0026	-0,0808	0,0034			abgelehnt
ifo- Geschäftsklimaindex	0,0057	0,0805	0,0025	26,09%	227,80%	T=15,68
	0,0042	-0,1028	0,0007			abgelehnt
	0,0006	0,1260	0,0015	112,68%	-233,53%	T=8,07
	-0,0044	0,0378	0,0025			abgelehnt
Kreditwachstum	0,0075	-0,0485	0,0006	28,99%	122,55%	T=4,37
	0,0053	0,0109	0,0021			abgelehnt
	0,0002	0,0029	0,0021	102,95%	84,84%	T=29,81
	-0,0073	0,0193	0,0010			abgelehnt
Spread 10Y	0,0041	0,0062	0,0032	101,85%	3107,29%	T=3,81
	-0,0001	-0,1861	0,0002			abgelehnt
	-0,0018	0,0017	0,0035	86,00%	98,97%	T=41,57
	-0,0128	0,1674	0,0001			abgelehnt
Offene Stellen	0,0052	0,0232	0,0031	34,37%	-39,98%	T=4,56
	0,0034	0,0325	0,0012			abgelehnt
	-0,0006	0,0765	0,0005	84,27%	9,70%	T=12,24
	-0,0036	0,0847	0,0022			abgelehnt

Anmerkung: Siehe auch die Erläuterungen zu Tabelle 7.6.

Führt man dagegen gleich zwei neue Regime ein, so ist es sinnvoll, dass zwei dieser Regime zu höheren absoluten Wachstumsraten führen. Das bedeutet näherungsweise, dass das Regime mit der am stärksten positiven Konstante der starken Expansion zugeordnet wird, das Regime mit der am stärksten negativen Konstante der starken Rezession. Die dazwischen liegenden Fälle stellen dann schwache Expansion und schwache Rezession dar.⁴ Natürlich ist eine solche Approximation nur dann zulässig, wenn der Regressorvektor klein dimensioniert ist, so dass die Konstante, wie in Abschnitt 7.3.2 ausgeführt, den für die Iteration der Zustandswahrscheinlichkeiten wichtigsten Term darstellt. Wie man Tabelle 7.6 entnehmen kann, bestätigen Welchtest-Resultate dieses Vorgehen – bis auf einen Fall (Baugenehmigungen). Wenn man die Anzahl der Regime gemäß dem Kriterium (7.17) ändert, so kann die Hypothese⁵

$$\begin{aligned}
 H_0 : & \quad \text{Gleichgerichtete Regime besitzen dieselbe Normalverteilung.} \\
 \Leftrightarrow & \quad \mu_{\text{Expansion}}^{\text{starke}} | \sigma_{\text{Expansion}}^{\text{starke}} = \mu_{\text{Expansion}}^{\text{schwache}} | \sigma_{\text{Expansion}}^{\text{schwache}} \\
 \wedge & \quad \mu_{\text{Rezession}}^{\text{starke}} | \sigma_{\text{Rezession}}^{\text{starke}} = \mu_{\text{Rezession}}^{\text{schwache}} | \sigma_{\text{Rezession}}^{\text{schwache}}
 \end{aligned} \tag{7.18}$$

für die Veröffentlichung vom August 2008 abgelehnt werden.⁶ Dies bestätigt den Wechsel zur Modellierung mit vier Regimen, obwohl zum Zeitpunkt des Wechsels die wirtschaftliche Erholung in Deutschland noch nicht eingesetzt hatte.

Nachdem eine Entscheidung darüber gefallen ist, dass ein Wechsel zu vier Regimen für den Untersuchungszeitraum geeignet ist, kann eine weitere interessante Frage analysiert

⁴Bei einem solchen Vorgehen kann eine mögliche Fehlallokation zwar nicht kategorisch ausgeschlossen werden. So mögen in einer einzelnen Spezifikation auch ein Expansions- und drei Rezessionsregime auftreten, aber die Wahl eines symmetrischen Ansatz spiegelt den Hauptfund aus Abschnitt 7.3.3 wider.

⁵Die Hypothese impliziert, unter gegebenen empirischen Varianzen zu testen, ob die Mittelwerte gleich sein können. Es handelt sich um bedingte Normalverteilungen, sodass eine Ablehnung der Hypothese gleicher Erwartungswerte als Ablehnung von Verteilungsgleichheit interpretiert werden kann.

⁶Betrachtet man S_t als die relevante Filtration, so lässt sich mit Eigenschaften der bedingten Erwartung zeigen, dass diese im vorhandenen Modell durch die Konstante sowie den exogenen Teil der Gleichungen abschätzbar ist:

$$\begin{aligned}
 E(y_t | S_t) &= E(E(y_t | S_t) | S_{t-1}) \\
 &= E\left(\beta_0^{S_t} + \beta_{1,y}^{S_t} E(y_{t-1} | S_t) + \beta_{1,x}^{S_t} x_{t-1} | S_{t-1}\right) \\
 &= \beta_0^{S_t} + \beta_{1,x}^{S_t} x_{t-1} + \beta_{1,y}^{S_t} E(y_{t-1} | S_{t-1}) \\
 &= \beta_0^{S_t} + \beta_{1,x}^{S_t} x_{t-1} + \beta_{1,y}^{S_t} \left(\beta_0^{S_{t-1}} + \beta_{1,x}^{S_{t-1}} x_{t-1} + \beta_{1,y}^{S_{t-1}} E(y_{t-2} | S_{t-2})\right).
 \end{aligned}$$

Da für jeden β -Term $\beta \ll 1$ gilt, kann der letzte Term als Produkt mehrerer β -Terme vernachlässigt werden. Eine ähnliche Herangehensweise identifiziert σ^2 als wesentlichen Teil der bedingten Varianz.

werden. Offensichtlich ist die starke Abwärtsbewegung der letzten Rezession mit den Turbulenzen des Finanzmarktes verbunden. Deshalb könnte man annehmen, dass in dieser Situation Finanzmarktvariablen eine höhere Vorhersagekraft entwickeln.

Transferiert man diese Aussage auf das vorgestellte Modell, ließe sich erwarten, dass die (relative) Abweichung zwischen den Koeffizienten gleichgerichteter Regime für Finanzmarktvariablen höher ist als für realwirtschaftliche. Tabelle 7.6 enthält die entsprechenden Resultate. Obwohl die größte Veränderung bei einer Finanzmarktvariable (3-Monats-Euribor) feststellbar ist, kann die obige Aussage im Allgemeinen nicht bestätigt werden. Sicherlich spielen Finanzmarktvariablen eine bedeutende Rolle für Konjunkturprognosen. Aber das Resultat aus Tabelle 7.6 deutet darauf hin, dass man diese Rolle nicht überschätzen sollte, z.B. indem man bei verschiedenen Spezifikationen ein relativ ausgeglichenes Verhältnis zwischen Finanzmarkt- und realwirtschaftlichen Variablen wählt.

7.4.2 Echtzeitvorhersagen

Dieser Abschnitt behandelt Einmonats-Echtzeitvorhersagen für die Industrieproduktion als monatlichem Stellvertreter der gesamtwirtschaftlichen Aktivität. Aus welcher Methodik diese Resultate entstanden wurde in den vorangegangenen Abschnitten beschrieben. Tabelle (7.8) enthält Maße für die Genauigkeit der Vorhersage. Abgebildet sind diese sowohl für die verschiedenen MS-Spezifikationen als auch für ihren Durchschnitt (Average). Dabei steht MAE für den mittleren Absolutwert, d.h.

$$MAE = \frac{1}{h} \sum_{t=T+1}^{T+h} |\mathcal{P}_{t|t-1} - b_t|, \quad (7.19)$$

RMSE für den mittleren quadratischen Fehler, d.h.

$$RMSE = \sqrt{\frac{1}{h} \sum_{t=T+1}^{T+h} (\mathcal{P}_{t|t-1} - b_t)^2}, \quad (7.20)$$

und Theil für den Theil-Koeffizienten, d.h.

$$Theil = \frac{\sqrt{\sum_{t=T+1}^{T+h} (\mathcal{P}_{t|t-1} - b_t)^2 / h}}{\sqrt{\sum_{t=T+1}^{T+h} (\mathcal{P}_{t|t-1})^2 / h} + \sqrt{\sum_{t=T+1}^{T+h} b_t^2 / h}}, \quad (7.21)$$

Tabelle 7.8: Maße für die Vorhersagegüte der verschiedenen MS-Spezifikationen

Stichprobe: 2007m2 – 2010m12	MAE	RMSE	Theil
Average	0,2948	0,3898	0,4060
Autoregressiv	0,3070	0,4439	0,4309
Inländische Auftragseingänge	0,3651	0,5007	0,4340
Ausländische Auftragseingänge	0,3141	0,4385	0,4203
Baugenehmigungen	0,3050	0,4394	0,4112
CDAX	0,3234	0,4569	0,4421
Corporate Spread	0,3609	0,4958	0,4921
Euribor-3M	0,3132	0,4485	0,4537
ifo-Geschäftsklimaindex	0,3564	0,4991	0,4860
Kreditwachstum	0,2559	0,4395	0,4128
Spread 10Y	0,2987	0,4585	0,4279
Offene Stellen	0,3774	0,5038	0,4564

Anmerkungen: Bildet man den Durchschnitt über die Einzelvorhersagen und ändert man die Anzahl der eingebetteten Regime gemäß dem Kriterium (7.17), so führt dies zu den besten Werten. Aber durch die verspätete Rezessionserkennung kann der Theil-Koeffizient keine Werte unterhalb von 40% erreichen.

wobei letzterer auf das Einheitsintervall normiert ist, und 0 die perfekte Vorhersage bezeichnet. Hierbei bezeichnet b_t die binäre Variable, welche die Konjunkturphase gemäß des Benchmark-Modells abbildet. $\mathcal{P}_{t|t-1}$ aggregiert die MS-Zustandswahrscheinlichkeiten, welche Rezessionen zugeordnet werden. Vergleiche auch Abschnitt 7.4.1.

Wie man sehen kann, übertrifft die Average-Spezifikation leistungsmäßig jede einzelne Spezifikation. In diesem Zusammenhang erwähnt Timmermann (2006), „dass einfache Kombinationen, die Korrelationen zwischen den Vorhersagefehlern ignorieren, oft besser abschneiden gegenüber vermeintlich weiterentwickelten“. Darüber hinaus zieht die Average-Spezifikation, wie sie in Tabelle 7.8 enthalten ist, ihren Vorteil nicht nur aus den verschiedenen Spezifikationen, sondern auch aus der Tatsache, dass sich die Anzahl eingebetteter Regime auf dem Echtzeitpfad gemäß Kriterium (7.17) verändert. Dieser Wechsel findet für die Vorhersage September 2008 auf Grundlage der Augustveröffentlichung statt.⁷

⁷Der Titel der Gemeinschaftsdiagnose führender deutscher Wirtschaftsforschungsinstitute im Herbst 2008 lautet „Deutschland am Rande einer Rezession“. Im Vergleich dazu kann die Echtzeiteinführung neuer Regime für September 2008 so interpretiert werden, dass Deutschland nicht länger am Rand einer Rezession war, sondern in der Mitte einer Rezession steckte, die sich als strukturell tiefer andeutete als alle anderen Rezessionen zuvor.

Ein Grund, warum der Theil-Koeffizient nicht Werte unterhalb von 40% erreicht, wird in Abbildung 7.4 ersichtlich.⁸ Einerseits übertrifft die MS-Rezessionswahrscheinlichkeit die 50%-Grenze 5 Monate zu spät (August 2008) verglichen mit dem Rezessionsstart (März 2008), wie er wiederum später (Oktober 2008) von der Benchmark-Methode datiert wird. Andererseits reagiert das MS-Modell 3 Monate früher als die Benchmark-Methode, da die Augustprognose im Juli bekannt ist. Zusätzlich liegt die Juliprognose über 30%. Ein solches erstes Indiz für eine herannahende Rezession kann durch die nicht-parametrische Benchmark-Methode nicht zur Verfügung gestellt werden, da hier nur eine binäre Entscheidung (1=Rezession, 0=Expansion) getroffen werden kann. Beide Modelle datieren die letzte Rezession kontinuierlich zwischen August 2008 und April 2009. Verglichen mit dem Ende der Rezession im April 2009, wie es die Benchmarkmethode im November 2009 datiert, sinkt das MS-Modell erstmals 3 Monate später, d.h. im Juli 2009, unter die 50%-Schwelle. Betrachtet man dagegen den Zeitpunkt, in dem das Rezessionsende feststand, übertrifft das MS-Modell die Benchmark-Methode um 5 Monate.⁹

7.5 Zusammenfassung

Dieses Kapitel des Forschungsberichts beschäftigt sich mit der Anwendung eines Markov-Switching-Modells auf monatliche Echtzeitdaten für Deutschland. Während die Eignung der Methode für Konjunkturanalysen seit der Einführung der Modellklasse durch Hamilton (1989) grundsätzlich bekannt ist, lassen sich die Forschungsergebnisse dieses Kapitels wie folgt zusammenfassen:

Sieht man sich einer beschränkten Echtzeitdatenlage gegenüber, so ist es sinnvoll, die Idee von Timmermann (2006), dass Einzelvorhersagen Verzerrungen aufgrund von Fehlspezifizierung unterliegen, mit der Markov-Switching-Modellklasse zu verbinden. Um eine potenzielle Verzerrung zu reduzieren, werden also die Ergebnisse gemittelt. Erzeugt man auf diese Weise Echtzeitvorhersagen, so können verschiedene makroökonomische und Finanzmarktindikatoren berücksichtigt werden, statt sich auf eine rein autoregressive Schätzung zu verlassen. Dies erlaubt Aussagen über einen einzelnen Indikator oder eine

⁸Ein anderer Grund besteht darin, dass die MS-Wahrscheinlichkeiten im Vergleich zur binären Benchmark ein gewisses Intervall abdecken. So würden die Maße zur Vorhersagegüte allein dadurch verbessert, dass man auf Grundlage der 50%-Schwelle eine binäre Reihe aus den MS-Wahrscheinlichkeiten erzeugt.

⁹Das hier datierte Ende der Rezession muss technisch gesehen werden. Zum Beispiel bedeutet dies nicht, dass zu diesem Zeitpunkt das Vorkrisenniveau der Produktion wieder erreicht ist.

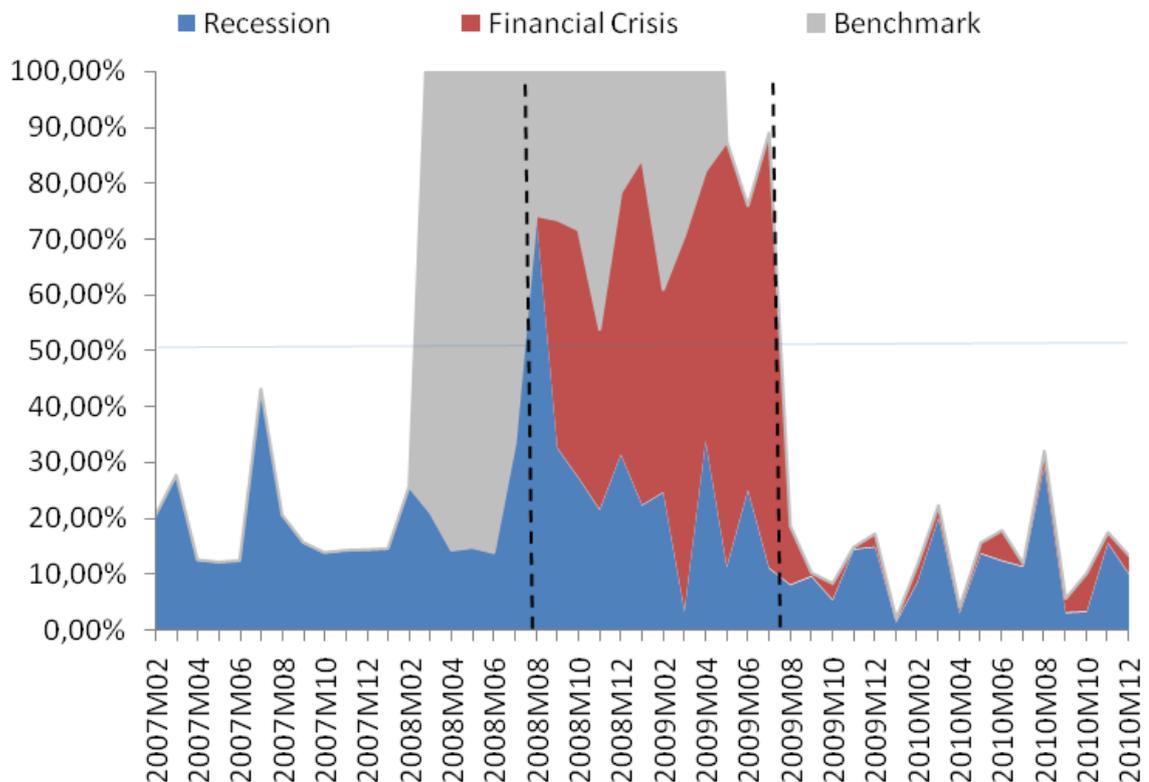


Abbildung 7.4: Einmonats-Echtzeitvorhersagen für Regimewahrscheinlichkeiten der Industrieproduktion. Die Regime Wahrscheinlichkeiten sind Durchschnittswerte der Einmonats-Echtzeitvorhersagen der verschiedenen MS Spezifikationen. Für September 2008 wechselt das Modell von 2 zu 4 Regimen gemäß dem in Abschnitt 7.4.1 beschriebenen Kriterium. Dabei deutet das neue Regime klar auf die Größenordnung des kommenden Abschwungs hin. Definiert man eine Rezession auf einer 50% Schwelle, so wird der Abschwung stabil zwischen August 2008 und Juli 2009 vorhergesagt (gestrichelte Linien). Dies stellt zwar eine Verspätung zu dem tatsächlichen Rezessionsstart dar, wie er im Nachhinein von der Benchmark Methode datiert wird. Allerdings wird die Rezession im Zeitverlauf drei Monate früher erkannt.

Indikatorgruppe.

Das Kreditwachstum, wie es aus den Bilanzen von Finanzdienstleistungsinstituten monatlich für die ESZB-Statistiken gemeldet wird, ist innerhalb einer MS-Spezifikation signifikant, wenn man Interbankengeschäfte berücksichtigt. Es verdient somit weitere Beachtung als ein möglicher Indikator des Konjunkturzyklus.

Diejenigen Frühindikatoren, die sofort verfügbar sind, verhalten sich im MS-Modell nicht allgemein leistungsfähiger als solche, die einer verzögerten Datenveröffentlichung unterliegen. Anhand dieses Ergebnisses ist eine einseitige Ausrichtung von Konjunkturmodellen auf unmittelbar vorhandene Daten (Finanzmarkt- und Umfragedaten) nicht zu empfehlen.

Trotz der Intuition, dass die letzte Krise ihren Ursprung auf den Finanzmärkten hatte, verhalten sich Finanzmarktindikatoren in den MS-Echtzeitvorhersagen des Konjunkturzyklus nicht allgemein sensitiver als realwirtschaftliche.

Erlaubt man dem MS-Modellansatz zusätzlich, die Anzahl eingebetteter Regime in Echtzeit zu verändern, so stabilisiert dies die Vorhersageresultate. Führt man hierfür ein Kriterium ein, so lässt sich der Zeitpunkt bestimmen, ab dem die Rezession nach der Finanzmarktkrise die vorangegangenen strukturell übertraf. In unserer Analyse geschieht dies für September 2008. Da die Vorhersage hierfür im August stattfindet, liegt die Informationsbereitstellung ein Monat vor der Insolvenzerklärung der Investment Bank Lehman Brothers.

Auf der Grundlage der Referenzzeitreihe dieses Forschungsberichts werden in der Beobachtungshistorie zwischen März 1994 und September 2010 sechs Rezessions- bzw. Stagnationsphasen durch das MS datiert. Dies deckt sich mit dem Ergebnis der in Abschnitt 4.3 eingeführten nicht-parametrischen Datierungsmethode und erweitert den Vorschlag von Schirwitz (2009) für eine deutsche Konjunkturzyklus-Chronologie. Dabei ergibt sich eine grundsätzliche Übereinstimmung insofern, dass die Autorin auf der Ebene ihrer einzelnen Modellergebnisse ebenfalls Rezessionsregime in den Jahren 1993, 1995, 1998, 2001, 2003 und sogar für die vermeintlich strittigste Phase in 2004 findet. Leichte Verschiebungen des Rezessionsbeginns und -endes sind darauf zurückzuführen, dass der vorliegende Forschungsbericht im Gegensatz zu Quartalsdaten des Bruttoinlandsprodukts die monatlich verfügbare Industrieproduktion als Referenzzeitreihe verwendet. Darüber hinausgehende Unterschiede zu den von der Autorin aus den einzelnen Modellergebnissen abgeleiteten Konsens-Konjunkturzyklen lassen sich unter anderem dadurch erklären, dass zwei ihrer Modelle (Boldin, MS aggregated) eine zusammenhängende Rezession von Februar 2001 bis Februar 2005 konstatieren. Diese Datierung einer solchen lang andauernden Rezessionsphase erscheint jedoch zumindest ausgehend vom letzten Informationsstand dieses Forschungsberichts (November 2010) eher zweifelhaft.

Bezüglich der Echtzeitvorhersagen zwischen Februar 2007 und Dezember 2010 übertrifft das vorgestellte MS-Modell die nicht-parametrische ex post-Datierung aus Abschnitt 4.3 in der Früherkennung von Rezessionsanfang und -ende, während sich ähnliche Merk-

male zeigen: Die Erkennung erfolgt einerseits zu spät (ex post) verglichen mit dem im Nachhinein festgelegten Rezessionsbeginn. Andererseits liegt jedoch keine irrtümliche Erkennung einer zusätzlichen Rezessionsphase im Vorhersagezeitraum vor. Schließlich sind mehrere Erweiterungen des hier vorgestellten Markov-Switching-Ansatzes möglich, so etwa die Berücksichtigung unterschiedlicher Vorhersagehorizonte.

8 Wendepunktbestimmung mit dynamischen Probit-Modellen

8.1 Einleitung

Modelle mit binären abhängigen Variablen sind während der letzten zwanzig Jahre weithin zur Schätzung und Vorhersage von Rezessionen benutzt worden. Zu den aktuellsten Studien, in denen diese Methodik angewandt wurde, gehören Moneta (2005), Wright (2006), Haltmaier (2008), Kauppi & Saikkonen (2008), Rudebusch & Williams (2009) und Nyberg (2010).

Seit der Veröffentlichung der Arbeit von Estrella & Hardouvelis (1991) hat es jedoch einen bedeutenden Fortschritt in der Entwicklung dieses ökonometrischen Ansatzes gegeben: Estrella & Hardouvelis (1991) machten zur Schätzung und Vorhersage von Rezessionen in den Vereinigten Staaten noch von sogenannten „statischen“ Probitmodellen Gebrauch, d.h. Modelle, in denen die binäre Rezessionsvariable nur durch verzögerte exogene Variablen erklärt wird. Stand der Forschung sind inzwischen jedoch dynamische autoregressive Probitmodelle von Kauppi & Saikkonen (2008), in denen sowohl verzögerte endogene Variablen als auch verzögerte Werte der latenten Probit-Variable einbezogen werden. Diese sind von einer deutlich höheren Komplexität.

In diesem Kapitel wird dem Ansatz von Proaño (2010) gefolgt, der eine alternative Spezifikation von dynamischen Probitmodellen vorschlägt, die insbesondere im Hinblick auf die Vorhersage von Rezessionen unter Echtzeitbedingungen attraktiv ist. Die in diesem Kapitel vorgestellte Methodologie unterscheidet sich jedoch in verschiedenen Aspekten von diesen Arbeiten. Zum einen da sie auf der Vorstellung beruht, dass die Kombination verschiedener Prognosen einen höheren Informationsgehalt hat als die Einzelprognosen (Timmermann 2006). Dies wird in diesem Kapitel noch im Detail diskutiert. Zum anderen weil sie so spezifiziert wurde, dass die resultierenden Prognosen in wahrer Echtzeit

bestimmt werden.

Es ist hervorzuheben, dass der in diesem Kapitel vorgestellte Konjunkturindikator so programmiert wurde, dass es dem Anwender ermöglicht wird, auf einfache Weise einen klaren Einblick in den Informationsstand eines beliebigen Rechenstandes zu erlangen. Gleichzeitig wird die Spezifikation aller Probit-Gleichungen automatisch für jeden untersuchten Rechenstand neu und unabhängig von früheren Spezifikationen gewählt und geschätzt, was die Angemessenheit der daraus resultierenden Schätzungen aus ökonomischer Perspektive gewährleistet.

Dieses Kapitel ist folgendermaßen aufgebaut: In Abschnitt 8.2 wird die theoretische Grundlage für den Konjunkturindikator dieses Kapitels vorgestellt und die Echtzeitimplikationen für die Spezifikation von dynamischen Probitmodellen diskutiert. Abschnitt 8.3 beschreibt den zu Grunde liegenden Datensatz, sowie die genaue Implementation des Probit-Indikators. In Abschnitt 8.4 werden sowohl die Insample-Anpassung sowie die Prognosegüte dieses Konjunkturindikators diskutiert, und Abschnitt 8.5 fasst die Ergebnisse dieses Kapitels zusammen.

8.2 Dynamische Probit-Modelle

Wie üblich in der oben diskutierten Literatur zur Schätzung und Vorhersage von Rezessionen, definieren wir die binäre Rezessionsindikator-Zeitreihe b_t , welche den Zustand der Volkswirtschaft im Konjunkturzyklus darstellen soll, folgendermaßen:

- $b_t = 1$, wenn die Volkswirtschaft zum Zeitpunkt t eine Rezessionsphase durchläuft,
- $b_t = 0$, wenn die Volkswirtschaft zum Zeitpunkt t eine expansive Phase durchläuft.

Sei Ω_{t-h} die Informationsmenge zum Zeitpunkt $t-h$, wobei h den Vorhersagehorizont beschreibt. Seien weiterhin E_{t-1} und $Prob_{t-1}(\cdot)$ der bedingte Erwartungswert und die bedingte Wahrscheinlichkeit für die Informationsmenge Ω_{t-1} . Dann ergibt sich unter der Annahme, dass für b_t eine Bernoulli-Verteilung $b_t|\Omega_{t-1} \sim B(p_t)$ vorliegt, für einen einperiodigen Vorhersagehorizont $h = 1$ die folgende bedingte Wahrscheinlichkeit p_t dafür, dass b_t den Wert 1 annimmt:

$$E_{t-1}(b_t) = Prob_{t-1}(b_t = 1) = p_t = \Phi(\varphi_t)$$

wobei φ_t die lineare Modellgleichung der Variablen darstellt, die in der Informationsmenge Ω_{t-1} enthalten sind. Die Verteilungsfunktion $\Phi(\cdot)$ stellt die Verbindung zwischen φ_t und der bedingten Wahrscheinlichkeit $Prob_{t-1}(b_t = 1)$ gemäß der Bernoulliverteilung her und ist in Probitmodellen durch eine Standardnormalverteilung gegeben.

Unter Echtzeitbedingungen gibt es jedoch einige Restriktionen, die die Bestandteile der Informationsmenge Ω_{t-h} betreffen. Wie in Proaño (2010) werden hierfür zwei weitere Verzögerungsarten explizit definiert: Der Rezessionserkennungslag r , d.h. die Anzahl der notwendigen Verzögerungen, die der verwendete Algorithmus braucht, um einen Wendepunkt datieren zu können, sowie der Datenverfügbarkeitslag s , wobei normalerweise $r > s$ gilt. Mit Hilfe dieser beiden expliziten Lag-Arten kann die Informationsmenge, die einer Vorhersage über h Perioden zu Grunde liegt, weiter spezifiziert werden als:

$$\Omega_{t-h} = \{b_{t-j}, b_{t-j-1}, \dots, \mathbf{x}_{t-k}, \mathbf{x}_{t-k-1}, \dots\}, \text{ mit } j \geq h+r, \quad k \geq h+s \quad (8.1)$$

Hierbei beschreibt b_{t-j} einen Skalar, der die Werte des binären Rezessionsindikators enthält, und \mathbf{x}_{t-k} einen Vektor, der eine Reihe zusätzlicher erklärender Variablen beinhaltet, die zum Zeitpunkt $t-h$ verfügbar sind. Implizit benennt Dueker (1997, p.45) diesen Zusammenhang folgendermaßen:

„Three months is probably a minimum recognition lag time for recessions. It would be clearly not reasonable to include last month’s value of the recession binary variable as an explanatory variable, because it takes more time to recognize that the economy has entered a recession. [...] If three months seems less than the minimum recognition lag time for recessions, then one can concentrate on the results for 6, 9 and 12 months.”

Die explizite Spezifikation dieser Lag-Untergrenzen ist keinesfalls trivial: Wie z.B. in Stark & Croushore (2002) angemerkt wird, muss im Gegensatz zu einer ökonometrischen Analyse mit ex-post Daten bei der Gestaltung eines ökonometrischen Instruments zur Vorhersage von Rezessionen unter Echtzeitbedingungen sowohl der Rezessionsidentifikationslag als auch der bei den meisten makroökonomischen Zeitreihen vorliegende Datenverfügbarkeitslag beachtet werden.¹

Unter Berücksichtigung dieser echtzeitbedingten Beschränkungen für die Lag-Länge lautet z.B. die lineare Modellgleichung für die h -Schritt-Prognose einer Rezession zum

¹Siehe dazu auch Kapitel 2 und 4.

Zeitpunkt t im Falle eines statischen Probitmodells wie z.B. bei Estrella & Hardouvelis (1991) oder Estrella & Mishkin (1998) folgendermaßen:

$$\varphi_t = c + \mathbf{x}'_{t-k} \beta, \text{ mit } k \geq h + s. \quad (8.2)$$

Der Mehrwert einer expliziten Formulierung der echtzeitbedingten Lag-Untergrenzen ist auch im Falle dynamischer Probitmodelle offensichtlich, bei denen Lags der binären Rezessionsvariablen als zusätzliche erklärende Variablen aufgenommen werden, um einer möglichen Autokorrelation innerhalb der binären Rezessionszeitreihe gerecht zu werden (Dueker 1997). Unter Echtzeitbedingungen kann die dem Probitmodell $E_{t-h}(b_t) = \Phi(\varphi_t)$ zu Grunde liegende lineare Funktion wegen des Identifikationslags der binären Rezessionsvariablen r (und wegen des Datenverfügbarkeitslags s) nur wie folgt spezifiziert werden:

$$\varphi_t = \sum_{j=h+r}^q \delta_j b_{t-j} + \mathbf{x}'_{t-k} \beta, \quad k \geq h + s \quad (8.3)$$

Wie diese Gleichung deutlich zeigt, kann der dynamische Probitansatz von Dueker (1997) nur einer möglichen Autokorrelation der binären Variablen einer Ordnung von höher als $h+r$ gerecht werden. Entsprechend ergibt sich, dass die Fähigkeit des dynamischen Probitansatzes, die Autokorrelation niedriger Ordnung der binären Reihe b aufzufangen, (und damit auch dessen Vorhersagekraft) umso niedriger sein wird, je länger der Rezessionserkennungslag ist (Proaño 2010).

Kürzlich entwickelten Kauppi & Saikkonen (2008) eine dynamisch-autoregressive Spezifikation, nach der nicht nur Verzögerungen der binären Rezessionsreihe b_t im Modell enthalten sind, sondern auch Verzögerungen der linearen Funktion φ_t . Für eine Ein-Schritt-Prognose ergibt sich:

$$\varphi_t = \sum_{i=1}^p \alpha_i \varphi_{t-i} + \sum_{j=1}^q \delta_j b_{t-j} + \mathbf{x}'_{t-1} \beta \quad (8.4)$$

Wie in Nyberg (2010) diskutiert wird, kann durch rekursive Substitution im speziellen Fall $\varphi_t = \alpha_1 \varphi_{t-1} + \mathbf{x}'_{t-k} \beta'$ gezeigt werden, dass das dynamische autoregressive Modell aus (8.4) durch

$$\varphi_t = \delta_1 \sum_{i=1}^{\infty} \alpha_1^{i-1} b_{t-i} + \sum_{i=1}^{\infty} \alpha_1^{i-1} \mathbf{x}'_{t-i} \beta, \quad (8.5)$$

also durch eine Erweiterung „unendlicher“ Ordnung des dynamischen Probitmodells $\varphi_t = \delta_1 b_{t-1} + \mathbf{x}'_{t-1} \beta$, gegeben ist.

Bei expliziter Berücksichtigung der Rezessionserkennungs- und Datenverfügbarkeits-Laguntergrenzen ist die tatsächlich unter Echtzeitbedingungen durchführbare dynamische autoregressive Probitspezifikation wie folgt (wieder unter Annahme von $h = p = 1$ und $q = 0$ aus Gründen der Vereinfachung der Notation):

$$\varphi_t = \delta_1 \sum_{i=1+r}^{\infty} \alpha_1^{i-1-r} b_{t-i} + \sum_{i=1+s}^{\infty} \alpha_1^{i-1-s} \mathbf{x}'_{t-i} \beta \quad (8.6)$$

Für den allgemeineren Fall einer h -Schritt-Prognose ergibt sich:

$$\varphi_t = \sum_{i=1}^p \alpha_i \varphi_{t-i} + \sum_{j=h+r}^q \delta_j b_{t-j} + \mathbf{x}'_{t-k} \beta, \quad k \geq h+s \quad (8.7)$$

Aus (8.6) wird deutlich, dass es stark von den Laggrenzen r und s abhängt, inwieweit die dynamische autoregressive Spezifikation in der Lage ist, eine mögliche Autokorrelation niedriger Ordnung in b_t (wie auch in φ_t) aufzufangen.

Da normalerweise der Informationslag s viel kürzer ist als der Rezessionserkennungslag r , schlägt Proaño (2010) vor, in die lineare Modellgleichung –und daher in die Informationsmenge Ω_{t-h} – den Vektor der zyklischen Referenzzeitreihen aufzunehmen und so die Autokorrelation niedriger Ordnung in der binären Rezessionsreihe b_t zu berücksichtigen. Für den Fall einer h -Schritt-Prognose bedeutet dies:

$$\varphi_t = \sum_{j=h+r}^q \delta_j b_{t-j} + \sum_{i=h+s}^q \mathbf{y}'_{t-i} \alpha_i + \sum_{j=h+s}^q \mathbf{x}'_{t-j} \beta_j \quad (8.8)$$

Proaño (2010) motiviert diese alternative Spezifikation mit der Anmerkung, dass die binäre Rezessionsreihe gemäß des verwendeten Wendepunkt-Datierungsalgorithmus durch eine nichtlineare Transformation der Referenzreihen im Vektor \mathbf{y} generiert wird.

Die folgenden Abschnitte beschreiben die Anwendung eines dynamischen Probitmodells zur Vorhersage von Rezessionen mit deutschen makroökonomischen Daten.

8.3 Rezessionsprognosen mit Probit-Modellen in Echtzeit

8.3.1 Datensatz

Für die folgende Analyse wurde eine Reihe von makroökonomischen Variablen und Finanzmarktindikatoren monatlicher Frequenz herangezogen, welche wie in der gesamten vorliegenden Studie den Datenbanken der Deutsche Bundesbank Deutsche Bundesbank (2011, 2010a,b) und des ifo-Instituts für Wirtschaftsforschung (ifo Institut für Wirtschaftsforschung 2011) entnommen wurden. Die makroökonomischen Indikatoren, die als erklärende Variablen dienen, sind der monatliche Produktionsindex (IP1AA020), die Inlands- und Auslands-Auftragseingänge in der Industrie und die Anzahl offener Stellen. Die entsprechenden Echtzeitmatrizen unterliegen zum Zeitpunkt der Bearbeitung einer Datenverfügbarkeitslücke von zwei Monaten. Der monatliche Produktionsindex ist auch die Referenzzeitreihe für die Bestimmung der binären Rezessionsvariable, vgl. 4.3.1. Die binäre Rezessionsvariable wurde anhand des im Kapitel 4 diskutierten Wendepunkterkennungsverfahrens erzeugt.

Die verwendeten Finanzmarktindikatoren wurden so gewählt, dass sie die Finanzmärkte möglichst umfassend darstellen. Zum einen wurde gemäß Bernanke (1990) und Friedman & Kuttner (1992) der Spread zwischen der durchschnittlichen Rendite auf Unternehmensanleihen aller Laufzeiten und der Durchschnittsrendite deutscher Staatsanleihen herangezogen (Renditeaufschlag Unternehmensanleihen – *corporate spread*). Zum anderen wurden die Wachstumsraten des CDAX genutzt, um die Entwicklung am deutschen Aktienmarkt abzubilden. Des Weiteren wurde wie bereits in Stock & Watson (1989), Estrella & Hardouvelis (1991), Estrella & Mishkin (1998) sowie kürzlich in Kauppi & Saikkonen (2008) und Nyberg (2010) der Spread zwischen Zinsen langfristiger und kurzfristiger Anleihen als Regressor aufgenommen. Aufgrund des fehlenden Konsenses darüber, welcher Renditespread die besten Prognoseeigenschaften aufweist, und wegen weiterer später zu erläuternder Gründe wurden verschiedene Spezifikationen mit den Spreads zwischen dem Dreimonats-Euribor und dem Zins 1-, 2-, 3-, 5- und 10-jähriger Anleihen geschätzt. Außerdem wurde als zusätzlicher Regressor der kurzfristige Zinssatz aufgenommen. Diesbezüglich haben Ang et al. (2006) anhand eines dynamischen Faktormodells gezeigt, dass die zwei Hauptfaktoren der Zinsstrukturkurve –welche in ihrer Arbeit 90% der Variation der gesamten Zinsstruktur erklären– über alle Laufzeiten stark mit dem Kurzfristzins

und dem 10-Jahres-Zinsspread korreliert sind. Zusätzlich zeigt Wright (2006), dass Probitmodelle mit dem kurzfristigen Dreimonatszinssatz von US-T-Bills sowie mit dem Spread zwischen 10-jährigen T-Bonds und dreimonatigen T-Bills bessere Ergebnisse bezüglich des mittleren quadratischen Fehlers (MQF/MSE) hervorbringen als Probitspezifikationen, die nur den Spread nutzen.

8.3.2 Modellspezifikation

Um Multikollinearität aufgrund der starken Korrelation zwischen den Zinsspreads unterschiedlicher Laufzeiten zu vermeiden, wurden verschiedene dynamische Probitmodelle spezifiziert und geschätzt. In allen Modellen sind die Variablen Industrieproduktion, offene Stellen, der ifo-Geschäftsklimaindex, der CDAX-Preisindex (alle Variablen als monatliche Wachstumsraten in %), sowie der Unternehmensanleihsread und der Dreimonats-Euribor enthalten. Dazu werden alternativ jeweils der 1-, 2-, 3-, 5- oder 10-Jahres-Zinsspread (relativ zum Dreimonats-Euribor) aufgenommen.

Aufgrund der Tatsache, dass die ausgewählten erklärenden Variablen verschiedene Vorlaufzeiten haben können, diese aber im Vorfeld nicht bekannt sind, erscheint es sinnvoll, mehrere Lags auf Signifikanz zu testen. Grundsätzlich unterliegen Ökonometriker bei der Spezifikation der Modelle einem Abwägungsproblem: Einerseits entstehen beim Weglassen von scheinbar irrelevanten Variablen „Modellsuchkosten“, da die Anpassung sinkt und die betreffenden Variablen mit einer gewissen Wahrscheinlichkeit doch wichtig sind. Andererseits treten bei der Berücksichtigung vieler Variablen „Inferenzkosten“ auf, die durch den Effizienzverlust bei der Aufnahme eigentlich insignifikanter Variablen entstehen. Um das Problem einer willkürlichen Wahl der Modellspezifikation zu vermeiden, die auf einer ad-hoc Auswahl der Lags oder der erklärenden Variablen im Allgemeinen basieren würde, wurde jede Spezifikation sowohl mit einem *general-to-specific* als auch mit einem *specific-to-general* Ansatz geschätzt.

Im *general-to-specific* Auswahlverfahren (siehe Campos et al. 2005) wurde der Erklärungswert jedes Lags einer jeden erklärenden Variablen mit Hilfe eines Likelihood-Ratio (LR) -Tests auf redundante Regressoren untersucht. Die LR-Statistik wurde dabei folgendermaßen berechnet:

$$LR = -2(\mathcal{L}_{\mathcal{R}} - \mathcal{L}_{\mathcal{U}}),$$

wobei $\mathcal{L}_{\mathcal{M}}$ und $\mathcal{L}_{\mathcal{N}}$ die maximierten Werte der (Gaußschen) log-likelihood-Funktion des kompletten Modells und der reduzierten Version sind.² Im Gegensatz dazu wurde in der *specific-to-general* Modellauswahl der zusätzliche Erklärungswert jedes Lags der erklärenden Variable mit Hilfe eines LR-Tests auf vernachlässigte Variablen gemessen, wobei unter der Nullhypothese gilt, dass der Koeffizient der hinzugefügten Variable (bzw. ihres Lags) verschwindet.

Im zweiten Schritt folgen wir Timmermann (2006), indem wir die geschätzten Rezessionswahrscheinlichkeiten aus den verschiedenen dynamischen Probitmodellen verwenden und daraus den ungewichteten Durchschnitt bilden. Insgesamt ergibt sich damit eine einzelne Zeitreihe der Rezessionswahrscheinlichkeiten. Wie kürzlich von Aiolfi et al. (2010) dargelegt wurde, ist dieses Verfahren, das verschiedene Prognosen kombiniert, einer einzelnen Prognose überlegen. Der Grund dafür ist die Reduktion negativer Effekte aufgrund von Instabilitäten einzelner Modelle, wodurch sich die Robustheit der Schätzergebnisse verbessert.

8.4 Schätzergebnisse

Für die Evaluierung des hier vorgestellten Probit-Prognosemodells wurden die 47 Rechenstände der Monate 2007m1 bis 2010m11 genutzt. Das bedeutet, dass für diesen Untersuchungszeitraum jeweils 470 einzelne Probit-Regressionsgleichungen für jeden Prognosehorizont spezifiziert wurden.

Wie bereits erwähnt wurde der entwickelte Konjunkturindikator so programmiert, dass er automatisch für jeden Rechenstand

- eine neue binäre Rezessionsvariable anhand der aktuellsten Daten der Referenzzeitreihe (Industrieproduktion) berechnet, und
- komplett neue Probit-Regressionsgleichungen einerseits anhand der neuen Information der neuesten Beobachtungen und andererseits anhand der neuesten Revisionen der makroökonomischen Zeitreihen schätzt.

²Unter der Nullhypothese dieses χ^2 -Tests mit einem Freiheitsgrad ist der Koeffizient der redundanten Variable (bzw. des jeweiligen Lags) Null. Ein Verwerfen dieser Hypothese führt zur Beibehaltung der getesteten Variable (bzw. ihres Lags) in der Modellspezifikation.

8.4.1 In-Sample-Evaluierung

Wir wenden uns in diesem Abschnitt zunächst der In-Sample-Evaluierung für einen bestimmten Rechenstand zu. Abbildungen 8.1 bis 8.3 und Tabellen 8.1, 8.2 und 8.3 fassen die Ergebnisse der dynamischen Probit-Regressionsgleichungen für die 1-, 2- und 3-Monats-Prognosen für den Rechenstand 2010m9, die durch die general-to-specific (bezeichnet mit G) und die specific-to-general (S) Ansätze generiert wurden, zusammen.

Wie diese Tabellen verdeutlichen, werden zu jedem Rechenstand zehn verschiedene Probit-Regressionsgleichungen spezifiziert, welche –wie in diesem Fall– sehr ähnlich aussehen oder, wie z.B. im Zwischenbericht dieses Projektes für den Rechenstand 2010m5, deutlich heterogener sein können. Dieser Punkt ist besonders erwähnenswert weil, wie bereits angemerkt, die Spezifikation dieser Probit-Regressionsgleichungen für jeden Prognosehorizont zu jedem Rechenstand vollständig automatisiert ist (gemäß der im Abschnitt 8.3.2 beschriebenen Methode). So liefert dieser Ansatz anhand der Daten des Rechenstands 2010m9 eine recht stabile Spezifikation, welche sowohl von der Wahl des Bund-Spreads (Eins-, Zwei-, Drei-, Fünf- oder Zehn-Jahres-Zinsspread gegenüber dem Drei-Monats-Euribor) als auch vom verwendeten Lag-Auswahlverfahren (*general-to-specific* oder *specific-to-general*) unabhängig zu sein scheint, und in welcher der ifo-Indikator und der CDax den höchsten Erklärungsgehalt zu haben scheinen. Im Gegensatz hierzu war im Zwischenbericht eine höhere Heterogenität in den Schätzergebnisse für die drei Prognosehorizonte zu finden, welche in einem höheren Maß den Einsatz des erläuterten Kombinationsansatzes nahelegte. So schienen damals die Signifikanzen der meisten Variablen (und deren Lags) stark vom jeweiligen Zinsspread als auch vom verwendeten Lag-Auswahlverfahren (*general-to-specific* oder *specific-to-general*) abzuhängen. Es fand sich jedoch eine gewisse Konstanz in den Signifikanzen bestimmter Variablen (Lags), welche aber vom gewählten Prognosehorizont abhing. Bei einem einmonatigen Prognosehorizont waren beispielsweise die dritten Lags (relativ zum Endpunkt) der offenen Stellen und der Auslandsaufträge sowie das siebte Lag der monatlichen Wachstumsrate des CDAX in allen Spezifikation durchgehend signifikant.

Die Tatsache, dass bei diesem spezifischen Rechenstand eine recht ähnliche Regressorenauswahl für alle Prognosehorizonte getroffen wurde, führt dazu, dass die verschiedenen Regressionsgleichungen ähnliche Prognosen liefern, wie in Abb. 8.1 deutlich zu erkennen ist.

Jede dieser (in den meisten Fällen verschiedenen) Probit-Regressionsgleichungen liefert

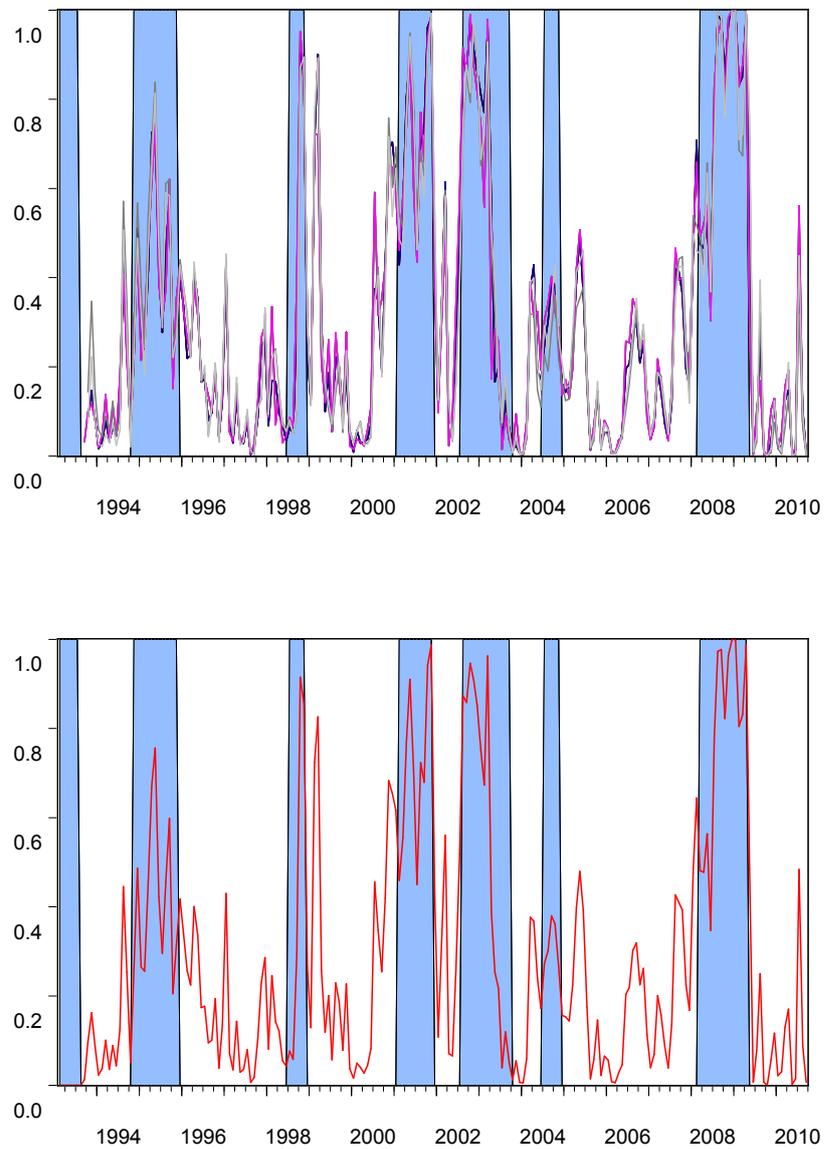


Abbildung 8.1: In-Sample-Anpassung der geschätzten und durchschnittlichen Rezessionswahrscheinlichkeiten, Einmonats-Prognosehorizonte, Rechenstand 2010m9.

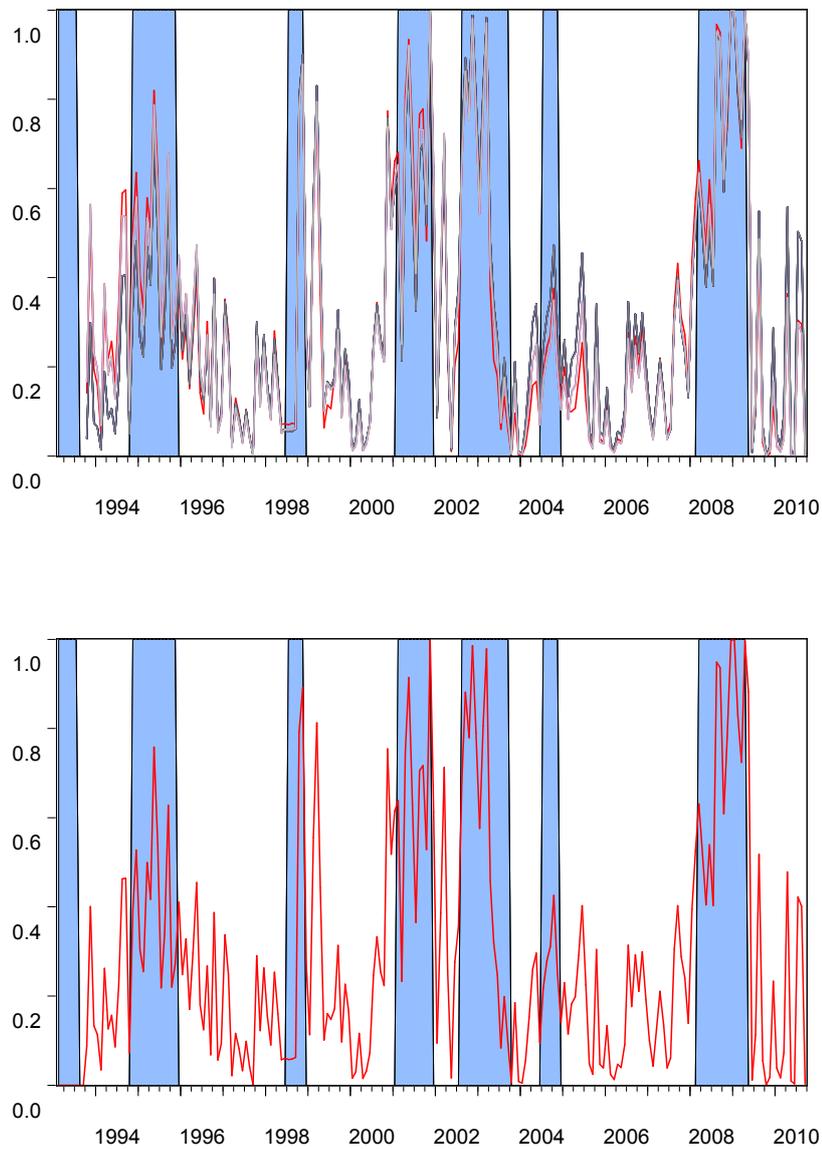


Abbildung 8.2: In-Sample-Anpassung der geschätzten und durchschnittlichen Rezessionswahrscheinlichkeiten, Zweimonats-Prognosehorizonte, Rechenstand 2010m9.

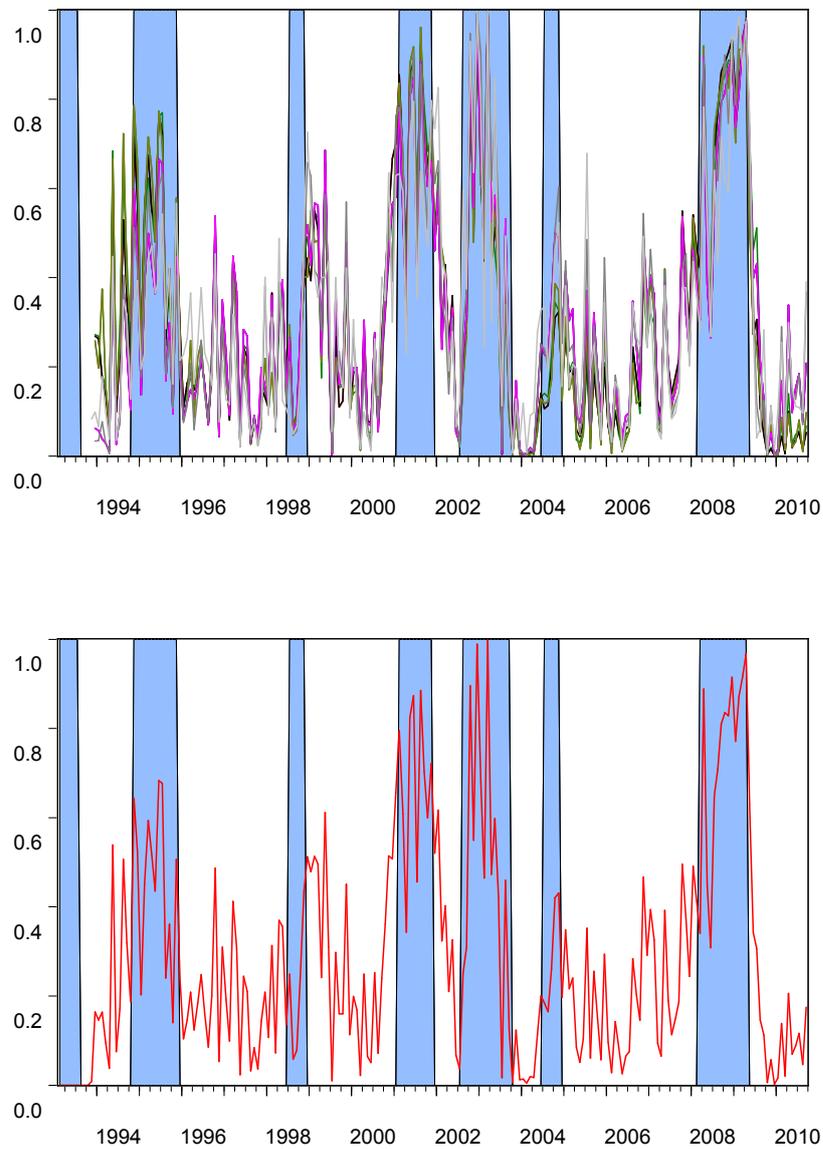


Abbildung 8.3: In-Sample-Anpassung der geschätzten und durchschnittlichen Rezessionswahrscheinlichkeiten, Dreimonats-Prognosehorizonte, Rechenstand 2010m9.

Tabelle 8.1: Dynamische Probitregressionen, Ein-Monats-Prognosehorizont, Rechenstand 2010m9

	YC1G	YC1S	YC2G	YC2S	YC3G	YC3S	YC5G	YC5S	YC10G	YC10S
Stichprobe: 1993m2 – 2010m8.										
Recess'ind.	-	-	-	-	-	-	-	-	-	-
Ind'prod.	-	-	-	-	-	-	-	-	-	-
Aufträge Incl.	-	-	-	-	-	-	-	-	-	-
Aufträge Ausl.	-	-	-	-	-	-	-	-	-	-
off. Stellen	-	-	-	-	-	-	-	-	-	-
ifo-idx	1;2	1;2	1;2	1;2	1;2	1;2	1;2	1;2	1;2	1;2
Unt' anleithesprd.	-	-	-	-	-	-	-	-	-	-
CDax	1;2;3;6;7	1;2;3;6	1;2;3;6	1;2;3;6	1;2;3;6;7	1;2;3;6;7	1;2;3;6;7	1;2;3;6	2;3;6;7	2;3;6;7
Euribor 3M	-	-	-	-	-	-	-	-	2	-
YC1Y	-	-	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
YC2Y	n.a.	n.a.	-	-	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
YC3Y	n.a.	n.a.	n.a.	n.a.	-	-	n.a.	n.a.	n.a.	n.a.
YC5Y	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	-	-	n.a.	n.a.
YC10Y	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	-	8
Pseudo- R^2	0,349	0,331	0,349	0,331	0,349	0,349	0,349	0,331	0,337	0,332
SSR	25,395	25,851	25,395	25,851	25,395	25,395	25,395	25,851	25,801	26,211
Mittl. Log-Lik.	-0,401	-0,412	-0,401	-0,412	-0,401	-0,401	-0,401	-0,412	-0,409	-0,412
AIC	0,883	0,894	0,883	0,894	0,883	0,883	0,883	0,894	0,898	0,895
SC	1,016	1,010	1,016	1,010	1,016	1,016	1,016	1,010	1,031	1,011
HQC	0,937	0,941	0,937	0,941	0,937	0,937	0,937	0,941	0,952	0,942

Tabelle 8.2: Dynamische Probitregressionen, Zwei-Monats-Prognosehorizont, Rechenstand 2010m9

	Stichprobe: 1991m1 – 2010m5; EQ_...									
	YC1G	YC1S	YC2G	YC2S	YC3G	YC3S	YC5G	YC5S	YC10G	YC10S
Rezess'ind.	8	-	-	-	-	-	-	-	-	-
Ind'prod.	-	-	-	-	-	-	-	-	-	-
Aufträge Inl.	-	-	-	-	-	-	-	-	-	-
Aufträge Ausl.	-	-	-	-	-	-	-	-	-	-
off. Stellen	-	-	-	-	-	-	-	-	-	-
ifo-idx	1;2	2	2	2	2	2	2	2	2	2
Unt'anlehesprd.	-	-	-	-	-	-	-	-	-	-
CDax	2;3;6;7;8	2;3;6;7	2;3;6;7	2;3;6;7	2;3;6;7	2;3;6;7	2;3;6;7	2;3;6;7	2;3;6;7	2;3;6;7
Euribor 3M	8	-	-	-	-	8	-	8	-	8
YC1Y	-	-	n.a.							
YC2Y	n.a.	n.a.	-	-	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
YC3Y	n.a.	n.a.	n.a.	n.a.	-	-	n.a.	n.a.	n.a.	n.a.
YC5Y	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	-	-	n.a.	n.a.
YC10Y	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	-	-
Pseudo-R ²	0,314	0,292	0,292	0,292	0,292	0,307	0,292	0,307	0,292	0,307
SSR	26,611	27,820	27,820	27,820	27,820	27,008	27,820	27,008	27,820	27,008
Mittel. Log-Lik.	-0,423	-0,436	-0,436	-0,436	-0,436	-0,427	-0,436	-0,427	-0,436	-0,427
AIC	0,927	0,934	0,934	0,934	0,934	0,925	0,934	0,925	0,934	0,925
SC	1,060	1,033	1,033	1,033	1,033	1,042	1,033	1,042	1,033	1,042
HQC	0,981	0,974	0,974	0,974	0,974	0,972	0,974	0,972	0,974	0,972

Tabelle 8.3: Dynamische Probitregressionen, Drei-Monats-Prognosehorizont, Rechenstand 2010m9

	Stichprobe: 1993m2 – 2010m8									
	YC1G	YC1S	YC2G	YC2S	YC3G	YC3S	YC5G	YC5S	YC10G	YC10S
Rezess'ind.	9	9	9	9	9	9	9	9	9	-
Ind'prod.	-	-	-	6	-	6	-	-	6	6
Aufträge Incl.	-	-	-	-	-	-	-	-	-	-
Aufträge Ausl.	-	-	-	-	-	-	-	-	-	-
off. Stellen	-	-	-	-	-	-	-	-	-	-
ifo-idx	4	4	4	4	4	4	4	4	4	4
Unt' anleithesprd.	-	-	-	-	-	-	-	-	-	-
CDax	3;6;8;9	3;6;8;9	3;6;8;9	3;6;8;9	3;6;8;9	3;6;8;9	3;6;8;9	3;6;8;9	3;4;6;8;9	3;4;6;8
Euribor 3M	3	9	3	9	-	9	-	-	-	-
YC1Y	-	-	n.a.	n.a.						
YC2Y	n.a.	n.a.	-	-	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
YC3Y	n.a.	n.a.	n.a.	n.a.	-	-	n.a.	n.a.	n.a.	n.a.
YC5Y	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	-	-	n.a.	n.a.
YC10Y	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	-	-
Pseudo-R ²	0,285	0,284	0,285	0,297	0,257	0,297	0,257	0,257	0,282	0,240
SSR	27,037	27,256	27,037	26,672	28,516	29,314	28,516	28,516	27,783	30,374
Mittl. Log-Lik.	-0,443	-0,443	-0,443	-0,435	-0,460	-0,412	-0,460	-0,460	-0,445	-0,469
AIC	0,968	0,969	0,968	0,963	0,992	0,963	0,992	0,992	0,981	1,058
SC	1,102	1,102	1,102	1,113	1,109	1,113	1,109	1,109	1,132	1,127
HQC	1,022	1,023	1,022	1,024	1,093	1,024	1,093	1,093	1,042	1,058

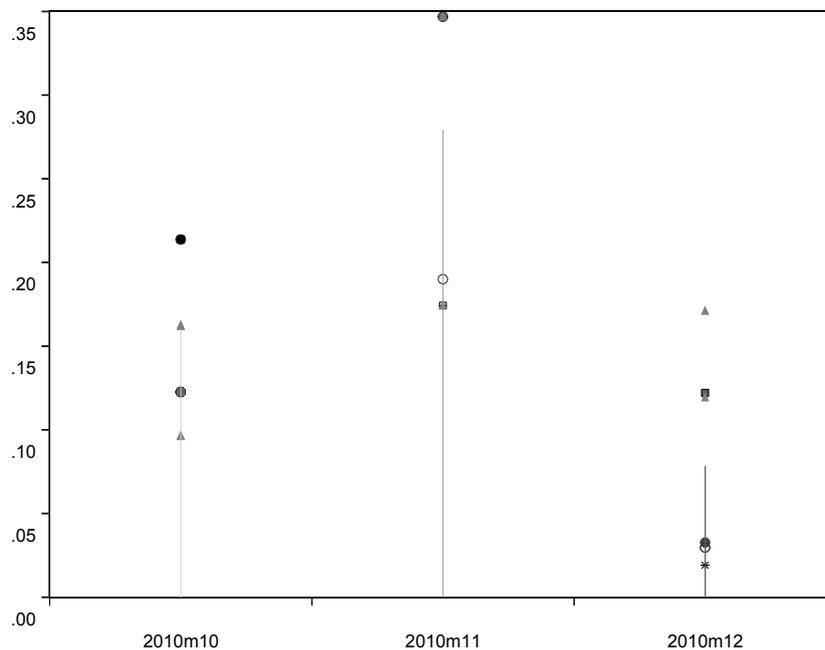


Abbildung 8.4: Beispiel der geschätzten und durchschnittlichen Rezessionswahrscheinlichkeiten out-of-sample, Ein-, Zwei- und Dreimonats-Prognosehorizonte, Rechenstand 2010m9.

für jeden Prognose-Horizont einen Prognosewert, welcher dann mit den anderen Prognosewerten kombiniert wird, um eine Durchschnittsprognose für jeden Prognosehorizont zu liefern. Abbildung 8.4 zeigt beispielhaft die Ein-, Zwei- und Drei-Monatsprognosen für den Rechenstand 2010m09.

Wie Abbildung 8.4 deutlich zeigt, berechnet der Probit-Indikator für jeden Horizont den einfachen Durchschnitt der prognostizierten Rezessionswahrscheinlichkeiten der einzelnen Probit-Regressionsgleichungen.

Um die gemeinsame Signifikanz aller Lags einer Variablen statistisch zu testen, wurde ein Likelihood-Ratio-Test verwendet, welcher die Differenz der Likelihood-Funktionen einer unrestringierten und einer restringierten Gleichung vergleicht. Unter der Null-Hypothese hätte keiner der restringierten Lag-Koeffizienten einen Einfluss, weshalb der Wert der maximierten Likelihood-Funktion kaum verändert wäre.

Die Ergebnisse dieser Tests für die Regressionsgleichungen PBYC1G und PBYC1S für den Ein-Monats-Prognosehorizont sowie für PBYC10G und PBYC10S für den Drei-

Monats-Prognosehorizont für alle Rechenstände von 2007m1 bis 2010m11 sind in den Tabellen 8.4 bis 8.11 beispielhaft zusammengefasst.

Die Tabellen 8.4 bis 8.11 zeigen deutlich, dass die automatisierte Lag-Selektionsprozedur nicht nur eine hohe individuelle, sondern auch eine hohe *gemeinsame* Signifikanz der ausgewählten Lags bei der großen Mehrheit der untersuchten Rechenstände erreicht, was die Nutzung dieser Spezifikationsstrategie bestätigt. Der Vergleich zwischen Tabellen 8.4 und 8.6, und zwischen 8.8 und 8.10 zeigt darüber hinaus, dass der *general-to-specific*- und der *specific-to-general*-Ansatz durchaus zu verschiedenen Spezifikationen führen und es deshalb vorteilhaft ist, die durch beide Ansätze erlangten Informationen zu kombinieren anstatt nur einen Ansatz zu nutzen. So wird z.B. beim *specific-to-general*-Ansatz in der Gleichung EQ_YC1G für einige Rechenstände die Industrieproduktion als eine wichtige erklärende Variable identifiziert (sowie auch in geringerem Umfang die Variable *corporate spread*), während dies beim *general-to-specific*-Ansatz in der Gleichung EQ_YC1S nicht der Fall ist. Gleichmaßen identifiziert der *general-to-specific*-Ansatz in der Gleichung EQ_YC10G bei einigen der Rechenstände die Lags der Auftragseingänge aus dem Inland als wichtige erklärende Variablen, während dies beim *specific-to-general*-Ansatz in Gleichung EQ_YC10S nicht der Fall ist.

Um den eigentlichen Beitrag jedes Lags einer ausgewählten Variable zum Niveau der latenten Variable der einzelnen Probit-Regressionsgleichungen zu ermitteln, wurde das Produkt jedes Regressors mit seinem dazugehörigen Koeffizienten berechnet, und durch den geschätzten Wert der latenten Variable des jeweiligen Prognosehorizonts geteilt. Dadurch wird der Beitrag jedes Regressors als Prozent der geschätzten latenten Variable dargestellt. Tabelle 8.12 zeigt am Beispiel des Rechenstands 2010m9 den Beitrag aller Regressoren für alle Probit-Spezifikationen für den Ein-Monats-Prognosehorizont.

Es gibt jedoch nicht nur Unterschiede zwischen den verschiedenen Spezifikationen und den Lag-Selektions-Prozeduren. So ist z.B. hervorzuheben, dass in dem untersuchten Zeitraum bestimmte Lags der binären Rezessionsvariablen, das ifo-Geschäftsklima, der CDax und der Drei-Monats-Euribor bei einem Großteil der Rechenstände und für alle drei Prognosehorizonte durch beide Auswahlprozeduren als statistisch signifikant erkannt werden.

Beim Vergleich mit den Ergebnissen früherer verwandter empirischer Studien ergeben sich vor allem zwei interessante Punkte. Zum einen bestätigen die Resultate den bereits in anderen Arbeiten (z.B. Harvey 1989, Stock & Watson 1989, und kürzlich Haltmaier 2008) dokumentierten Prognosewert der Aktienkursentwicklung für die zukünftige wirtschaftliche Aktivität. Zum anderen scheint die Vorhersagekraft der Zinsspreads (unabhängig von

Tabelle 8.4: Likelihood-Ratio-Tests PBY1CG IMF – Teil 1

	RECES_	IPROD_	AE_INL_	AE_AUSL_	VAC	IFO_R1	C_	CDAX	EURIBOR	YC1Y
	IND	OB	BUBA	BUBA		SPREAD	3M			
2007M01	0,0080					0,0041		0,0052	0,0000	0,0004
2007M02	0,0077					0,0043		0,0050	0,0000	0,0004
2007M03	0,0074					0,0043		0,0048	0,0000	0,0003
2007M04	0,0100					0,0016		0,0049	0,0000	0,0000
2007M05	0,0095					0,0015		0,0050	0,0000	0,0000
2007M06	0,0093					0,0014		0,0050	0,0000	0,0000
2007M07	0,0092					0,0013		0,0050	0,0000	0,0000
2007M08								0,0016	0,0000	0,0000
2007M09								0,0016	0,0000	0,0000
2007M10								0,0016	0,0000	0,0000
2007M11								0,0015	0,0000	0,0000
2007M12								0,0014	0,0000	0,0000
2008M01								0,0014	0,0000	0,0000
2008M02								0,0014	0,0000	0,0000
2008M03								0,0014	0,0000	0,0000
2008M04	0,0129					0,0035		0,0015	0,0000	0,0000
2008M05	0,0235					0,0021		0,0000	0,0000	0,0000
2008M06	0,0234					0,0021		0,0000	0,0000	0,0000
2008M07	0,0234					0,0021		0,0000	0,0000	0,0000
2008M08	0,1152					0,0181		0,0000	0,0000	0,0000
2008M09						0,0124		0,0000	0,0000	0,0038
2008M10						0,0120		0,0000	0,0000	0,0032
2008M11						0,0000		0,0000	0,0000	
2008M12						0,0000		0,0000	0,0000	

Tabelle 8.5: Likelihood-Ratio-Tests PBY1CG IMF – Fortsetzung

	RECES_ IND	IPROD_ OB	AE_INL_ BUBA	AE_AUSL_ BUBA	VAC	IFO_R1	C_ SPREAD	CDAX	EURIBOR 3M	YCIY
2009M01						0,0000			0,0000	
2009M02						0,0000			0,0000	
2009M03						0,0002		0,0023	0,0020	
2009M04	0,0465					0,0000		0,0000	0,0102	
2009M05	0,0465					0,0000		0,0000	0,0102	
2009M06	0,0465					0,0000		0,0000	0,0102	
2009M07	0,0447					0,0000		0,0000	0,0074	
2009M08	0,0436					0,0000		0,0000	0,0055	
2009M09	0,0436					0,0000		0,0000	0,0055	
2009M10	0,0436					0,0000		0,0000	0,0055	
2009M11						0,0000		0,0000		
2009M12						0,0011		0,0000		
2010M01						0,0009		0,0000		
2010M02						0,0002		0,0000	0,0611	
2010M03						0,0006		0,0000		
2010M04						0,0006		0,0000		
2010M05						0,0002		0,0000	0,0600	
2010M06						0,0003		0,0000		
2010M07						0,0011		0,0000	0,0360	
2010M08						0,0011		0,0000	0,0355	
2010M09						0,0002		0,0000		
2010M10						0,0002		0,0000		
2010M11						0,0002		0,0000		

Tabelle 8.6: Likelihood-Ratio-Tests PBY1CS IMF – Teil 1

	RECES_	IPROD_	AE_INL_	AE_AUSL_	VAC	IFO_R1	C_	CDAX	EURIBOR 3M	YC1Y
	IND	OB	BUBA	BUBA			SPREAD			
2007M01	0,0273					0,0001	0,0108	0,0038	0,0000	0,0031
2007M02	0,0267					0,0001	0,0106	0,0037	0,0000	0,0031
2007M03	0,0030	0,0252				0,0064		0,0037	0,0000	0,0000
2007M04	0,0068	0,0665				0,0037		0,0041	0,0000	0,0000
2007M05	0,0064	0,0656				0,0034		0,0043	0,0000	0,0000
2007M06	0,0062	0,0645				0,0032		0,0043	0,0000	0,0000
2007M07	0,0061	0,0643				0,0031		0,0043	0,0000	0,0000
2007M08								0,0020	0,0000	0,0000
2007M09								0,0020	0,0000	0,0000
2007M10								0,0020	0,0000	0,0000
2007M11								0,0019	0,0000	0,0000
2007M12								0,0019	0,0000	0,0000
2008M01								0,0018	0,0000	0,0000
2008M02								0,0018	0,0000	0,0000
2008M03								0,0018	0,0000	0,0000
2008M04	0,0057	0,0552				0,0116		0,0019	0,0000	0,0000
2008M05	0,0361					0,0094		0,0000	0,0000	0,0000
2008M06	0,0362					0,0077		0,0000	0,0000	0,0000
2008M07	0,0362					0,0077		0,0000	0,0000	0,0000
2008M08	0,0234					0,0021		0,0000	0,0000	0,0000
2008M09	0,0236					0,0020		0,0000	0,0000	0,0000
2008M10	0,0237					0,0019		0,0000	0,0000	0,0000
2008M11						0,0000		0,0000	0,0000	0,0000
2008M12	0,0400					0,0000		0,0000	0,0000	0,0000

Tabelle 8.7: Likelihood-Ratio-Tests PBY1CS IMF – Fortsetzung

	RECES_ IND	IPROD_ OB	AE_INL_ BUBA	AE_AUSL_ BUBA	VAC	IFO_R1	C_ SPREAD	CDAX	EURIBOR 3M	YCIY
2009M01						0,0000			0,0000	
2009M02						0,0000			0,0000	
2009M03	0,0021					0,0000	0,0128		0,0000	
2009M04						0,0003	0,0000			
2009M05						0,0003	0,0000			
2009M06						0,0003	0,0000			
2009M07						0,0002	0,0000			
2009M08						0,0002	0,0000			
2009M09						0,0002	0,0000			
2009M10						0,0002	0,0000			
2009M11						0,0086	0,0000			
2009M12		0,1356				0,0000	0,0000		0,0168	
2010M01		0,0788				0,0000	0,0000		0,0185	
2010M02		0,0280				0,0000	0,0000		0,0166	
2010M03		0,0279				0,0000	0,0000		0,0166	
2010M04		0,0269				0,0000	0,0000		0,0165	
2010M05		0,0284				0,0000	0,0000		0,0174	
2010M06						0,0004	0,0000			
2010M07						0,0003	0,0000			
2010M08						0,0003	0,0000			
2010M09						0,0003	0,0000			
2010M10						0,0002	0,0000			
2010M11						0,0002	0,0000			

Tabelle 8.8: Likelihood-Ratio-Tests PBY10CG 3MF – Teil 1

	RECES_	IPROD_	AE_INL_	AE_AUSL_	VAC	IFO_R1	C_	CDAX	EURIBOR	YC10Y
	IND	OB	BUBA	BUBA			SPREAD		3M	
2007M01	0,0000				0,0290	0,0000	0,0103	0,0001	0,0000	
2007M02	0,0000				0,0289	0,0000	0,0103	0,0001	0,0000	
2007M03	0,0000				0,0288	0,0000	0,0101	0,0001	0,0000	
2007M04	0,0000				0,0191	0,0000	0,0114	0,0004	0,0000	
2007M05	0,0000				0,0419	0,0000	0,0287	0,0039	0,0000	
2007M06	0,0000					0,0000		0,0044	0,0000	
2007M07	0,0044					0,0001		0,0047		
2007M08						0,0012		0,0001		
2007M09						0,0011		0,0001		
2007M10						0,0009		0,0001		
2007M11								0,0000		
2007M12								0,0000		
2008M01								0,0000		
2008M02								0,0000		
2008M03								0,0000		
2008M04	0,0001					0,0000	0,0225	0,0011	0,0000	0,0010
2008M05	0,0014		0,0238			0,0001	0,0292	0,0002	0,0000	0,0023
2008M06	0,0015		0,0232			0,0001	0,0264	0,0002	0,0000	0,0016
2008M07	0,0042		0,0245			0,0002	0,0393	0,0001	0,0000	0,0011
2008M08	0,0043		0,0214			0,0002	0,0401	0,0001	0,0000	0,0008
2008M09	0,0044		0,0170			0,0002	0,0395	0,0001	0,0000	0,0006
2008M10	0,0028		0,0608			0,0004	0,0840	0,0005	0,0000	0,0005
2008M11	0,0024					0,0002	0,1235	0,0022	0,0000	0,0417
2008M12	0,0687					0,0035		0,0011	0,0000	

Tabelle 8.9: Likelihood-Ratio-Tests PB10CG 3MF – Fortsetzung

	RECES_	IPROD_	AE_INL_	AE_AUSL_	VAC	IFO_R1	C_	CDAX	EURIBOR	YC10Y
	IND	OB	BUBA	BUBA			SPREAD		3M	
2009M01						0,0026		0,0048		
2009M02						0,0023		0,0012		
2009M03	0,0002					0,0124		0,0000	0,0000	
2009M04	0,0037					0,0039		0,0000	0,0114	
2009M05	0,0040					0,0029		0,0000	0,0113	
2009M06		0,3509						0,0000	0,0363	
2009M07		0,3688						0,0000	0,0362	
2009M08		0,1034				0,0195		0,0000		
2009M09		0,1036				0,0187		0,0000		
2009M10		0,1023				0,0186		0,0000		
2009M11	0,0281	0,1904			0,0177	0,0007		0,0000	0,0275	
2009M12	0,0122	0,2047			0,0472	0,0013		0,0000	0,0115	
2010M01	0,0128	0,1472			0,0477	0,0014		0,0000	0,0096	
2010M02	0,0126	0,1272			0,0494	0,0012		0,0000	0,0087	
2010M03	0,0126	0,1171			0,0495	0,0012		0,0000	0,0085	
2010M04		0,0264				0,0024		0,0000	0,0258	
2010M05		0,0262				0,0024		0,0000	0,0256	
2010M06		0,0262				0,0024		0,0000	0,0254	
2010M07		0,0126				0,0022		0,0000	0,0165	
2010M08		0,0006				0,0013		0,0000	0,0063	
2010M09	0,0095	0,0674				0,0003		0,0000		
2010M10	0,0029					0,0001		0,0000		
2010M11	0,0030					0,0001		0,0000		

Tabelle 8.10: Likelihood-Ratio-Tests PBY10CS 3MF – Teil 1

	RECES_	IPROD_	AE_INL_	AE_AUSL_	VAC	IFO_R1	C_	CDAX	EURIBOR	YC10Y
	IND	OB	BUBA	BUBA			SPREAD		3M	
2007M01	0,0000				0,0515	0,0000	0,0192	0,0001	0,0000	
2007M02	0,0000				0,0513	0,0000	0,0191	0,0001	0,0000	
2007M03	0,0000				0,0511	0,0000	0,0188	0,0001	0,0000	
2007M04	0,0000					0,0000		0,0065	0,0000	
2007M05	0,0000				0,0373	0,0000	0,0417	0,0014	0,0000	
2007M06	0,0000				0,0591	0,0000		0,0010	0,0000	
2007M07	0,0000					0,0000		0,0036	0,0000	
2007M08								0,0000		
2007M09								0,0000		
2007M10								0,0000		
2007M11								0,0000		
2007M12								0,0000		
2008M01								0,0000		
2008M02								0,0000		
2008M03								0,0000		
2008M04						0,0021		0,0039		
2008M05	0,0068					0,0001	0,0302	0,0036	0,0000	
2008M06	0,0001					0,0001		0,0092	0,0000	0,0373
2008M07						0,0030		0,0050		
2008M08	0,0014					0,0014		0,0016	0,0000	0,0033
2008M09	0,0013					0,0003		0,0023	0,0000	0,0357
2008M10	0,0018					0,0018		0,0019	0,0000	0,0020
2008M11						0,0041		0,0047		
2008M12	0,0004					0,0017		0,0336	0,0000	

Tabelle 8.11: Likelihood-Ratio-Tests PBY10CS 3MF – Fortsetzung

	RECES_	IPROD_	AE_INL_	AE_AUSL_	VAC	IFO_RI	C_	CDAX	EURIBOR	YC10Y
	IND	OB	BUBA	BUBA			SPREAD		3M	
2009M01						0,0026		0,0048		
2009M02						0,0023		0,0012		
2009M03	0,0017					0,0014		0,0001		0,0328
2009M04	0,0027					0,0030		0,0000	0,0093	
2009M05	0,0029					0,0022		0,0000	0,0093	
2009M06	0,0204					0,0133		0,0000		
2009M07	0,0211					0,0123		0,0000		
2009M08	0,0215					0,0090		0,0000		
2009M09	0,0217					0,0085		0,0000		
2009M10	0,0217					0,0085		0,0000		
2009M11	0,0013					0,0004		0,0000	0,0111	
2009M12	0,0051	0,2196				0,0006		0,0000	0,0100	
2010M01	0,0059	0,0796				0,0005		0,0000	0,0090	
2010M02	0,0058	0,0679				0,0004		0,0000	0,0089	
2010M03	0,0057	0,0646				0,0004		0,0000	0,0089	
2010M04	0,0058	0,0645				0,0004		0,0000	0,0089	
2010M05	0,0053	0,0661				0,0004		0,0000	0,0085	
2010M06	0,0053	0,0661				0,0004		0,0000	0,0085	
2010M07		0,0001				0,0008		0,0000	0,0071	
2010M08		0,0002				0,0007		0,0000	0,0055	
2010M09		0,0546				0,0012		0,0000		
2010M10		0,0254				0,0023		0,0000		
2010M11		0,0230				0,0007		0,0000		

Tabelle 8.12: Einfluss der einzelnen Regressoren auf die latente Variable im Probit-Modell beim Ein-Monats-Prognosehorizont, Rechenstand 2010m9

	CONST	IFO_R1 DLN1	IFO_R1 DLN2	CDAX DLN1	CDAX DLN2	CDAX DLN3	CDAX DLN6	CDAX DLN7
EQ_YC1G	-0,542780	0,047259	0,467389	-0,055904	0,077637	-0,002572	0,251304	-0,006618
	CONST	IFO_R1 DLN1	IFO_R1 DLN2	CDAX DLN1	CDAX DLN2	CDAX DLN3	CDAX DLN6	
EQ_YC1S	-0,541119	0,046028	0,446138	-0,061846	0,081289	-0,002621	0,269183	
	CONST	IFO_R1 DLN1	IFO_R1 DLN2	CDAX DLN1	CDAX DLN2	CDAX DLN3	CDAX DLN6	CDAX DLN7
EQ_YC2G	-0,542780	0,047259	0,467389	-0,055904	0,077637	-0,002572	0,251304	-0,006618
	CONST	IFO_R1 DLN1	IFO_R1 DLN2	CDAX DLN1	CDAX DLN2	CDAX DLN3	CDAX DLN6	
EQ_YC2S	-0,541119	0,046028	0,446138	-0,061846	0,081289	-0,002621	0,269183	
	CONST	IFO_R1 DLN1	IFO_R1 DLN2	CDAX DLN1	CDAX DLN2	CDAX DLN3	CDAX DLN6	CDAX DLN7
EQ_YC3G	-0,542780	0,047259	0,467389	-0,055904	0,077637	-0,002572	0,251304	-0,006618
	CONST	IFO_R1 DLN1	IFO_R1 DLN2	CDAX DLN1	CDAX DLN2	CDAX DLN3	CDAX DLN6	
EQ_YC3S	-0,541119	0,046028	0,446138	-0,061846	0,081289	-0,002621	0,269183	
	CONST	IFO_R1 DLN1	IFO_R1 DLN2	CDAX DLN1	CDAX DLN2	CDAX DLN3	CDAX DLN6	CDAX DLN7
EQ_YC5G	-0,542780	0,047259	0,467389	-0,055904	0,077637	-0,002572	0,251304	-0,006618
	CONST	IFO_R1 DLN1	IFO_R1 DLN2	CDAX DLN1	CDAX DLN2	CDAX DLN3	CDAX DLN6	
EQ_YC5S	-0,541119	0,046028	0,446138	-0,061846	0,081289	-0,002621	0,269183	
	CONST	IFO_R1 DLN1	IFO_R1 DLN2	CDAX DLN1	CDAX DLN2	CDAX DLN3	CDAX DLN6	EURIBOR 3M 2
EQ_YC10G	-0,966563	0,042566	0,393016	0,064214	-0,002082	0,213509	-0,0006030	-0,0333380
	CONST	IFO_R1 DLN1	IFO_R1 DLN2	CDAX DLN2	CDAX DLN3	CDAX DLN6	CDAX DLN7	
EQ_YC10S	-0,551716	0,050226	0,448954	0,072399	-0,002339	0,232476	-0,0006417	

Anmerkung: DLN = Wachstumsrate (approximativ, Differenz des Logarithmus)

der gewählten Laufzeit) weniger signifikant als allgemein angenommen.

8.4.2 Out-of-Sample-Evaluierung

Um die Güte der out-of-sample-Prognosen der geschätzten Probit-Regressionsgleichungen unter Echtzeit zu evaluieren, wurden für jeden Rechenstand und für jeden Prognosehorizont die bereits besprochenen Probitregressionen mit dem dem Rechenstand zugrundeliegenden Datenstand neu spezifiziert und geschätzt. Anschließend wurden, von jedem Rechenstand aus, die Wahrscheinlichkeiten einer Rezession bei den verschiedenen Prognosehorizonten aufgezeichnet, siehe Moneta (2005). Auf Basis der Rechenstände 2007m1 bis 2010m11 liegen als Ergebnis dieser Vorgehensweise anschließend die geschätzten out-of-sample Wahrscheinlichkeiten für den Berichtszeitraum 2007m2 bis 2010m12 für die Ein-Monatsprognose, für den Zeitraum 2007m3 bis 2011m1 für die Zwei-Monatsprognose und für den Zeitraum 2007m4 bis 2011m2 für die Drei-Monatsprognose vor.

Um die out-of-sample-Prognosegüte eines geschätzten Probitmodells \mathcal{M} zu quantifizieren benutzen wir drei Kennzahlen, die in der Literatur (siehe z.B. Rudebusch & Williams 2009) oft für die Vorhersagegenauigkeit herangezogen werden:

Die mittlere absolute Abweichung (mean absolute error – MAE):

$$MAE(\mathcal{M}, h) = \frac{1}{T} \sum_{t=1}^T |P_{t|t-h}^{\mathcal{M}} - b_t|,$$

die Wurzel des mittleren quadratischen Fehlers (root mean squared error – RMSE)

$$RMSE(\mathcal{M}, h) = \sqrt{\frac{1}{T} \sum_{t=1}^T (P_{t|t-h}^{\mathcal{M}} - b_t)^2},$$

sowie den Theilschen Ungleichheitskoeffizienten (Theil Inequality Coefficient)

$$Theil = \frac{\sqrt{\sum_{t=T+1}^{T+h} (P_{t|t-h}^{\mathcal{M}} - b_t)^2 / h}}{\sqrt{\sum_{t=T+1}^{T+h} (P_{t|t-h}^{\mathcal{M}})^2 / h} + \sqrt{\sum_{t=T+1}^{T+h} b_t^2 / h}}.$$

Letzterer liegt im Einheitsintervall $[0; 1]$, wobei eine 0 für eine perfekte Vorhersage und eine 1 für keinerlei Erklärungsgehalt steht. Wie die Ergebnisse in Tabelle 8.13 zeigen, liefern die erstellten dynamischen Probitmodelle sowie die geschätzten Wahrscheinlichkeitszeitreihen aus deren Kombination statistisch aussagekräftige Resultate. Außerdem haben

sie eine signifikante Vorhersagekraft und weisen auch gute out-of-sample-Eigenschaften auf.

Abbildung 8.5 zeigt die Echtzeit-Durchschnittsrezessionswahrscheinlichkeiten für die drei berechneten Prognosehorizonte. Wie zu erwarten ist, ist die Anpassung der out-of-sample prognostizierten Echtzeit-Rezessionswahrscheinlichkeiten nicht so gut wie in der Stützstichprobe, wo die geschätzten Wahrscheinlichkeiten in der Mehrheit der Fälle die Rezessionsperioden richtig und rechtzeitig erkennen. Auch bei diesem speziellen Prognosezeitraum jedoch bilden die geschätzten Rezessionswahrscheinlichkeiten der Probit-Regressionen bis zu einem gewissen Maß den Anfangszeitpunkt und Dauer der Rezession –so wie sie durch die binäre Rezessionvariable definiert sind– ab. Die Prognosegüte dieses Rezessionsindikators variiert jedoch mit dem Prognosehorizont: Während die Ein- und Zwei-Monatsprognosen rechtzeitig über den Wert von 50% Rezessionswahrscheinlichkeit steigen, scheint die Drei-Monatsprognose die Rezession nicht zu erkennen. Die Ein- und Zwei-Monatsprognosen fallen jedoch innerhalb von zwei Monaten wieder unter die 50% Marke, um erst nach fünf Monaten wieder über den Wert von 50% zu steigen.

Es ist jedoch darauf hinzuweisen, dass dieses Ergebnis nicht unbedingt repräsentativ für die Prognosegüte dieses Indikators ist, da die 2008er Rezession durchaus als ein Regime-Wechsel betrachtet werden könnte, bei dem viele bis dahin zu vernachlässigende makroökonomische und finanzmarktwirtschaftliche Zusammenhänge in den Vordergrund getreten sind. Eine Evaluierung dieses Indikators bei der nächsten Rezession würde einen besseren Einblick in die Performance dieser Prognosemethode liefern.

8.5 Schlussfolgerungen

In diesem Kapitel wurde ein Echtzeit-Indikator zur Vorhersage von Rezessionen vorgestellt, welcher auf der Kombination verschiedener dynamischer Probit-Gleichungen basiert. Bei der Implementierung dieses Prognose-Indikators wurde insbesondere auf eine hohe Transparenz und eine leichte Implementierbarkeit viel Wert gelegt, ohne die Komplexität der Vorhersage von Rezessionen unter Echtzeitbedingungen aus den Augen zu verlieren.

Die Tatsache, dass sich die daraus resultierenden Rezessionswahrscheinlichkeiten aus der Kombination von verschiedenen dynamischen Probit-Regressionsgleichungen ergeben, ist ein Aspekt, der insbesondere unter Echtzeitbedingungen von großem Vorteil ist.

Tabelle 8.13: Out-of-sample Prognose-Evaluierung in Echtzeit

	MAE	RMSE	Theil-Koeff.
Ein-Monats-Prognosehorizont			
EQ_YC1G	0,3126	0,4644	0,5001
EQ_YC1S	0,2807	0,4396	0,4655
EQ_YC2G	0,2236	0,3378	0,3391
EQ_YC2S	0,2377	0,3497	0,3514
EQ_YC3G	0,2235	0,3183	0,3145
EQ_YC3S	0,2121	0,3040	0,2999
EQ_YC5G	0,2521	0,3545	0,3460
EQ_YC5S	0,2230	0,3154	0,3153
EQ_YC10G	0,2630	0,3772	0,3724
EQ_YC10S	0,2280	0,3161	0,3003
AVG_PRB	0,2456	0,3339	0,3415
Zwei-Monats-Prognosehorizont			
EQ_YC1G	0,3588	0,4956	0,5647
EQ_YC1S	0,3276	0,4606	0,4850
EQ_YC2G	0,3524	0,4596	0,4858
EQ_YC2S	0,4076	0,5260	0,5635
EQ_YC3G	0,3209	0,4251	0,4414
EQ_YC3S	0,2985	0,4016	0,4144
EQ_YC5G	0,2387	0,3389	0,3538
EQ_YC5S	0,2993	0,4080	0,4246
EQ_YC10G	0,2575	0,3454	0,3450
EQ_YC10S	0,2766	0,3778	0,3585
AVG_PRB	0,3138	0,3929	0,4236
Drei-Monats-Prognosehorizont			
EQ_YC1G	0,2635	0,3801	0,4619
EQ_YC1S	0,1617	0,3015	0,3792
EQ_YC2G	0,1729	0,3111	0,4166
EQ_YC2S	0,2284	0,3833	0,4969
EQ_YC3G	0,1610	0,2912	0,3750
EQ_YC3S	0,2157	0,3515	0,4606
EQ_YC5G	0,2606	0,3444	0,3683
EQ_YC5S	0,2972	0,3854	0,4325
EQ_YC10G	0,2211	0,3058	0,3510
EQ_YC10S	0,2941	0,3776	0,3967
AVG_PRB	0,3447	0,4570	0,5779

Anmerkung: Anfänglicher Stützzeitraum bis zum Datenstand (Vintage) 2007m1, Prognosestichprobe Berichtsmonate 2007m2 (bzw. m3, m4) bis 2010m11 (bzw. m12; 2011m1), siehe Text.

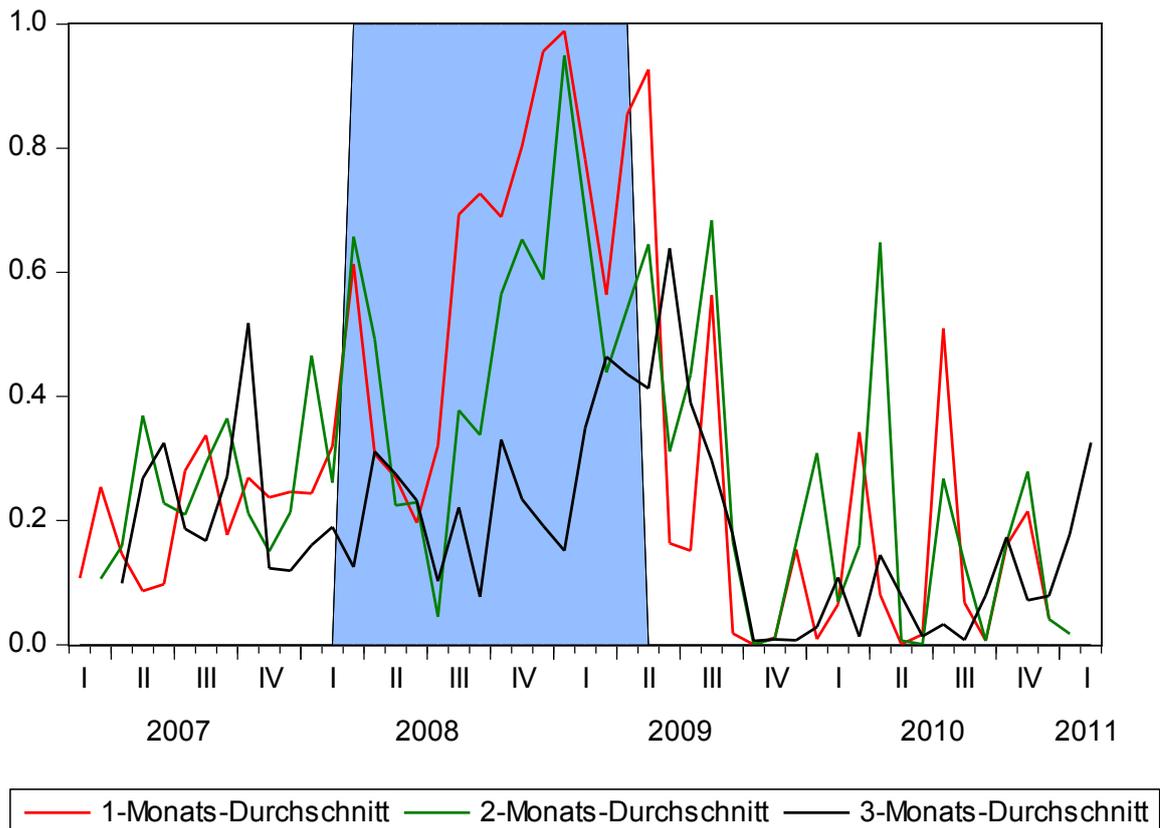


Abbildung 8.5: Echtzeit-Durchschnittsrezessionswahrscheinlichkeiten. Die Zeitachse beschreibt Berichtsmonate, die entsprechenden Publikationsperioden erhält man durch Abzug des jeweiligen Prognoseschritts (z.B. ein Signal für Berichtsmonat 2008m08 bei der 1-Monats-Prognose erhält man in 2008m07). Die schraffierte Fläche stellt den ex-post datierten Rezessionszeitraum dar (vgl. Kapitel 4).

Dies beruht hauptsächlich darauf, dass der Ansatz, aus den Rezessionswahrscheinlichkeiten verschiedener Probit-Regressionsgleichungen den Durchschnitt zu bilden, den Einfluss möglicher Ausreißer in den einzelnen Spezifikationen (welche sich auch aus der Vorläufigkeit der Daten ergeben könnte) reduziert.

Die in diesem Kapitel beschriebenen Ergebnisse zeigen darüber hinaus, dass die durchgeführten dynamischen Probitregressionen einige wichtige und interessante Erkenntnisse beitragen konnten, indem sie Information über den Nutzen verschiedener makroökonomischer Variablen und Finanzmarktindikatoren zur Rezessionsvorhersage mit verschiedenen Zeithorizonten unter Echtzeitbedingungen bereitstellen. Insbesondere sind hier nochmals die Vorhersagekraft von Aktienmarktentwicklungen für alle Vorhersagehorizonte sowie die weniger guten Ergebnisse für die Zinsspreads von Bundesanleihen verschiedener Laufzeiten zu nennen.

9 Fazit

Zunächst wurde mit den ersten Kapiteln des vorliegenden Berichts eine belastbare Grundlage gelegt für die weitere Untersuchung der zeitnahen Bestimmung von konjunkturellen Wendepunkten. Kapitel 2 stellte ausführlich die Aufbereitung der zur Verfügung stehenden Daten dar, in Kapitel 3 wurden die beobachteten Revisionshistorien mittels deskriptiver statistischer Verfahren untersucht, und Kapitel 4 beleuchtete detailliert, wie eine möglichst zeitnahe ex post-Datierung der Wendepunkte unter Echtzeitbedingungen erfolgen kann.

Einige der aus diesen Teilen vorliegenden Ergebnisse sind durchaus bemerkenswert. So stellte sich z.B. heraus, dass Erstveröffentlichungen von revisionsanfälligen Variablen anscheinend verzerrte Schätzungen sind in dem Sinne, dass die folgenden Revisionen eher in eine bestimmte Richtung gehen (im Mittel Unterschätzung von Industrieproduktion und Auftragseingängen aus dem In- und Ausland, Überschätzung der Konsumentenpreise). Andererseits ergab aber die Betrachtung der Wendepunktatierung in Echtzeit, dass die Revisionen trotzdem im Großen und Ganzen die Erkennung und Bestimmung von Rezessionen nicht beeinträchtigen.

Von zentralem Interesse sind dann die im engeren Sinne ökonometrischen Verfahren zur Wendepunktbestimmung. In den Kapiteln 5 bis 8 wurden dafür die folgenden Modellklassen angewendet: lineare Einzelgleichungen mit direkten Multischritt-Prognosen, VARs mit Parameterrestriktionen, ein Markov-Switching-Modell und schließlich dynamische Probitgleichungen. All diese Modelle wurden streng unter den Restriktionen der Datenlage in Echtzeit spezifiziert und evaluiert. Die Verwendung der ersten beiden Modellklassen ist ein Novum in der Wendepunktliteratur, da diese linearen Modelle zunächst lediglich das Wachstum der stetigen Outputgröße prognostizieren. Erst durch die Analyse der gesamten Prognoseverteilungen (Dichtefunktionen) kann dort eine Aussage über die Rezessionswahrscheinlichkeiten abgeleitet werden. Die letzten beiden Modellklassen andererseits stellen nichtlineare Ansätze dar, die durch ihre explizite Erfassung von verschiedenen Regimen sich auf natürliche Weise für die Analyse von Wendepunkten anbieten. Der Markov-Switching-Modellansatz wurde für diese Studie verfeinert, so dass die

Zahl der Regime zeitabhängig variiert werden konnte. Dadurch konnten unterschiedliche Intensitäten von Rezessionen auch in der Prognose erfasst werden.

In Tabelle 9.1 vergleichen wir die jeweils beste Spezifikation jeder untersuchten Modellklasse. In den Spalten „MAE“, „RMSE“ und „Theil“ geben wir dabei zunächst die Evaluationsmaße an, die sich auf die Gegenüberstellung mit (den vergleichbaren Berichtsperioden) der ex post-Benchmarkreihe beziehen. Dabei ist nochmals zu beachten, dass die effektiven Prognosehorizonte der verschiedenen Modelle teils unterschiedlich sind, was sich auf die Performance bei dieser Evaluation auswirken könnte. So ist der durchschnittliche effektive Prognosehorizont beim Subset-VAR-Modell bezüglich der Zweitveröffentlichung (Rev. 1) -1, d.h. ausgehend vom Prognosezeitraum wird jeweils die wirtschaftliche Aktivität des Vormonats prognostiziert (backcast – ökonomisch ist dies natürlich trotzdem eine Prognose, da die Daten selbst erst später vorliegen). Gleiches gilt für die direkten Einzelgleichungen. Beim Probit- und beim Markov-Switching-Modell ist der effektive Horizont +1. Abgesehen von diesem Interpretationsproblem erreicht bei den Evaluationsstatistiken das Subset-VAR-Modell die besten Werte.

Als zweite wichtige Dimension des Modellvergleichs bieten sich die Zeitpunkte der Rezessions- und Aufschwungssignale an, die die verschiedenen Modellklassen liefern. Ein vollständiger Überblick über die Entwicklung der Signale aller Modellklassen im gesamten Evaluierungszeitraum ist in Abbildung 9.1 enthalten. Dort sind die prognostizierten Rezessionswahrscheinlichkeiten der jeweils besten Spezifikation jeder Modellklasse nochmals abgebildet, aber nun für die Publikationsperioden (Datenstände), in denen die Prognosen in Echtzeit erstellt werden konnten. Ebenfalls eingetragen ist die ex post-Rezessionsdatierung, aber nunmehr nicht bezogen auf die betroffenen Berichtsmonate, sondern ebenfalls auf die Publikationsperioden, in denen die Signale gemeldet wurden. Ein Teil der Information der Abbildung 9.1 ist auch noch in der Tabelle 9.1 (Spalte 5) ausgeführt, nämlich in welcher Publikationsperiode die verschiedenen Modellansätze jeweils zum ersten Mal die herannahende Rezession erkannten. Wie in der letzten Zeile zum Vergleich wiederholt wird, konnte auf Grund der veröffentlichten Daten im Oktober 2008 eine Rezession ex post festgestellt werden konnte, wenn man das im Kapitel 4 erläuterte Verfahren der Datierung anwendet.¹

¹In den Kapiteln 5 und 6 wurde als Hilfsmittel eine vereinfachte Definition angewendet, nach der die veröffentlichten Daten bereits im August 2008 die Eigenschaft einer Rezession aufweisen. Allerdings hat das in Kapitel 4 erläuterte flexible Verfahren u.a. den Vorteil, dass es den zwischenzeitlichen Anstieg des Outputs Ende 2008 nicht als Ende dieser Rezession darstellt. Daher erfolgt der Vergleich hier mit der zentralen ex post-Datierung aus Kapitel 4.

Tabelle 9.1: Modellvergleich

Modell (Spezifikation)	MAE	RMSE	Theil	Rez'signal zum Datenstand	implizierter Wendepunkt (Berichtsmonat)
Subset-VAR (Durchschnitt Rev. 1)	0,1569	0,3256	0,3156	2008m6	ab/nach 2007m12
Probit (1-Schritt/t+1, Durchschn.)	0,2456	0,3339	0,3415	2008m7 (2008m2)	2008m8 (2008m3)
Einzelgleichungen (1-Prognoseschritt/t-1)	0,1619	0,3684	0,3500	2008m8	ab/nach 2008m2
Markov-Switching (Durchschnitt)	0,2948	0,3898	0,4060	2008m7	2008m8
nachrichtlich: Benchmark ex post-Datierung	–	–	–	2008m10	2008m3

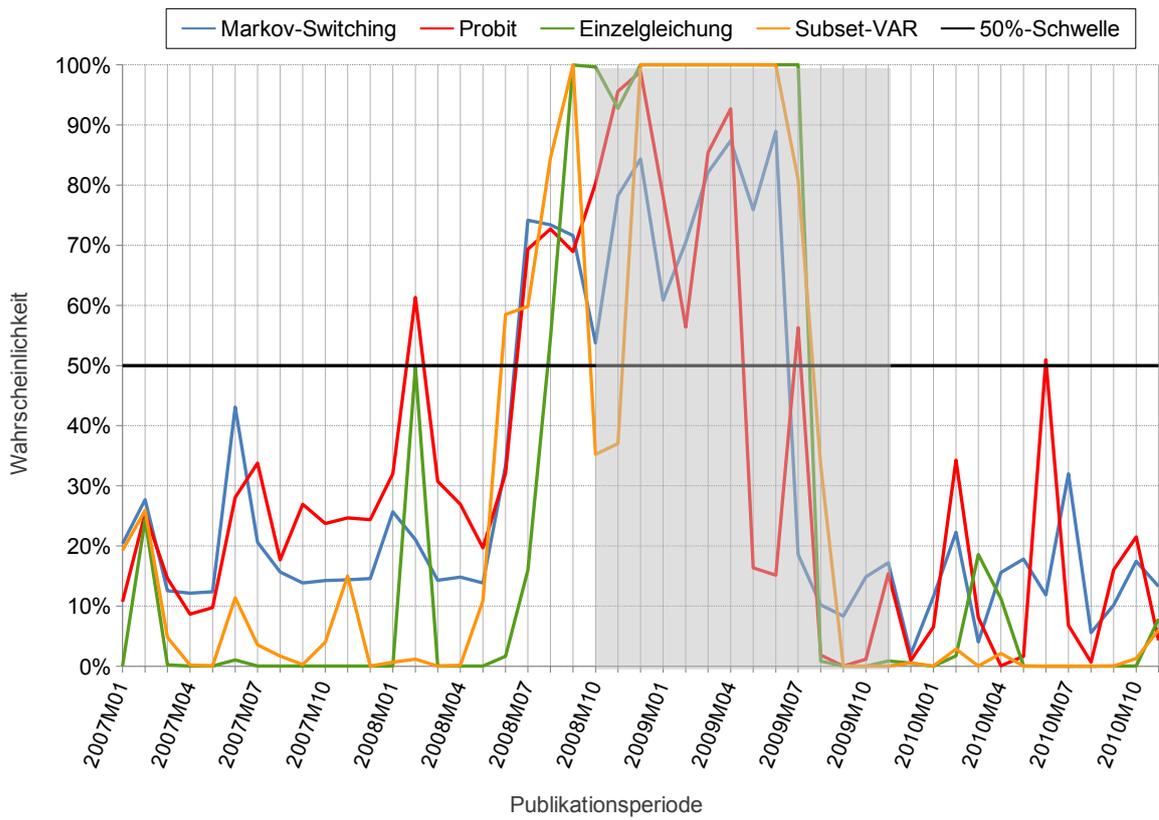
Anmerkungen: Die Datumsangaben in Klammern beim Probitmodell beziehen sich auf die isolierte Spitze, die jedoch kein nachhaltiges Signal liefert (s. Abb. 8.5). „Durchschnitt“ bezieht sich teils auf unterschiedliche Prognosehorizonte (Subset-VAR), teils auf unterschiedliche Spezifikationen zum gleichen Horizont (Probit, Markov-Switching). „1-Schritt“ bedeutet entweder eine Prognose für den effektiven Berichtsmonat nach dem Monat der Datenveröffentlichung (Probitmodell), oder eine Prognose für den ersten Berichtsmonat, für den noch keine veröffentlichten Daten vorliegen (Einzelgleichungen). MAE = mittlerer absoluter Fehler, RMSE = Wurzel des mittleren quadratischen Fehlers, Theil = Theil'scher Koeffizient.

Sowohl in Abbildung 9.1 als auch in Tabelle 9.1 (Spalte 5) lässt sich ablesen, dass allgemein die angewendeten Prognoseverfahren in Echtzeit einen Vorlauf von zwei bis vier Monaten liefern, zumindest im Evaluationsfalle dieser letzten Rezession.² Eine Ausnahme stellt teilweise das Probitmodell dar, dessen erstes Rezessionssignal tatsächlich bereits im Februar 2008 gemeldet wurde. Allerdings ist auch deutlich zu erkennen, dass direkt nach diesem einmaligen Signal für einige Monate keine Rezession mehr angezeigt wurde. Insofern ist dieses Signal wohl als etwas zwiespältig anzusehen. Aber auch die Signale des ansonsten gut abschneidenden Subset-VAR-Modells weisen das Problem auf, dass sie gegen Ende 2008 zwischenzeitlich wieder unter die Rezessions-Wahrscheinlichkeitsschwelle von 50% fallen. Je nachdem wie stark man diese verschiedenen Teilschwächen gewichtet, gewinnt also entweder die Subset-VAR-Methode oder das dynamische Probitmodell diesen Vergleich.

Des Weiteren gibt Tabelle 9.1 für jedes Modell den impliziten Monat an, in dem laut aufgetretenem Signal der Wendepunkt stattgefunden hat. (Hierbei handelt es sich um die effektiven Berichtsperioden, nicht um die Publikationszeitpunkte.) Das zeitliche Verhältnis von Rezessionssignal und impliziertem Wendepunkt hängt ab vom jeweiligen Prognosehorizont und ggf. vom Publikationslag. Bei den linearen Modellklassen (Einzelgleichungen und VARs) ist ohne weiteres nicht exakt anzugeben, welcher Monat innerhalb des jeweils zugrunde gelegten Fünfmonatszeitraums genau den Wendepunkt darstellt; hier nennt die Tabelle daher den Beginn des entsprechenden Fünfmonatszeitraums als frühestmögliche Wendepunktperiode. Es ist zu erkennen, dass die explizit unter dem Rezessionserkennungslag „leidenden“ linearen Methoden dies in diesem Fall dadurch zu kompensieren scheinen, dass sie die Wendepunkte vergleichsweise früh datieren. Jedoch fällt keine der implizierten Wendepunktperioden eklatant aus dem Rahmen, wenn man bedenkt, dass auch noch ex post die Wendepunkt datierung kontrovers diskutiert wird und nicht abschließend geklärt ist.

Schließlich wenden wir uns der Frage zu, die im Untertitel dieses Forschungsauftrags steht: Welche Indikatoren geben frühzeitig zuverlässige Signale? Insgesamt müssen wir aus den empirischen Untersuchungen dieser Studie schließen, dass sich kein einzelner Indikator und keine Indikatorgruppe besonders hervorheben lässt. Sowohl realwirtschaftliche Indikatoren wie solche der Finanzmärkte, „harte“ Messdaten und „weiche“ Umfra-

²Bei der Interpretation der Evaluierung ist generell Vorsicht geboten, denn es liegen auch hier nur relativ kurze Zeitreihen vor, in denen die Wendepunkt ereignisse relativ selten sind. Teilweise könnten die Ergebnisse auch zufällig sein und sich bei anderen Datenkonstellationen anders darstellen.



Die Phase der Rezessionssignale, die durch das nicht-parametrische Datierungsverfahren mit veröffentlichten Daten festgestellt werden, ist grau unterlegt.

Abbildung 9.1: Vergleich der Rezessionssignale bezogen auf die Publikationsperiode.

geergebnisse, verzögert und sofort verfügbare Variablen, revisionsanfällige Größen und feststehende Daten, all diese zur Verfügung stehenden Indikatoren spielen grundsätzlich eine Rolle bei der Bestimmung der wirtschaftlichen Entwicklung und ihrer Wendepunkte. Innerhalb der verschiedenen betrachteten Modellklassen entstanden natürlich leichte Nuancierungen bezüglich mancher Indikatoren. Ein einheitliches Bild oder insgesamt belastbare Unterschiede zwischen den Indikatorgruppen ergaben sich daraus allerdings nicht. Dies erscheint durchaus plausibel, da für die Studie von vornherein solche Indikatoren in Betracht gezogen wurden, die in der einschlägigen Literatur als vorlaufend charakterisiert werden. Auch für andere Prognoseprobleme der Makroökonomie ist die Nutzung einer breiten Informationsbasis das mittlerweile übliche Verfahren. Gemäß unseren Forschungsergebnissen ist es daher unabhängig von der konkret gewählten ökonometrischen Methodik für praktische Anwendungen der Wendepunktbestimmung sinnvoll, viele Indikatoren zu kombinieren.

Literaturverzeichnis

- Aiolfi, M., Capistrán, C. & Timmermann, A. (2010), Forecast combinations, CREATES Research Paper 21, School of Economics and Management, Aarhus University.
- Ang, A., Piazzesi, M. & Wei, M. (2006), ‘What does the yield curve tell us about GDP growth?’, *Journal of Econometrics* **131**, 359–403.
- Banbura, M., Giannone, D. & Reichlin, L. (2010), ‘Large bayesian vector auto regressions’, *Journal of Applied Econometrics* **25**(1), 71–92.
URL: <http://ideas.repec.org/a/jae/japmet/v25y2010i1p71-92.html>
- Benner, J. & Meier, C. (2004), ‘Prognosegüte alternativer Frühindikatoren für die Konjunktur in Deutschland’, *Jahrbücher für Nationalökonomie und Statistik* **224**(6), 637–652.
- Bernanke, B. (1990), ‘On the predictive power of interest rates and interest rate spreads’, *New England Economic Review* **Nov./Dec.**, 51–68.
- Biggs, M., Mayer, T. & Pick, A. (2009), Credit and economic recovery, Working Paper 218, De Nederlandsche Bank.
- Brock, W., Dechert, W., Scheinkman, J. & LeBaron, B. (1996), ‘A test for independence based on the correlation dimension’, *Econometric Reviews* **15**, 197–235.
- Bry, G. & Boschan, C. (1971), *Cyclical analysis of time series: Selected procedures and computer programs*, National Bureau of Economic Research, Inc.
- Burns, A. & Mitchell, W. (1946), *Measuring business cycles*, Columbia University Press.
- Business Cycle Dating Committee (2008), Determination of the 2007 peak in economic activity, Technical report, National Bureau of Economic Research.
URL: <http://www.nber.org/cycles/dec2008.html>

- Campos, J., Ericsson, N. R. & Hendry, D. (2005), General-to-specific modeling: An overview and selected bibliography, International Finance Discussion Papers 838, Boards of Governors of the Federal Reserve System.
- Carrasco, M., Hu, L. & Ploberger, W. (2005), Optimal test for Markov switching.
URL: <http://www.econ.sinica.edu.tw/upload/file/20060725.pdf>
- Carstensen, K., Wohlrabe, K. & Ziegler, C. (2011), 'Predictive ability of business cycle indicators under test – a case study for the euro area industrial production', *Jahrbücher für Nationalökonomie und Statistik* **231**(1), 82–106.
- Chauvet, M. & Piger, J. (2008), 'A comparison of the real time performance of business cycle dating methods', *Journal of Business and Economic Statistics* **26**, 42–49.
- Chauvet, M. & Potter, S. (2005), 'Forecasting recessions using the yield curve', *Journal of Forecasting* **24**, 77–103.
- Clements, M. & Hendry, D. F. (1999), *Forecasting non-stationary economic time series*, MIT Press.
- Cottrell, A. & Lucchetti, R. (2010), *Gretl User's Guide*. Version 1.9; available at <http://gretl.sourceforge.net/#man>.
URL: <http://gretl.sourceforge.net/#man>
- Darne, O. & Ferrara, L. (2009), Identification of slowdowns and accelerations for the euro area economy, Documents de Travail 239, Banque de France.
- de Bondt, G. & Hahn, E. (2010), 'Predicting recessions and recoveries in real time. The euro area-wide leading indicator (ALI)', *ECB Working Paper* (1246).
- Deutsche Bundesbank (1999), Der Übergang vom Saisonbereinigungsverfahren Census X-11 zu Census X-12-Arima, Technical report. Monatsbericht September.
- Deutsche Bundesbank (2009), 'Metainformationen zur Zeitreihe Produktion im Produzierenden Gewerbe'.
URL: <http://www.bundesbank.de/download/statistik/realtimedaten/methnotes/PD>
- Deutsche Bundesbank (2010a), 'MFI Balances Sheets Deutsche Bundesbank'.
URL: http://www.bundesbank.de/statistik/statistik_eszb_tabellen.php

Deutsche Bundesbank (2010b), 'Realtime Data Deutsche Bundesbank'.

URL: http://www.bundesbank.de/statistik/statistik_realtime.php

Deutsche Bundesbank (2011), 'Finanzmarktdaten Deutsche Bundesbank'.

URL: http://www.bundesbank.de/statistik/statistik_zinsen.php

Dickey, D. A. & Fuller, W. A. (1979), 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American Statistical Association* **74**, 427–431.

Dickey, D. A. & Fuller, W. A. (1981), 'Likelihood ratio statistics for autoregressive time series with a unit root', *Econometrica* **49**, 1057–1072.

Diebold, F. X. & Rudebusch, G. D. (1991), 'Forecasting output with the Composite Leading Index: A real-time analysis', *Journal of the American Statistical Association* **86**, 603–610.

Drechsel, K. & Scheufele, R. (2010), 'Should we trust in leading indicators? Evidence from the recent recession', *IWH Discussion Paper* (10-10).

Drechsel, K. & Scheufele, R. (2011), 'The performance of short-term forecasts of the German economy before and during the 2008/2009 recession', *International Journal of Forecasting im Druck*.

Dueker, M. (1997), Strengthening the case for the yield curve as a predictor of U.S. recessions, Review, Federal Reserve Bank of St. Louis.

Durland, J. M. & McCurdy, T. H. (1994), 'Duration-dependent transitions in a Markov model of U.S. GNP growth', *Journal of Business and Economic Statistics* **1**(3), 279–288.

Estrella, A. & Hardouvelis, G. A. (1991), 'The term structure as a predictor of real economic activity', *Journal of Finance* **46**, 555–576.

Estrella, A. & Mishkin, F. S. (1998), 'Predicting U.S. recessions: Financial variables as leading indicators', *Review of Economics and Statistics* **80**(1), 45–61.

Estrella, A., Rodrigues, A. P. & Schich, S. (2003), 'How stable is the predictive power of the yield curve? Evidence from Germany and the United States', *Review of Economics and Statistics* **85**(3), 629–644.

- Fichtner, F., Ruffer, R. & Schnatz, B. (2009), 'Leading indicators in a globalised world', *ECB Working Paper*.
- Forni, M., Hallin, M., Lippi, M. & Reichlin, L. (1999), 'The generalized dynamic factor model: Identification and estimation', *CEPR Discussion Paper* (2338).
- Friedman, B. M. & Kuttner, K. N. (1992), 'Money, income, prices and interest rates', *American Economic Review* **82**(3), 472–492.
- Friedman, M. (1993), 'The plucking model of business fluctuations revisited', *Economic Inquiry* **31**(2), 171–177.
- Fritsche, U. & Stephan, S. (2002), 'Leading indicators of German business cycles. an assessment of properties', *Jahrbücher für Nationalökonomie und Statistik* **222**(3), 289–315.
- Fuller, W. A. (1976), *Introduction to statistical time series*, Wiley, New York.
- Garcia, R. (1998), 'Asymptotic null distribution of the likelihood ratio test in Markov switching models', *International Economic Review* **39**, 763–788.
- Geweke, J. (1977), The dynamic factor analysis of economic time series, in 'Latent variables in socio-economic models', A.S. Goldberger, ed. North-Holland, Amsterdam.
- Granger, C. (1969), 'Investigating causal relations by econometric models and cross-spectral methods', *Econometrica* **37**(3), 424–438.
- Haltmaier, J. (2008), Predicting cycles in economic activity, International Finance Discussion Papers 926, Board of Governors of the Federal Reserve System, Washington D.C.
- Hamilton, J. D. (1989), 'A new approach to the economic analysis of nonstationary time series and the business cycle', *Econometrica* **57**(2), 357–384.
- Hamilton, J. D. (1990), 'Analysis of time series subject to changes in regime', *Journal of Econometrics* **45**, 39–70.
- Hamilton, J. D. (1994), *Time series analysis*, Princeton University Press.
- Hamilton, J. D. (2010), Calling recessions in real time. Preliminary.
URL: <http://dss.ucsd.edu/jhamilto/realtime.pdf>

- Hansen, B. E. (1996), 'Inference when a nuisance parameter is not identified under the null hypothesis', *Econometrica* **64**, 413–430.
- Harding, D. & Pagan, A. (2002), 'Dissecting the cycle: A methodological investigation', *Journal of Monetary Economics* **49**(2), 365–381.
- Harvey, C. (1989), 'Forecasts of economic growth from the bond and the stock markets', *Financial Analysts Journal* **Sept./Oct.**, 38–45.
- Hinze, J. (2003), 'Prognoseleistung von Frühindikatoren. Die Bedeutung von Frühindikatoren für Konjunkturprognosen – eine Analyse für Deutschland', *HWWA Discussion Paper* (236).
- ifo Institut für Wirtschaftsforschung (2011), 'ifo Geschäftsklimaindex'.
URL: <http://www.cesifo-group.de>
- Kauppi, H. & Saikkonen, P. (2008), 'Predicting U.S. recessions with dynamic binary response models', *The Review of Economics and Statistics* **90**(4), 777–791.
- Kim, C. (1994), 'Dynamic linear models with Markov-switching', *Journal of Econometrics* **60**, 1–22.
- Knetsch, T. & Reimers, H.-E. (2006), How to treat benchmark revisions? The case of German production and orders statistics, Discussion Paper 38, Deutsche Bundesbank.
- Krolzig, H. M. (1997), *Markov switching vector autoregressions. Statistical inference and application to business cycle analysis*, Springer Science+Business Media.
- Levanon, G. (2010), 'Evaluating and comparing leading and coincident economic indicators', *Business Economics* **45**(1), 16–27.
- Meeks, R. (2011), 'Do credit shocks drive output fluctuations? Evidence from corporate spreads and defaults'. Revised draft.
- Moneta, F. (2005), 'Does the yield spread predict recessions in the euro area', *International Finance* **8**, 263–301.
- NBER (2010), 'US business cycle expansions and contractions'.
URL: <http://www.nber.org/cycles.html>

- Neftci, S. N. (1982), 'Optimal prediction of cyclical downturns', *Journal of Economic Dynamics and Control* **4**, 225 – 241.
- Newey, W. & West, K. (1987), 'A simple, positive, semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix', *Econometrica* **55**(3), 703–708.
- Nyberg, H. (2010), 'Dynamic probit models and financial variables in recession forecasting', *Journal of Forecasting* **29**(1-2), 215–230.
- OECD (2010), 'Realtime Data OECD'.
URL: <http://stats.oecd.org/mei/default.asp?rev=1>
- Perlin, M. (2009), MS_Regress – A package for Markov regime switching models in MATLAB. MATLAB Central File Exchange, Available in <http://www.mathworks.com/matlabcentral/fileexchange/authors/21596>.
- Proaño, C. (2010), Recession forecasting with dynamic probit models under real time conditions, IMK Working Paper 10-2010, IMK at the Hans Boeckler Foundation, Macroeconomic Policy Institute.
- Rietzler, K. (2001), Erstellung und Anwendung eines Kriterienkatalogs zur Beurteilung der Eignung verschiedener Saisonbereinigungsverfahren zur gesamtwirtschaftlichen Konjunkturdiagnose in der Europäischen Union, Technical report, DIW Berlin.
- Robinsonov, N. & Wohlrabe, K. (2008), 'Freedom of choice in macroeconomic forecasting: An illustration with German industrial production and linear models', *Ifo Working Paper* (57).
- Rudebusch, G. D. & Williams, J. C. (2009), 'Forecasting recessions: The puzzle of the enduring power of the yield curve', *Journal of Business and Economic Statistics* **27**(4), 492–503.
- Sargent, T. & Sims, C. (1977), Business cycle modeling without pretending to have too much a priori economic theory, in 'New methods in business cycle research', C. Sims, ed. Federal Reserve Bank of Minneapolis.
- Schirwitz, B. (2009), 'A comprehensive German business cycle chronology', *Empirical Economics* **37**(2), 287–301.

- Schumacher, C. & Breitung, J. (2008), 'Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data', *International Journal of Forecasting* **24**(3), 386–398.
- Stark, T. & Croushore, D. (2002), 'Forecasting with a real-time data set for macroeconomists', *Journal of Macroeconomics* **24**(4), 507–531.
- Stock, J. H. & Watson, M. W. (1989), New indexes of coincident and leading economic indicators, in 'NBER Macroeconomics Annual', Vol. 4, O. Blanchard and S. Fischer, eds. The MIT Press, Cambridge, MA.
- Stock, J. H. & Watson, M. W. (1993), A procedure for predicting recessions with leading indicators: Econometric issues and recent experience, in 'Business Cycles, Indicators and Forecasting', J. Stock and M. Watson, eds. NBER Book Series Studies in Business Cycles.
- Timmermann, A. (2006), Forecast combinations, in 'Handbook of Economic Forecasting', Vol. 1, G. Elliot, C. Granger and A. Timmermann, eds., pp. 135–196. North-Holland, Amsterdam.
- Wright, J. (2006), The yield curve and predicting recessions, Technical report, Division of Monetary Affairs, Federal Reserve Board.
- Zentrum für Europäische Wirtschaftsforschung (2011), 'ZEW-Konjunkturerwartungen'.
URL: <http://www.zew.de/de/publikationen/Konjunkturerwartungen/Konjunkturerwartungen.php3>

Publisher: Hans-Böckler-Stiftung, Hans-Böckler-Str. 39, 40476 Düsseldorf, Germany
Phone: +49-211-7778-331, IMK@boeckler.de, <http://www.imk-boeckler.de>

IMK Study is an online publication series available at:
http://www.boeckler.de/imk_5023.htm

ISSN: 1861-2180

The views expressed in this paper do not necessarily reflect those of the IMK or the Hans-Böckler-Foundation.

All rights reserved. Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.

**Hans Böckler
Stiftung** 

Fakten für eine faire Arbeitswelt.
