

Universität Hohenheim
Institut für Volkswirtschaftslehre
Fachgebiet Statistik und Ökonometrie I

Steuerschätzung und Analyse der Prognosegüte für die Bundesrepublik Deutschland

Dissertation
zur Erlangung des Grades eines
Doktors der Wirtschaftswissenschaften
(Dr. oec.)

vorgelegt der
Fakultät Wirtschafts- und Sozialwissenschaften
der Universität Hohenheim

von
Dipl.-Volkswirtin Ulrike Berberich
Stuttgart-Hohenheim

2012

Die vorliegende Arbeit wurde im Jahr 2012 von der Fakultät Wirtschafts- und Sozialwissenschaften der Universität Hohenheim als Dissertation zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaften (Dr. oec.) angenommen.

Datum der mündlichen Doktorprüfung: 26. Juli 2012

Dekan: Prof Dr. Dirk Hachmeister

Prüfungsvorsitz: Prof. Dr. Robert Jung

Erstgutachter: Prof. Dr. Gerhard Wagenhals

Zweitgutachterin: Prof. Dr. Nadine Riedel

Danksagung

Mein besonderer Dank gilt meinem Doktorvater Herrn Prof. Dr. Gerhard Wagenhals, der mich nicht nur fachlich sondern auch persönlich all die Jahre hindurch sehr unterstützt hat. Frau Prof. Dr. Nadine Riedel danke ich für ihre kritischen Anmerkungen und das Erstellen des Zweitgutachtens.

Außerdem bedanke ich mich bei Uli Scheurle sowohl für die kritischen fachlichen Kommentare und Anregungen zu dieser Arbeit, als auch für die interessanten persönlichen Gespräche während unserer zahlreichen gemeinsamen Fahrten von Tübingen nach Hohenheim. Mein herzlicher Dank gilt auch all meinen anderen Kollegen des Lehrstuhls für Statistik und Ökonometrie I, vor allem Martina Rabe, Wolf Dieter Heinbach, Stefanie Schröpfer, Eva Schlenker, Frauke Wolf und Harald Weiß, die mir durch ein sehr angenehmes Arbeitsklima und ein stets offenes Ohr besonders auch in schwierigen Phasen einen guten Rückhalt boten.

Ebenso danke ich unseren ehemaligen wissenschaftlichen Hilfskräften Eva Schlenker, Stefanie Seifert und Markus Spindler. Sie haben in mühevoller Arbeit die vom Statistischen Bundesamt in Papierform bereit gestellten Steuerdaten in Excel eingegeben. Außerdem bedanke ich mich bei allen anderen ehemaligen und aktuellen wissenschaftlichen Hilfskräften, die mich im Laufe der Zeit mit Recherchen, Latex-Hilfestellungen und Korrekturlesen unterstützt haben, insbesondere bei Claudia Ehnert.

Mein größter Dank gilt meiner Familie und meinen Freunden. Sie haben mir in jeder Lebenslage stets Rückhalt und moralische Unterstützung geboten. In besonderem Maße danke ich dabei meiner Tante Ingrid, der der Abschluss meiner Dissertation so sehr am Herzen lag, dass sie mir an ihrem Sterbebett das Versprechen abnahm, nun nicht mehr abzubereiten. Diese Arbeit ist ihr gewidmet.

Tübingen, Oktober 2012

Ulrike Berberich

Inhaltsverzeichnis

| | |
|---|-----------|
| Tabellenverzeichnis | III |
| Abbildungsverzeichnis | XII |
| Abkürzungsverzeichnis | XIV |
| Verzeichnis der wichtigsten Symbole | XVI |
| 1. Einführung | 1 |
| 1.1. Problemstellung | 1 |
| 1.2. Aufbau der Arbeit | 4 |
| 2. Grundlagen | 5 |
| 2.1. Der Arbeitskreis „Steuerschätzungen“ | 5 |
| 2.1.1. Der institutionelle Rahmen des AKS | 5 |
| 2.1.2. Vorgehensweise des AKS | 6 |
| 2.2. Methoden der Steuerschätzung und damit verbundene Schwierigkeiten | 10 |
| 2.2.1. Die Determinanten der Präzision der Steuerprognose | 10 |
| 2.2.2. Überblick über die Methoden der Steueraufkommensprognose | 15 |
| 2.2.3. Schwierigkeiten bei der Steuerschätzung | 19 |
| 2.3. Bisherige empirische Untersuchungen zur Qualität der AKS-Prognosen | 25 |
| 3. Eigene Schätzung des Steueraufkommens mit (S)ARIMA-Modellen | 33 |
| 3.1. Saisonale ARIMA-Modelle | 34 |
| 3.2. Datengrundlage und Datenaufbereitung | 41 |
| 3.3. Schätzung des Steueraufkommens anhand der Box-Jenkins-Methode . | 42 |
| 3.3.1. Phase I: Identifikation | 43 |
| 3.3.2. Phase II: Schätzen und Testen | 54 |
| 3.3.3. Phase III: Prognose als Anwendung | 62 |

| | |
|---|------------|
| 4. Analyse der Prognosegüte | 68 |
| 4.1. Modellansatz zur Untersuchung der Prognoserationalität | 68 |
| 4.1.1. Test auf Unverzerrtheit | 71 |
| 4.1.2. Test auf Informationseffizienz | 81 |
| 4.2. Untersuchung der Prognoserationalität der SARIMA-Schätzungen . . | 89 |
| 4.2.1. Test auf Unverzerrtheit der SARIMA-Prognosen | 89 |
| 4.2.2. Test auf Informationseffizienz der SARIMA-Prognosen | 98 |
| 4.2.3. Fazit zur Prognoserationalität der SARIMA-Schätzungen . . . | 113 |
| 4.3. Untersuchung der Prognoserationalität der AKS-Schätzungen | 115 |
| 4.3.1. Test auf Unverzerrtheit der AKS-Prognosen | 116 |
| 4.3.2. Test auf Informationseffizienz der AKS-Prognosen | 123 |
| 4.3.3. Fazit zur Prognoserationalität der AKS-Schätzungen | 134 |
| 4.4. Vergleich der Prognosegenauigkeit anhand deskriptiver Gütemaße . . | 137 |
| 4.4.1. Vergleich bei gleichem Informationsstand zum möglichst ähn- | |
| lichen Prognosezeitpunkt | 141 |
| 4.4.2. Vergleich unter Verwendung aller verfügbarer Informationen | |
| zu allen Prognosezeitpunkten | 151 |
| 4.4.3. Fazit zum Vergleich der Prognosegenauigkeit anhand deskrip- | |
| tiver Gütemaße | 165 |
| 5. Fazit und Ausblick | 167 |
| A. Ergebnistabellen und Schaubilder zu den SARIMA-Schätzungen | 177 |
| B. Ergebnistabellen zur Analyse der SARIMA-Prognosen | 180 |
| C. Ergebnistabellen zur Analyse der AKS-Prognosen | 197 |
| D. Ergebnistabellen zur deskriptiven Analyse der Prognosegüte | 228 |
| Literaturverzeichnis | 232 |

Tabellenverzeichnis

| | |
|--|-----|
| 4.1. Ergebnisse des Tests auf Unverzerrtheit der Prognosen anhand des bevorzugten Modells mit $H = 4$ für Variante I | 90 |
| 4.2. Ergebnisse des Tests auf Unverzerrtheit der Prognosen anhand des bevorzugten Modells mit $H = 4$ für Variante II | 91 |
| 4.3. Ergebnisse des Tests auf Unverzerrtheit der Prognosen anhand des bevorzugten Modells mit $H = 8$ für Variante I | 92 |
| 4.4. Ergebnisse des Tests auf Unverzerrtheit der Prognosen anhand des bevorzugten Modells mit $H = 8$ für Variante II | 93 |
| 4.5. Ergebnisse des Tests auf Unverzerrtheit der Prognosen anhand des bevorzugten Modells mit $H = 12$ für Variante I | 94 |
| 4.6. Ergebnisse des Tests auf Unverzerrtheit der Prognosen anhand des bevorzugten Modells mit $H = 12$ für Variante II | 95 |
| 4.7. Ergebnisse des Tests auf Unverzerrtheit der Prognosen anhand des bevorzugten Modells mit $H = 16$ für Variante I | 97 |
| 4.8. Ergebnisse des Tests auf Unverzerrtheit der Prognosen anhand des bevorzugten Modells mit $H = 16$ für Variante II | 98 |
| 4.9. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 1 der Prognosen mit $H = 4$ anhand des bevorzugten Modells . . | 99 |
| 4.10. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 1 der Prognosen mit $H = 4$ anhand des bevorzugten Modells . . | 100 |
| 4.11. Ergebnisse des separaten Tests auf Informationseffizienz bezüglich Fall 2 der Prognosen mit $H = 4$ anhand des bevorzugten Modells . . | 102 |

| | |
|--|-----|
| 4.12. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 1 der Prognosen mit $H = 8$ anhand des bevorzugten Modells . . . | 103 |
| 4.13. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 1 der Prognosen mit $H = 8$ anhand des bevorzugten Modells . . . | 104 |
| 4.14. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 2 der Prognosen mit $H = 8$ anhand des bevorzugten Modells . . . | 105 |
| 4.15. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 2 der Prognosen mit $H = 8$ anhand des bevorzugten Modells . . . | 107 |
| 4.16. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 1 der Prognosen mit $H = 12$ anhand des bevorzugten Modells . . . | 108 |
| 4.17. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 1 der Prognosen mit $H = 12$ anhand des bevorzugten Modells . . . | 109 |
| 4.18. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 2 der Prognosen mit $H = 12$ anhand des bevorzugten Modells . . . | 111 |
| 4.19. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 2 der Prognosen mit $H = 12$ anhand des bevorzugten Modells . . . | 112 |
| 4.20. Ergebnisse des Tests auf Unverzerrtheit der Prognosen des Arbeits- kreises „Steuerschätzungen“ mit $H = 5$ für Variante I | 117 |
| 4.21. Regressionsergebnisse der horizont-spezifischen Verzerrungen $\hat{\alpha}_h$ der AKS-Prognosen mit $H = 5$ | 118 |
| 4.22. Ergebnisse des Tests auf Unverzerrtheit der Prognosen des Arbeits- kreises „Steuerschätzungen“ mit $H = 5$ für Variante II | 119 |
| 4.23. Ergebnisse des Tests auf Unverzerrtheit der Prognosen des Arbeits- kreises „Steuerschätzungen“ mit $H = 4$ für Variante I | 120 |
| 4.24. Ergebnisse des Tests auf Unverzerrtheit der Prognosen des Arbeits- kreises „Steuerschätzungen“ mit $H = 4$ für Variante II | 121 |
| 4.25. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 5$ | 124 |
| 4.26. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 5$ | 126 |

| | |
|---|---------------|
| 4.27. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 5$ ohne die Jahre 1991 bis 1995 | . 127 |
| 4.28. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 5$ ohne die Jahre 1991 bis 1995 | . 128 |
| 4.29. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 2 der AKS-Prognosen mit $H = 5$ | 130 |
| 4.30. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 2 der AKS-Prognosen mit $H = 5$ | 131 |
| 4.31. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 2 der AKS-Prognosen mit $H = 5$ ohne die Jahre 1991 bis 1995 | . 132 |
| 4.32. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 2 der AKS-Prognosen mit $H = 5$ ohne die Jahre 1991 bis 1995 | . 133 |
| 4.33. Prozentfehler PE_{th} der AKS-Prognosen | 143 |
| 4.34. Horizontale Treffsicherheitsanalyse gemäß Methode 1 | 144 |
| 4.35. Vertikale Treffsicherheitsanalyse gemäß Methode 1 | 147 |
| 4.36. Gepoolte Treffsicherheitsanalyse gemäß Methode 1 | 150 |
| 4.37. Horizontale Treffsicherheitsanalyse gemäß Methode 2 mit Vier-Jahres- Prognosen | 153 |
| 4.38. Horizontale Treffsicherheitsanalyse gemäß Methode 2 ohne Vier-Jahres- Prognosen | 156 |
| 4.39. Vertikale Treffsicherheitsanalyse gemäß Methode 2 | 159 |
| 4.40. Gepoolte Treffsicherheitsanalyse gemäß Methode 2 mit Vier-Jahres- Prognosen | 162 |
| 4.41. Gepoolte Treffsicherheitsanalyse gemäß Methode 2 ohne Vier-Jahres- Prognosen | 164 |
| A.1. Schätzergebnisse des SARIMA(1,0,0)(1,0,1) ₄ -Modells | 177 |
| A.2. Schätzergebnisse des SARIMA(0,0,1)(1,0,1) ₄ -Modells | 178 |
| A.3. Schätzergebnisse des SARIMA(2,0,1)(1,0,1) ₄ -Modells | 178 |

| | |
|--|-----|
| B.1. Regressionsergebnisse der horizont-spezifischen Verzerrungen $\hat{\alpha}_h$ der Prognosen anhand des bevorzugten Modells mit $H = 4$ | 180 |
| B.2. Ergebnisse des Tests auf Unverzerrtheit der Prognosen anhand des Alternativmodells mit $H = 4$ für Variante I | 181 |
| B.3. Ergebnisse des Tests auf Unverzerrtheit der Prognosen anhand des Alternativmodells mit $H = 4$ für Variante II | 181 |
| B.4. Regressionsergebnisse der horizont-spezifischen Verzerrungen $\hat{\alpha}_h$ der Prognosen anhand des bevorzugten Modells mit $H = 8$ | 182 |
| B.5. Ergebnisse des Tests auf Unverzerrtheit der Prognosen anhand des Alternativmodells mit $H = 8$ für Variante I | 183 |
| B.6. Ergebnisse des Tests auf Unverzerrtheit der Prognosen anhand des Alternativmodells mit $H = 8$ für Variante II | 183 |
| B.7. Regressionsergebnisse der horizont-spezifischen Verzerrungen $\hat{\alpha}_h$ der Prognosen anhand des bevorzugten Modells mit $H = 12$ | 184 |
| B.8. Ergebnisse des Tests auf Unverzerrtheit der Prognosen anhand des Alternativmodells mit $H = 12$ für Variante I | 185 |
| B.9. Ergebnisse des Tests auf Unverzerrtheit der Prognosen anhand des Alternativmodells mit $H = 12$ für Variante II | 185 |
| B.10. Ergebnisse des Tests auf Unverzerrtheit der Prognosen anhand des Alternativmodells mit $H = 16$ für Variante I | 186 |
| B.11. Ergebnisse des Tests auf Unverzerrtheit der Prognosen anhand des Alternativmodells mit $H = 16$ für Variante II | 186 |
| B.12. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 1 der Prognosen mit $H = 4$ anhand des Alternativmodells | 187 |
| B.13. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 1 der Prognosen mit $H = 4$ anhand des Alternativmodells | 187 |
| B.14. Ergebnisse des separaten Tests auf Informationseffizienz bezüglich Fall 2 der Prognosen mit $H = 4$ anhand des Alternativmodells | 188 |
| B.15. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 1 der Prognosen mit $H = 8$ anhand des Alternativmodells | 189 |

| | |
|--|-----|
| B.16. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 1 der Prognosen mit $H = 8$ anhand des Alternativmodells | 190 |
| B.17. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 2 der Prognosen mit $H = 8$ anhand des Alternativmodells | 191 |
| B.18. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 2 der Prognosen mit $H = 8$ anhand des Alternativmodells | 192 |
| B.19. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 1 der Prognosen mit $H = 12$ anhand des Alternativmodells | 193 |
| B.20. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 1 der Prognosen mit $H = 12$ anhand des Alternativmodells | 194 |
| B.21. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 2 der Prognosen mit $H = 12$ anhand des Alternativmodells | 195 |
| B.22. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 2 der Prognosen mit $H = 12$ anhand des Alternativmodells | 196 |
| C.1. Regressionsergebnisse der horizont-spezifischen Verzerrungen $\hat{\alpha}_h$ der AKS-Prognosen mit $H = 4$ | 197 |
| C.2. Ergebnisse des Tests auf Unverzerrtheit der Prognosen des Arbeits- kreises "Steuerschätzungen" mit $H = 3$ für Variante I | 198 |
| C.3. Ergebnisse des Tests auf Unverzerrtheit der Prognosen des Arbeits- kreises "Steuerschätzungen" mit $H = 3$ für Variante II | 198 |
| C.4. Ergebnisse des Tests auf Unverzerrtheit der Prognosen des Arbeits- kreises "Steuerschätzungen" mit $H = 2$ für Variante I | 199 |
| C.5. Ergebnisse des Tests auf Unverzerrtheit der Prognosen des Arbeits- kreises "Steuerschätzungen" mit $H = 2$ für Variante II | 199 |
| C.6. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne das Jahr 1990 mit $H = 5$ für Variante I | 200 |
| C.7. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne das Jahr 1990 mit $H = 5$ für Variante II | 200 |
| C.8. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne das Jahr 1990 mit $H = 4$ für Variante I | 201 |

| | |
|---|-----|
| C.9. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne das Jahr 1990 mit $H = 4$ für Variante II | 201 |
| C.10. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne das Jahr 1990 mit $H = 3$ für Variante I | 202 |
| C.11. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne das Jahr 1990 mit $H = 3$ für Variante II | 202 |
| C.12. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne das Jahr 1990 mit $H = 2$ für Variante I | 203 |
| C.13. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne das Jahr 1990 mit $H = 2$ für Variante II | 203 |
| C.14. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne die Jahre 1991 bis 1995 mit $H = 5$ für Variante I | 204 |
| C.15. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne die Jahre 1991 bis 1995 mit $H = 5$ für Variante II | 204 |
| C.16. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne die Jahre 1991 bis 1995 mit $H = 4$ für Variante I | 205 |
| C.17. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne die Jahre 1991 bis 1995 mit $H = 4$ für Variante II | 205 |
| C.18. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne die Jahre 1991 bis 1995 mit $H = 3$ für Variante I | 206 |
| C.19. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne die Jahre 1991 bis 1995 mit $H = 3$ für Variante II | 206 |
| C.20. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne die Jahre 1991 bis 1995 mit $H = 2$ für Variante I | 207 |
| C.21. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne die Jahre 1991 bis 1995 mit $H = 2$ für Variante II | 207 |
| C.22. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne das Jahr 1991 mit $H = 5$ für Variante I | 208 |
| C.23. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne das Jahr 1991 mit $H = 5$ für Variante II | 208 |

| | |
|--|-----|
| C.24. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne das Jahr 1991 mit $H = 4$ für Variante I | 209 |
| C.25. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne das Jahr 1991 mit $H = 4$ für Variante II | 209 |
| C.26. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne das Jahr 1991 mit $H = 3$ für Variante I | 210 |
| C.27. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne das Jahr 1991 mit $H = 3$ für Variante II | 210 |
| C.28. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne das Jahr 1991 mit $H = 2$ für Variante I | 211 |
| C.29. Ergebnisse des Tests auf Unverzerrtheit der AKS-Prognosen ohne das Jahr 1991 mit $H = 2$ für Variante II | 211 |
| C.30. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 4$ | 212 |
| C.31. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 4$ | 212 |
| C.32. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 3$ | 213 |
| C.33. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 3$ | 213 |
| C.34. Ergebnis des separaten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 2$ | 214 |
| C.35. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 5$ ohne 1990 | 214 |
| C.36. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 5$ ohne 1990 | 215 |
| C.37. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 4$ ohne 1990 | 215 |
| C.38. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 4$ ohne 1990 | 216 |

| | |
|---|-----|
| C.39. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 3$ ohne 1990 | 216 |
| C.40. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 3$ ohne 1990 | 217 |
| C.41. Ergebnis des separaten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 2$ ohne 1990 | 217 |
| C.42. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 5$ ohne 1991 | 218 |
| C.43. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 5$ ohne 1991 | 218 |
| C.44. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 4$ ohne 1991 | 219 |
| C.45. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 4$ ohne 1991 | 219 |
| C.46. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 3$ ohne 1991 | 220 |
| C.47. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 3$ ohne 1991 | 220 |
| C.48. Ergebnis des separaten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 2$ ohne 1991 | 221 |
| C.49. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 4$ ohne die Jahre 1991 bis 1995 . | 221 |
| C.50. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 4$ ohne die Jahre 1991 bis 1995 . | 222 |
| C.51. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 3$ ohne die Jahre 1991 bis 1995 . | 222 |
| C.52. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 3$ ohne die Jahre 1991 bis 1995 . | 223 |
| C.53. Ergebnis des separaten Tests auf Informationseffizienz bezüglich Fall 1 der AKS-Prognosen mit $H = 2$ ohne die Jahre 1991 bis 1995 | 223 |

| | |
|---|-----|
| C.54. Ergebnis des separaten Tests auf Informationseffizienz bezüglich Fall 2 der AKS-Prognosen mit $H = 4$ | 224 |
| C.55. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 2 der AKS-Prognosen mit $H = 5$ ohne 1990 | 224 |
| C.56. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 2 der AKS-Prognosen mit $H = 5$ ohne 1990 | 225 |
| C.57. Ergebnis des separaten Tests auf Informationseffizienz bezüglich Fall 2 der AKS-Prognosen mit $H = 4$ ohne 1990 | 225 |
| C.58. Ergebnisse der separaten Tests auf Informationseffizienz bezüglich Fall 2 der AKS-Prognosen mit $H = 5$ ohne 1991 | 226 |
| C.59. Ergebnisse des gepoolten Tests auf Informationseffizienz bezüglich Fall 2 der AKS-Prognosen mit $H = 5$ ohne 1991 | 226 |
| C.60. Ergebnis des separaten Tests auf Informationseffizienz bezüglich Fall 2 der AKS-Prognosen mit $H = 4$ ohne 1991 | 227 |
| C.61. Ergebnis des separaten Tests auf Informationseffizienz bezüglich Fall 2 der AKS-Prognosen mit $H = 4$ ohne die Jahre 1991 bis 1995 | 227 |
| D.1. Prozentfehler PE_{th} der Prognosen anhand des SARIMA(1,0,0)(1,0,1) ₄ - Modells gemäß Methode 1 | 228 |
| D.2. Prozentfehler PE_{th} der Prognosen anhand des SARIMA(0,0,1)(1,0,1) ₄ - Modells gemäß Methode 1 | 229 |
| D.3. Prozentfehler PE_{th} der Prognosen anhand des SARIMA(1,0,0)(1,0,1) ₄ - Modells gemäß Methode 2 | 230 |
| D.4. Prozentfehler PE_{th} der Prognosen anhand des SARIMA(0,0,1)(1,0,1) ₄ - Modells gemäß Methode 2 | 231 |

Abbildungsverzeichnis

| | |
|---|-----|
| 2.1. Leibrechts Modell der Determinanten der Präzision (Leibrecht (2003), S. 135) | 11 |
| 2.2. Methoden der Steueraufkommensprognose (eigene Darstellung) | 16 |
| 3.1. Abbildung des Gesamtsteueraufkommens der BRD vom ersten Quartal 1961 bis zum vierten Quartal 2010 in Millionen Euro | 43 |
| 3.2. Abbildung des logarithmierten Gesamtsteueraufkommens der BRD vom ersten Quartal 1961 bis zum vierten Quartal 2010 | 45 |
| 3.3. Abbildung der Wachstumsraten des Gesamtsteueraufkommens der BRD vom ersten Quartal 1961 bis zum vierten Quartal 2010 | 49 |
| 3.4. Abbildung der empirischen Autokorrelationsfunktion der Wachstumsraten des Gesamtsteueraufkommens | 52 |
| 3.5. Abbildung der empirischen partiellen Autokorrelationsfunktion der Wachstumsraten des Gesamtsteueraufkommens | 54 |
| 3.6. Abbildung der empirischen ACF sowie der empirischen PACF des SARIMA(1,0,0)(1,0,1) ₄ -Modells | 59 |
| 3.7. Abbildung der Residuen des SARIMA(1,0,0)(1,0,1) ₄ -Modells | 61 |
| 4.1. Schematische Darstellung der Struktur der Prognosefehler (eigene Darstellung) | 76 |
| 4.2. Abbildung der Werte des MAPE _t der AKS-Prognosen sowie der Prognosen des SARIMA(1,0,0)(1,0,1) ₄ -Modells gemäß Methode 1 | 146 |

| | | |
|------|---|-----|
| 4.3. | Abbildung der Werte des MAPE_h der AKS-Prognosen sowie der Prognosen des $\text{SARIMA}(0,0,1)(1,0,1)_4$ -Modells gemäß Methode 1 | 149 |
| 4.4. | Abbildung der Werte des MAPE_t der AKS-Prognosen sowie der Prognosen des $\text{SARIMA}(1,0,0)(1,0,1)_4$ -Modells gemäß Methode 2 unter Berücksichtigung der Vier-Jahres-Prognosen | 155 |
| 4.5. | Abbildung der Werte des MAPE_t der AKS-Prognosen sowie der Prognosen des $\text{SARIMA}(1,0,0)(1,0,1)_4$ -Modells gemäß Methode 2 ohne Berücksichtigung der Vier-Jahres-Prognosen | 158 |
| 4.6. | Abbildung der Werte des MAPE_h der AKS-Prognosen sowie der Prognosen des $\text{SARIMA}(0,0,1)(1,0,1)_4$ -Modells gemäß Methode 2 | 161 |
| A.1. | Abbildung der empirischen ACF sowie der empirischen PACF des $\text{SARIMA}(0,0,1)(1,0,1)_4$ -Modells | 179 |
| A.2. | Abbildung der empirischen ACF sowie der empirischen PACF des $\text{SARIMA}(2,0,1)(1,0,1)_4$ -Modells | 179 |

Abkürzungsverzeichnis

| | |
|----------------------------------|--|
| ACF | Autokorrelationsfunktion |
| ADF-Test | Augmented-Dickey-Fuller-Test |
| AEUV | Vertrag über die Arbeitsweise der Europäischen Union |
| AIC | Akaike Information Criterion |
| AKS | Arbeitskreis „Steuerschätzungen“ |
| AO | additive outlier |
| AR(p)-Modell | autoregressives Modell der Ordnung p |
| ARIMA-Modell | Autoregressive Integrated Moving Average Modell |
| BHO | Bundeshaushaltsordnung |
| BIC | Bayesian Information Criterion |
| BIP | Bruttoinlandsprodukt |
| BLG | Bruttolöhne und -gehälter |
| BLUP | beste lineare unverzerrte Prognose |
| BMF | Bundesministerium der Finanzen |
| BMWi | Bundesministerium für Wirtschaft und Technologie |
| BRD | Bundesrepublik Deutschland |
| DIW | Deutsches Institut für Wirtschaftsforschung in Berlin |
| EU | Europäische Union |
| FAZ | Frankfurter Allgemeine Zeitung |
| GG | Grundgesetz der Bundesrepublik Deutschland |
| GLS | verallgemeinerte Kleinstquadrat Regression (Generalized Least Squares) |

| | |
|----------------------------------|---|
| HEGY-Test | Test auf saisonale Einheitswurzeln nach Hyllenberg, Engle, Granger und Yoo |
| HGrG | Haushaltsgrundsätzegesetz |
| ifo-Institut | Leibniz-Institut für Wirtschaftsforschung an der Universität München |
| IfW | Institut für Weltwirtschaft in Kiel |
| IO | innovational outlier |
| IWF | Internationaler Währungsfonds |
| IWH | Institut für Wirtschaftsforschung Halle |
| KPSS-Test | Test auf Stationarität nach Kwiatkowski, Phillips, Schmidt und Shin |
| MA(q)-Modell | Moving Average Modell der Ordnung q |
| MAPE | mittlerer absoluter Prozentfehler (Mean Absolute Percentage Error) |
| ML | Maximum Likelihood |
| MSE | mittlerer quadratischer Fehler (Mean Squared Error) |
| MSPE | mittlerer quadratischer Prozentfehler (Mean Squared Percentage Error) |
| OECD | Organisation für wirtschaftliche Zusammenarbeit und Entwicklung |
| PACF | partielle Autokorrelationsfunktion |
| PE | Prozentfehler (Percentage Error) |
| RMSPE | Wurzel aus dem mittleren quadratischen Prozentfehler (Root Mean Squared Percentage Error) |
| RWI | Rheinisch-Westfälisches Institut für Wirtschaftsforschung in Essen |
| SARIMA-Modell | saisonales Autoregressive Integrated Moving Average Modell |
| StWG | Stabilitäts- und Wachstumsgesetz |
| VGR | Volkswirtschaftliche Gesamtrechnung |

Verzeichnis der wichtigsten Symbole

| | |
|----------------------|--|
| A_t | tatsächlicher Wert des Steueraufkommens zum Zeitpunkt t |
| A^* | T -dimensionaler Spaltenvektor der tatsächlichen Werte des Steueraufkommens |
| A, B, \dots, H | Komponentenmatrizen |
| c, \hat{c} | Konstante, geschätzt |
| $\text{Corr}(\cdot)$ | Korrelation |
| $\text{Cov}(\cdot)$ | Kovarianz |
| d | Ordnung der Differenzenbildung |
| D | Ordnung der saisonalen Differenzenbildung |
| e_{th} | Prognosefehler der Prognose des Steueraufkommens für das Prognoseziel t mit dem Prognosehorizont h |
| e | TH -dimensionaler Spaltenvektor der Prognosefehler des Steueraufkommens |
| $E(\cdot)$ | Erwartungswert |
| $\hat{E}(\cdot)$ | lineare Projektion mit Konstante |
| F_{th} | Prognosewert des Steueraufkommens für das Prognoseziel t mit dem Prognosehorizont h |
| F | TH -dimensionaler Spaltenvektor der Prognosewerte des Steueraufkommens |
| $F(\cdot)$ | Verteilungsfunktion |

| | |
|---|---|
| h | Prognosehorizont |
| H | maximaler Prognosehorizont |
| i_j | j -dimensionaler Spaltenvektor aus Einsen |
| I_j | j -dimensionale Einheitsmatrix |
| $I(d)$ | integriert der Ordnung d |
| j | Zeitspanne |
| k | Quantil der Standardnormalverteilung |
| L | Wert der Likelihood am Maximum |
| L | Lag-Operator |
| n | Anzahl der zu schätzenden Parameter |
| $\stackrel{(a)}{\sim} N(\cdot; \cdot)$ | (asymptotisch) normalverteilt |
| p | Ordnung des AR-Prozesses |
| P | Ordnung des saisonalen AR-Prozesses |
| $\hat{P}(\cdot)$ | lineare Projektion |
| q | Ordnung des MA-Prozesses |
| Q | Ordnung des saisonalen MA-Prozesses |
| Q_1, \dots, Q_4 | erstes, \dots , viertes Quartal |
| $r_{t;h-1;h}$ | Änderung zwischen der Prognose für das Prognoseziel t mit dem Prognosehorizont $h - 1$ und derjenigen für dasselbe Prognoseziel mit dem darauf folgenden Prognosehorizont h |
| r | Vektor der Prognoseänderungen |
| r_{-1} | Vektor der verzögerten Prognoseänderungen |
| $\{r_{t-1;h-1;h}, r_{t-2;h-1;h}, \dots\}$ | Informationsmenge der Änderungen gleichbleibender aufeinanderfolgender Prognosehorizonte für verschiedene vergangene Prognoseziele |
| $\{r_{t;h-1+i;h+i}\}$ | Informationsmenge der vergangenen Änderungen der Prognosen desselben Prognoseziels |
| s | saisonale Periodizität |
| $se(\cdot)$ | Standardfehler |

| | |
|----------------------------|--|
| t | Zeitpunkt bzw. Prognoseziel |
| T | Anzahl an Beobachtungen bzw. Anzahl der Prognoseziele |
| u_{th} | makro-ökonomischer Schock |
| δ_d | Vektor von Dummyvariablen für die Prognosehorizonte |
| u_{t+s} | Prognosefehler der besten linearen unverzerrten Prognose |
| U' | modifiziertes Theils U Maß für alle Prognoseziel t und Prognosehorizonte h |
| U'_h | modifiziertes Theils U Maß für den Prognosehorizont h |
| U'_t | modifiziertes Theils U Maß für das Prognoseziel t |
| v_{th} | zusammengesetzter Störterm beim Test auf Unverzerrtheit |
| v, \hat{v} | TH -dimensionaler Spaltenvektor des zusammengesetzten Störterms, geschätzt |
| $\text{Var}(\cdot)$ | Varianz |
| x_t | alle Beobachtungen bis zum Zeitpunkt t |
| X | Regressormatrix |
| \bar{y} | arithmetisches Mittel |
| y_t | Beobachtungswert zum Zeitpunkt t |
| $\{y_t\}$ | Folge der Beobachtungswerte |
| (y_1, y_2, \dots, y_T) | Stichprobenrealisation |
| Y_t, \hat{Y}_t | Zufallsvariable Y zum Zeitpunkt t , geschätzt |
| $\{Y_t\}$ | Folge der Zufallsvariable Y über die Zeit; stochastischer Prozess |
| \tilde{Y}_t | durch Differenzenbildung gefilterte Zeitreihe |
| $\alpha, \hat{\alpha}$ | Verzerrung des Steueraufkommens, geschätzt |
| $\alpha_h, \hat{\alpha}_h$ | Verzerrung für den Prognosehorizont h , geschätzt |
| $\alpha^H, \hat{\alpha}^H$ | H -dimensionaler Spaltenvektor aller horizont-spezifischen Verzerrungen, geschätzt |
| $\beta_k, \hat{\beta}_k$ | partieller Autokorrelationskoeffizient, geschätzt |
| $\gamma, \hat{\gamma}$ | Koeffizient zum Test auf Informationseffizienz, geschätzt |
| $\gamma_0, \hat{\gamma}_0$ | Varianz, geschätzt |

| | |
|--|---|
| $\gamma_j, \hat{\gamma}_j$ | Kovarianz, geschätzt |
| $\delta, \hat{\delta}$ | Differenz der horizont-spezifischen Verzerrungen, geschätzt |
| $\delta^h, \hat{\delta}^h$ | Parametervektor für die Prognosehorizonte, geschätzt |
| Δ | Differenz |
| $\varepsilon_t, \hat{\varepsilon}_t$ | Störterm zum Zeitpunkt t, geschätzt (Residuum) |
| ε_{th} | idiosynkratischer Störterm beim Test auf Unverzerrtheit |
| $\theta_q, \hat{\theta}_q$ | Koeffizient des MA-Prozesses, geschätzt |
| $\theta_{sQ}, \hat{\theta}_{sQ}$ | Koeffizient des saisonalen MA-Prozesses, geschätzt |
| $\Theta_q(L)$ | MA Lag-Polynom |
| $\Theta_q(L^s)$ | saisonales MA Lag-Polynom |
| Θ | Hilfsmatrix für die Korrelationsstruktur der Prognosefehler |
| $\kappa, \hat{\kappa}$ | Parametervektor, geschätzt |
| λ_{th} | aggregierte makro-ökonomische Schocks zwischen Prognosezeitpunkt und Prognoseziel |
| $\mu, \hat{\mu}$ | Erwartungswert, geschätzt |
| $\hat{\rho}_0$ | Schätzer für σ_ε^2 |
| $\hat{\rho}_1, \hat{\rho}_1^*$ | Schätzer für $\sigma_u^2, \sigma_{u^*}^2$ |
| $\sigma^2, \hat{\sigma}^2$ | Varianz, geschätzt |
| $\sigma_\varepsilon^2, \hat{\sigma}_\varepsilon^2$ | Varianz des idiosynkratischen Störterms, geschätzt |
| $\sigma_u^2, \hat{\sigma}_u^2$ | Varianz des makro-ökonomischen Schocks, geschätzt |
| $\Sigma, \hat{\Sigma}$ | Varianz-Kovarianz-Matrix des zusammengesetzten Störterms v , geschätzt |
| $\phi_p, \hat{\phi}_p$ | Koeffizient des AR-Prozesses, geschätzt |
| $\phi_{sP}, \hat{\phi}_{sP}$ | Koeffizient des saisonalen AR-Prozesses, geschätzt |
| $\Phi_p(L)$ | AR Lag-Polynom |
| $\Phi_P(L^s)$ | saisonales AR Lag-Polynom |
| $\varphi_j, \hat{\varphi}_j$ | Autokorrelation, geschätzt |
| $\overset{(a)}{\sim} \chi^2(\cdot)$ | (asymptotisch) Chi-Quadrat-verteilt |

| | |
|----------------------------|---|
| Ψ | Matrix, die die Korrelationsstruktur der Prognosen bzw. Prognosefehler beschreibt |
| Ψ_i | Koeffizienten des $MA(\infty)$ -Prozesses |
| $\Psi(L)$ | $MA(\infty)$ Lag-Polynom |
| ω_t | Störterm beim Test auf Informationseffizienz |
| ω | Störtermvektor beim Test auf Informationseffizienz |
| $\Upsilon, \hat{\Upsilon}$ | Varianz-Kovarianz-Matrix des Störtermvektors ω , geschätzt |

1. Einführung

1.1. Problemstellung

Die Vorausschätzung der Steuereinnahmen stellt die Grundlage für die Haushaltsaufstellung sowie die mittelfristige Finanzplanung von Bund, Ländern und Kommunen dar (vgl. Bundesministerium der Finanzen (2005), S. 5). In Deutschland werden seit 1955 die Prognosen des Steueraufkommens vom Arbeitskreis „Steuerschätzungen“ (AKS) erstellt. Da die Steuereinnahmen gegenwärtig mehr als drei Viertel der gesamten Einnahmen der öffentlichen Haushalte (ohne die Sozialversicherung) ausmachen, kommt ihrer Schätzung eine hohe **finanzpolitische Bedeutung** zu. So betragen beispielsweise im Jahr 2010 die Steuereinnahmen des Bundes 226,2 Milliarden Euro, was bei 259,3 Milliarden Euro Einnahmen insgesamt einen Anteil von 87 Prozent ausmacht (vgl. Bundesministerium der Finanzen (2011b), S. 2).

Außerdem stehen die offiziellen Prognosen des AKS vor allem dann im Blickpunkt des öffentlichen Interesses, wenn das Steueraufkommen deutlich überschätzt wird, so dass die Steuerausfälle zu Haushaltlöchern führen, die wiederum Konsolidierungsmaßnahmen beziehungsweise eine höhere Kreditaufnahme nach sich ziehen (vgl. beispielsweise auch Körner (1995), S. 14). Unterschätzungen sorgen demgegenüber für eine bessere Haushaltslage als von der Regierung erwartet und liefern den Politikern Spielraum für Steuerentlastungen, die Finanzierung anderer Wahlversprechen oder die Rückführung der Staatsverschuldung. Im Mai 2010 war beispielsweise das öffentliche Interesse an den Vorhersagen des AKS besonders groß, „weil die Bun-

desregierung hiervon ihr weiteres Vorgehen in der Steuerpolitik abhängig gemacht hatte“ (Lehmann (2010), S. 1).

Doch neben der finanzpolitischen Relevanz ergibt sich die Notwendigkeit der Steuerprognosen auch aus **verfassungs- und haushaltsrechtlichen Vorschriften**. So beschreiben der Artikel 110 des Grundgesetzes (GG) sowie Paragraph 11 der Bundeshaushaltsordnung (BHO) die gesetzlichen Verpflichtungen des Bundes in Bezug auf die kurzfristige Vorausschätzung der Steuereinnahmen, die zur Aufstellung des Bundeshaushalts benötigt wird. Die Länder haben gemäß den Paragraphen 1 und 8 des Haushaltsgrundsätzegesetzes (HGrG) die Verpflichtung, diese Vorschriften ebenfalls in den jeweiligen Länderhaushaltsordnungen umzusetzen. Des Weiteren ziehen die allgemeinen Haushaltsgrundsätze die Notwendigkeit einer möglichst exakten Steuerschätzung nach sich. Dabei stehen vor allem die Grundsätze der Vollständigkeit, Vorherigkeit, Genauigkeit und Jährlichkeit eines Budgets im Mittelpunkt, wonach unter anderem alle erwarteten Einnahmen und voraussichtlichen Ausgaben **vor** Verabschiedung eines Etats **möglichst exakt** zu veranschlagen sind (vgl. beispielsweise Diller (1987), S. 293).

Entsprechendes gilt auch für die mittelfristige Finanzplanung, deren gesamte Planungsperiode fünf Jahre beträgt, wobei das erste Planungsjahr der Finanzplanung das laufende Haushaltsjahr ist (vgl. §50 HGrG). Denn Deutschland hat sich als Mitgliedsland der Europäischen Union gemäß Artikel 109 GG sowie Artikel 126 des Vertrags über die Arbeitsweise der Europäischen Union (AEUV) zum Stabilitäts- und Wachstumspakt verpflichtet. Dieser sieht vor, dass alle Mitgliedsländer jährlich ein Stabilitätsprogramm bei der Europäischen Kommission vorzulegen haben, was eine jährliche mittelfristige Steuerschätzung erforderlich macht. So finden sich in Paragraph 9 Absatz 1 des Stabilitäts- und Wachstumsgesetzes (StWG) die gesetzlichen Verpflichtungen des Bundes sowie in den Paragraphen 14 und 16 Absatz 2 StWG die analogen Folgerungen für die Länder und Gemeindeverbände.

Auch wenn beispielsweise eine starke Überschätzung des Steueraufkommens gegen den Haushaltsgrundsatz der Genauigkeit verstößt, ergeben sich daraus zwar keine juristischen Folgen, dennoch muss die Regierung in diesem Fall mit auftretenden **politischen Schwierigkeiten** rechnen. Denn weder ein höheres Haushaltsdefizit als vorhergesehen noch daraus erforderliche Konsolidierungsmaßnahmen werden von der Öffentlichkeit beziehungsweise den Wählern gerne gesehen. Im Jahre 1982 führte eine Folge drastisch nach unten korrigierter Steuerschätzungen nicht zuletzt zum Zusammenbruch der sozialliberalen Koalition, was den damaligen Regierungswechsel nach sich zog (vgl. von der Lippe (1986), S. 335).

Im Gegensatz zu den Staatsausgaben hat die Regierung allerdings nur einen sehr mittelbaren Einfluss auf die Steuereinnahmen. Schließlich hängen diese nicht nur vom geltenden Steuerrecht ab, sondern werden maßgeblich von der wirtschaftlichen Entwicklung und dem Verhalten der Steuerpflichtigen beeinflusst. Gerade bei Steuerrechtsänderungen sind die Reaktionen der Steuerzahler nur sehr begrenzt vorhersehbar. Des Weiteren spielt auch der umgekehrte Zusammenhang zwischen Konjunkturverlauf und Steueraufkommen eine zentrale Rolle. Denn die wirtschaftliche Entwicklung hängt wiederum von der Steuerbelastung der Zensiten ab.

Da die Steuerschätzungen also immer eine Erwartungsgröße und damit eine **unsichere Prognose** zukünftiger Steuereinnahmen darstellen, stellt sich unmittelbar die Frage, wie „gut“ beziehungsweise treffsicher die offiziellen Schätzungen des AKS tatsächlich sind, und ob es unter Umständen alternative Schätzmethode gibt, die das Steueraufkommen mit einer mindestens vergleichbaren, wenn nicht sogar höheren, Prognosequalität vorhersagen. Dabei ist ebenfalls von Interesse, ob die AKS-Schätzungen die für eine Prognose wünschenswerten Eigenschaften der Unverzerrtheit und Effizienz erfüllen. **Ziel dieser Arbeit** ist es, diese Fragestellungen aufzugreifen, indem saisonale ARIMA-Modelle als alternative Schätzmethode vorgeschlagen werden und deren Prognosequalität mit derjenigen der offiziellen AKS-Schätzungen verglichen wird.

1.2. Aufbau der Arbeit

In der vorliegenden Arbeit wird zunächst im zweiten Kapitel auf den institutionellen sowie methodischen Rahmen der Steuerprognosen des AKS eingegangen und dessen Vorgehensweise bei der Steuerschätzung erläutert. Anschließend werden die Determinanten der Präzision einer Steuerprognose sowie verschiedene Prognosemethoden vorgestellt und es werden die damit verbundenen Schwierigkeiten bei der Steuerschätzung diskutiert. Ein Überblick über die bisherigen empirischen Analysen zur Prognosequalität der AKS-Schätzungen schließt das Grundlagenkapitel ab.

Anschließend werden im dritten Kapitel als alternative Schätzmethode die so genannten saisonalen ARIMA-Modelle vorgeschlagen und schrittweise entwickelt. Auf die Beschreibung der Datengrundlage folgt sodann die Untersuchung der Zeitreihe des Gesamtsteueraufkommens der Bundesrepublik Deutschland (BRD) von 1961 bis 2010. Mit Hilfe des Box-Jenkins-Verfahrens werden geeignete Modellspezifikationen abgeleitet, anhand derer eigene Prognosen für die Jahre 2001 bis 2010 erstellt werden.

Darauf aufbauend wird im vierten Kapitel einerseits die Qualität dieser saisonalen ARIMA-Schätzungen sowie andererseits der offiziellen AKS-Prognosen anhand ökonomischer Modelle auf ihre Prognoserationalität, das heißt auf Unverzerrtheit und Informationseffizienz, untersucht. Hierbei kommt erstmalig ein so genannter gepoolter Ansatz zur Anwendung, bei dem nicht mehr für jeden Prognosehorizont separat getestet wird, sondern die Prognosen für alle Prognosehorizonte zusammen auf ihre Prognoserationalität untersucht werden. Ein Vergleich der horizontalen, vertikalen sowie gepoolten Treffsicherheit der eigenen SARIMA-Prognosen mit derjenigen der offiziellen AKS-Prognosen anhand deskriptiver Gütemaße schließt das vierte Kapitel ab.

Das fünfte Kapitel fasst die Ergebnisse zusammen, stellt die daraus abgeleiteten Politikempfehlungen dar und bietet einen Ausblick auf noch offene Forschungsfragen und Erweiterungsmöglichkeiten der vorliegenden Analyse.

2. Grundlagen

2.1. Der Arbeitskreis „Steuerschätzungen“

2.1.1. Der institutionelle Rahmen des AKS

Der Arbeitskreis „Steuerschätzungen“ wurde im Jahre 1955 als Beirat beim Bundesministerium der Finanzen (BMF) gegründet und setzt sich aus insgesamt etwa 35 Vertretern folgender Entsendungsstellen zusammen (vgl. Fox (2005), S. 245):

- aus dem BMF (federführend)
- aus dem Bundesministerium für Wirtschaft und Technologie (BMWV)
- von den fünf großen Wirtschaftsforschungsinstituten: Leibniz-Institut für Wirtschaftsforschung an der Universität München (ifo-Institut), dem Deutschen Institut für Wirtschaftsforschung (DIW) in Berlin, dem Rheinisch-Westfälischen Institut für Wirtschaftsforschung (RWI) in Essen, dem Institut für Weltwirtschaft (IfW) in Kiel sowie dem Institut für Wirtschaftsforschung Halle (IWH) in Halle
- aus dem Statistischen Bundesamt
- von der Deutschen Bundesbank
- aus dem Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung
- aus den Länderfinanzministerien
- aus der Bundesvereinigung kommunaler Spitzenverbände

Der Vorsitz obliegt dem zuständigen Referatsleiter im BMF. Die meisten Mitglieder des AKS blicken auf eine langjährige Erfahrung bei der Steuerschätzung zurück. Normalerweise finden zwei Sitzungen pro Jahr statt, nämlich im Mai und im November. Dabei werden im Mai die Steuerschätzungen für den mittelfristigen Zeitraum erstellt, das heißt für das laufende Jahr plus vier Folgejahre, welche die Grundlage für den Haushaltsentwurf des Folgejahres, für die jährliche Finanzplanung und für die Abstimmung der Finanzplanung zwischen Bund, Länder und Gemeinden im Finanzplanungsrat darstellen (vgl. Bundesministerium der Finanzen (2005), S. 11). Kurz vor Abschluss der Haushaltsberatungen im November erfolgen eine Überprüfung der Schätzungen für das laufende Jahr sowie eine Überarbeitung der Prognose für das kommende Jahr. Diese Prognosen bilden dann die Basis für die Ansätze des Haushaltsgesetzes des Bundes.

Im November 1969 wurde zudem der Unterausschuss „**Regionalisierung**“ gegründet, dessen Aufgabe es ist, die für die Gesamtheit der Länder geschätzten Steuereinnahmen auf die einzelnen Bundesländer aufzuteilen und auch den Länderfinanzausgleich zu berechnen. Die Federführung dieses Ausschusses liegt bis heute beim Land Baden-Württemberg (vgl. Bundesministerium der Finanzen (2005), S. 8 f.). Im November 1972 führte der AKS ergänzend zu den Steuerschätzungen so genannte **Methodensitzungen** ein, in denen Verbesserungen der Schätzmethodik erörtert werden. Die Sitzungen erfolgen in unregelmäßigen Abständen, oft im Vorfeld weitgehender Steuerrechtsänderungen, die eine Anpassung des Schätzinstrumentariums erforderlich machen.

2.1.2. Vorgehensweise des AKS

In vielen anderen Staaten wird die Steuerschätzung ausschließlich von den jeweiligen Finanzministerien durchgeführt, was des Öfteren den Verdacht der Öffentlichkeit nahelegt, sie verfolgten mit der Schätzung der Steuereinnahmen finanzpolitische Zielsetzungen. Nach einer Auseinandersetzung im Jahr 1954 zwischen dem BMF und dem ifo-Institut um das Ausmaß der als möglich gesehenen Einkommen-

steuersenkungen ist dagegen in Deutschland der AKS entstanden, um zukünftig „die Prognosen des Steueraufkommens durch die Mitwirkung unabhängiger Experten zu objektivieren und von tagespolitischen Ziel- und Wunschvorstellungen freizuhalten“ (Gebhardt (2001), S. 128). Laut der Jubiläumsausgabe „50 Jahre Arbeitskreis Steuerschätzungen“ des BMF aus dem Jahr 2005 ist die **Zielsetzung** des AKS,

„(...) möglichst den gesamten in Deutschland vorhandenen Sachverstand auf dem Gebiet der Steuerschätzung in einem Gremium zu versammeln. Die Mitglieder des AKS sollen allein aufgrund von fachlichen Erwägungen zu einem Schätzergebnis gelangen. Dieses Ergebnis mag zwar im Einzelfall mehr oder weniger vom späteren Ist-Ergebnis abweichen, doch ist diese Abweichung nicht auf politische Einflussnahme auf die Steuerschätzung zurückzuführen.“ (Bundesministerium der Finanzen (2005), S. 12)

Auch Körner (1983) erwähnt lobend, dass Strömungen politischer Willensbildung dank der unabhängigen Mitglieder im AKS nicht zum Tragen kommen, fügt allerdings den Nebensatz hinzu, „wenn auch der Bund bedingt durch Vorsitz und Einigungstechnik manchmal einen Hauch von Schätzführerschaft zu wahren weiß“ (Körner (1983), S. 250).

Allerdings stützt sich der AKS nicht allein auf einen einzelnen nachvollziehbaren methodischen Ansatz, sondern acht der Mitglieder, nämlich die fünf Wirtschaftsforschungsinstitute, die Bundesbank, der Sachverständigenrat und das BMF erarbeiten jeweils unabhängig voneinander mit eigenen Methoden und Modellen Schätzvorschläge für jede Einzelsteuer. Diese dürfen der Öffentlichkeit nicht bekannt gegeben werden, sondern werden dann bei der Zusammenkunft des AKS diskutiert und gemäß der Regeln eines bargaining-Prozesses angepasst, bis ein Konsens erreicht wird, der von allen Mitgliedern vertreten werden kann (vgl. Bundesministerium der Finanzen (2005), S. 11). Erst das Endergebnis dieser **Konsensgespräche** wird schließlich in einer Pressemitteilung des BMF der Öffentlichkeit zugänglich gemacht. Fachliche Darstellungen der Arbeitsweise des AKS gibt es kaum. Außerdem sind kritische

Analysen der Schätzergebnisse in der Literatur genauso dünn gesät wie unabhängige Arbeiten zu Methodenfragen der Steuerprognose. Wünschenswert wäre daher einerseits ein ausgeprägteres wissenschaftliches Interesse an Vorgehensweisen und Problemen der Steuerschätzung sowie andererseits mehr Transparenz durch eine stärkere Offenlegung des methodischen Vorgehens des Arbeitskreises „Steuerschätzungen“ (vgl. z. B. auch Fox (2005), S. 246).

Offiziell bekannt ist lediglich, dass die Schätzungen des AKS an die von der Bundesregierung vorgegebenen Grundannahmen der gesamtwirtschaftlichen Entwicklung gebunden sind, die Steuereinnahmen auf Grundlage des geltenden Steuerrechts geschätzt werden und die finanziellen Auswirkungen von geplanten Steuerrechtsänderungen gesondert in der Haushalts- und Finanzplanung berücksichtigt werden (s. Bundesministerium der Finanzen (2005), S. 10). Die so genannten **Eckdaten für die kurzfristige Wirtschaftsentwicklung** sind Zielprojektionen der Bundesregierung, die unter Federführung des BMWi und unter Beteiligung des BMF, des Statistischen Bundesamts und der Deutschen Bundesbank im interministeriellen Arbeitskreis „Gesamtwirtschaftliche Vorausschätzungen“ halbjährlich erstellt werden. Die entsprechenden Eckdaten für die mittelfristige Entwicklung werden allein vom BMWi unter Mitarbeit des BMF erarbeitet (vgl. Rürup und Schachler (1983), S. 217).

Im Gegensatz zu Prognosen, die die unter bestimmten Annahmen wahrscheinlichste Entwicklung schätzen, geben diese (kurz- und mittelfristigen) **Zielprojektionen** vielmehr die von der Bundesregierung „angestrebte, für wünschenswert und erreichbar gehaltene Entwicklung“ wieder und sind damit eine „Mischung aus Prognose und Plan“ (von der Lippe (1986), S. 336). Sie enthalten somit Aussagen über die wirtschaftspolitischen Ziele der Regierung und sollen dementsprechend auch eine Signalwirkung haben. Diese dem AKS vorgegebenen Eckdaten weisen die voraussichtliche Entwicklung des nominalen und realen Bruttoinlandsprodukts (BIP), die Aufteilung des nominalen BIP auf die Verwendungsaggregate (Konsumausgaben der privaten Haushalte und des Staates, Bruttoanlageinvestitionen, Vorratsveränderung,

Außenbeitrag), die prognostizierte Entwicklung der Unternehmens- und Vermögenseinkommen, die Bruttolöhne und -gehälter (BLG) sowie die BLG je beschäftigten Arbeitnehmer aus und werden vom AKS als Indikatoren für die Steuerbasis verschiedener Einzelsteuerschätzungen herangezogen (vgl. z. B. Fox (2005), S. 246).

Die Bindung des AKS an diese gesamtwirtschaftlichen Eckwerte hat zwar den „Vorteil einer Vereinheitlichung der für die Schätzungen erforderlichen Ausgangsdaten“ (Diller (1987), S. 294), bietet aber vor allem Anlass zur **Kritik**. Denn die mit der Gründung des AKS eigentlich angestrebte Objektivierung der Steuerschätzung und Unabhängigkeit von tagespolitischen Ziel- und Wunschvorstellungen (s.o.) kann durch die strikte Vorgabe der gesamtwirtschaftlichen Eckwerte als Ausgangsdaten der bedingten Prognosen des AKS stark hinterfragt werden (vgl. auch Schaft (1981), S. 488). Zwar liegen laut BMF die Projektionen der Regierung „innerhalb des Schätzspektrums anderer Institutionen“ (Bundesministerium der Finanzen (2005), S. 17) – darunter der Sachverständigenrat, die Bundesbank sowie mehrere Forschungsinstitute und auf internationaler Ebene die Europäische Union (EU), die Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) und der Internationale Währungsfond (IWF). Außerdem werden im AKS zu Beginn einer jeden Sitzung die vom Bund vorgelegten Annahmen ausführlich diskutiert und teilweise auch intensiv und lange hinterfragt (vgl. Fox (2005), S. 246).

Dennoch bleibt das **Risiko**, dass allein schon durch den Zielcharakter der gesamtwirtschaftlichen Projektion mehr von der angestrebten als der zu erwartenden Entwicklung bei den Prognosen ausgegangen wird, was zu einer Tendenz der Überschätzung des Steueraufkommens führen könnte (vgl. Gebhardt (2001), S. 129). So titulierte die Frankfurter Allgemeine Zeitung (FAZ) beispielsweise am 10. Mai 2005 „Wachstumsprognose regelmäßig zu optimistisch“ und behauptete damals „Die Bundesregierung hat einen zu hohen Ansatz gewählt, weil sie Optimismus verbreiten wollte oder vor Wahlen die Wirtschaftslage besser zeichnen wollte“ (Schäfers (2005)). Durch die Vorgabe der gesamtwirtschaftlichen Annahmen für die Steuerschätzung

des AKS übernimmt die Bundesregierung jedenfalls auch implizit die Hauptverantwortung für das Prognoseergebnis und eventueller Fehleinschätzungen.

2.2. Methoden der Steuerschätzung und damit verbundene Schwierigkeiten

Zunächst folgt eine Übersicht über die Bestimmungsfaktoren des Steueraufkommens. Anschließend werden die verschiedenen Methoden der Steueraufkommensprognose vorgestellt, und es wird auf die damit jeweils verbundenen Schwierigkeiten der Steuerschätzung eingegangen.

2.2.1. Die Determinanten der Präzision der Steuerprognose

Das Steueraufkommen hängt von einer Vielzahl von Bestimmungsfaktoren ab. Leibrecht (2003) systematisiert die verschiedenen Einflussfaktoren in „Ist- und Soll-Determinanten der Präzision einer kurzfristigen Steuerprognose“ und konstruiert das in Abbildung 2.1 dargestellte Modell (vgl. Leibrecht (2003), S. 135).

Dabei handelt es sich um eine Erweiterung des Modells von Bretschneider und Gorr (1987). Gemäß Leibrecht gehören das bestehende Steuerrecht, beschlossene und geplante Steuerrechtsänderungen, das Verhalten der Zensiten sowie die konjunkturellen Rahmenbedingungen zu den so genannten Ist-Determinanten, die sich wiederum gegenseitig stark beeinflussen, und sowohl einzeln als auch über interdependente Kanäle eine Vielzahl von Unsicherheiten bezüglich des Ist-Wertes des Steueraufkommens hervorrufen. Die so genannten Soll-Determinanten beeinflussen dagegen die Festlegung des Prognosewertes (Soll-Wert) und sind im Modell von Leibrecht die eingesetzten Prognosemethoden, der Ablauf der Prognose, die organisatorischen und institutionellen Faktoren sowie die politischen Faktoren. Dabei hängt die Wahl der Prognosemethode laut Leibrecht vom Steuerrecht ab, der Ablauf der Progno-

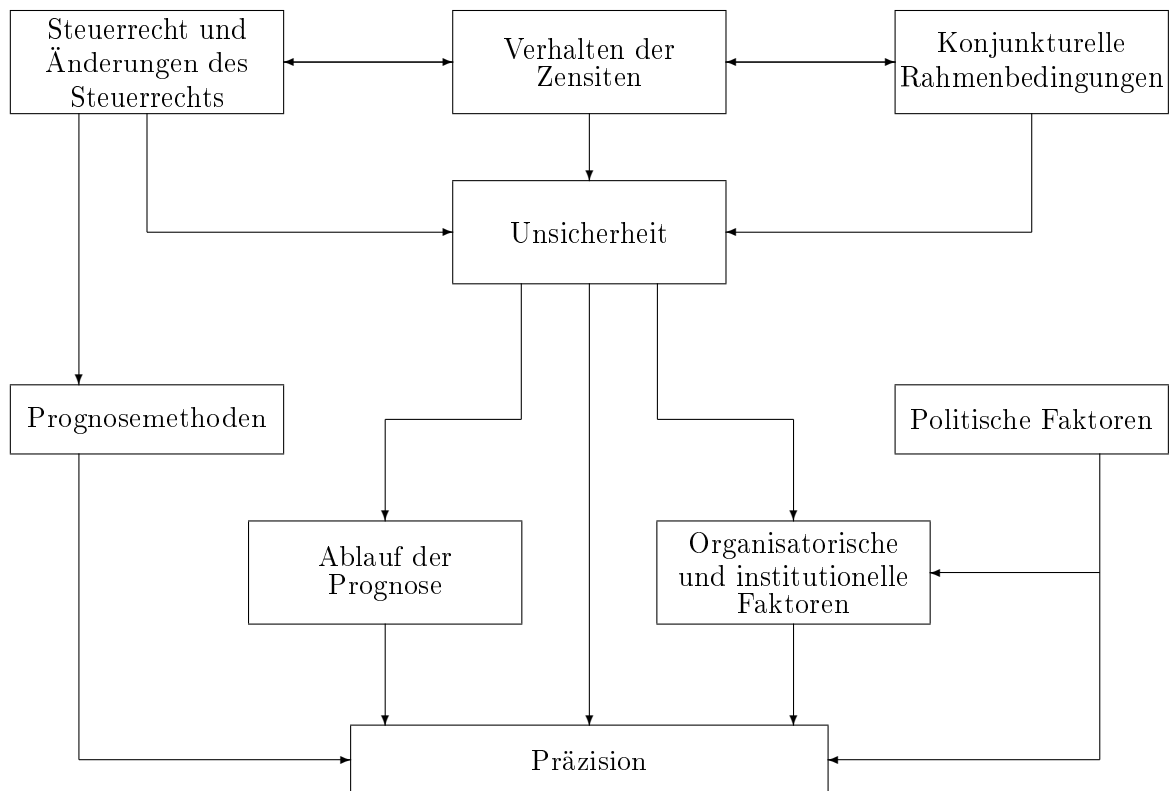


Abbildung 2.1.: Leibrechts Modell der Determinanten der Präzision (Leibrecht (2003), S. 135)

se sowie die organisatorischen und institutionellen Faktoren werden von den Ist-Determinanten mitbestimmt, und die politischen Faktoren beeinflussen ebenfalls die organisatorischen und institutionellen Faktoren.

Die Einflüsse der **Ist-Determinanten** lassen sich aufgrund der vielen Wechselwirkungen nur schwer systematisieren. Vor allem die Auswirkungen von Steuerrechtsänderungen auf das Verhalten der Zensiten und damit auf das Steueraufkommen erschweren die Prognose maßgeblich. Doch auch bei bestehendem Steuerrecht entsteht Unsicherheit bezüglich des Steueraufkommens, einerseits aufgrund von Zeitverzögerungen zwischen dem Entstehen der Steuerschuld und ihrer Kassenwirksamkeit, andererseits eventuell aufgrund des Steuertarifs (vgl. Leibrecht (2004), S. 43 f.).

Die **Zeitverzögerungen** lassen sich wiederum in Veranlagungsverzögerungen und Zahlungslag unterteilen. Dabei bezeichnet die Veranlagungsverzögerung den Zeitraum von der Entstehung der Steuerschuld bis zum Abschluss der Veranlagung durch die Finanzverwaltung. Sie besteht aus der Steuererklärungsverzögerung, die von den geltenden Fristen zur Abgabe der Steuererklärung abhängt, und der jeweiligen Bearbeitungsverzögerung durch das Finanzamt (vgl. Leibrecht (2003), S. 137). Unter dem Zahlungslag versteht man die Verzögerungen im effektiven Kassenzufluss einer Steuerschuld, die durch die Bestimmungen zu Voraus-, Abschluss-, Nach- und sonstigen Zahlungen bestimmt werden. So führt auch eine lange Bearbeitungszeit der Finanzbehörde neben der Veranlagungsverzögerung zu einem größeren Zahlungslag. Die zeitlichen Verzögerungen hängen schließlich stark vom nicht zwingend fristgerechten Verhalten der Zensiten ab, vor allem bezüglich der Steuererklärungsverzögerung aber auch durch Zahlungsverzögerungen (vgl. beispielsweise Flascha (1985), S. 64 ff.).

Außerdem entsteht Unsicherheit bei der Prognose der Steuerschuld eines Jahres dadurch, dass die Steuerpflichtigen die bestehenden **Spielräume zur Gestaltung der Bemessungsgrundlage** für sich nutzen können, z. B. durch Freibeträge, Verlustvor- oder -nachträge sowie durch Rückforderungen von Steuerguthaben. Auch die **Art des Steuertarifs** kann die Steuerprognose wesentlich beeinflussen, sofern kein proportionaler Steuersatz gilt. Liegt nämlich wie in Deutschland ein progressiver Steuertarif vor, führen reine Strukturänderungen in der Bemessungsgrundlage, die ebenfalls nur schwer prognostizierbar sind, zu Änderungen des Steueraufkommens (vgl. Leibrecht (2003), S. 138).

Die Auswirkungen von **Steuerrechtsänderungen** werden in direkte und indirekte Effekte beziehungsweise statische und dynamische Effekte untergliedert. **Statische Effekte** lassen die Wirkungen von Verhaltensänderungen der Zensiten nach Steuerrechtsänderungen außer Acht und umfassen lediglich die Veränderung in der Bemessungsgrundlage einer Steuer, die sich aufgrund der veränderten steuerrechtlichen Bestimmungen zur Bemessungsgrundlage ergibt. Ändert sich beispielsweise nur der

Steuersatz, so wird für die Schätzung der statischen Effekte die Bemessungsgrundlage der Steuer als konstant angenommen (vgl. Gravelle (1995), S. 463). Im Gegensatz dazu berücksichtigen die **dynamischen Effekte** die Auswirkungen der Verhaltensänderungen auf das Steueraufkommen, die durch die Steuerrechtsänderungen ausgelöst werden. Dabei trennt Gravelle die dynamischen Effekte weiterhin in mikroökonomische sowie makroökonomische Effekte und untergliedert letztere in nachfrage- und angebotsseitige Effekte. Mikroökonomische dynamische Effekte wirken sich über Veränderungen in der Struktur der Bemessungsgrundlage auf das Steueraufkommen aus, wie z. B. durch Steuerrechtsänderungen hervorgerufene Modifikationen in der Konsum- oder Einkommensstruktur der Steuerpflichtigen. Dagegen verändern makroökonomische dynamische Effekte von Steuerrechtsänderungen das Niveau makroökonomischer Größen, die in wechselseitiger Beziehung zum Steueraufkommen stehen. Sie werden auch als so genannte makroökonomische **Feedback-Effekte** bezeichnet (vgl. z. B. Auerbach (1996), S. 141).

Nicht zuletzt sorgen die **gesamtwirtschaftlichen Rahmenbedingungen** für große Unsicherheit bezüglich des Ist-Wertes des Steueraufkommens, da die Bemessungsgrundlage vieler Steuerarten direkt von der konjunkturellen Entwicklung der Volkswirtschaft abhängt. Dabei sind wiederum sowohl das Verhalten der Zensiten als auch wirtschaftspolitische Maßnahmen unmittelbare Einflussfaktoren. Umgekehrt beeinflusst die konjunkturelle Lage das Verhalten der Steuerpflichtigen und die Politik. Des Weiteren sorgen in offenen Volkswirtschaften exogene Schocks im Ausland für zusätzliche Risiken in Bezug auf das Steueraufkommen.

Zusätzlich zu den Ist-Determinanten können auch die **Soll-Determinanten** die Prognosegenauigkeit maßgeblich beeinflussen. Zu den **institutionellen Faktoren**, die wiederum von den politischen Gegebenheiten abhängen, zählt einerseits die Wahl der Verlustfunktion der prognostizierenden Institution bzw. Person(en). Bei bewusster so genannter konservativer Prognose liegt beispielsweise eine asymmetrische Verlustfunktion zur Prognose vor, bei der das Steueraufkommen systematisch unterschätzt und damit die Prognose verzerrt wird (vgl. Leibrecht (2003), S. 139).

Diese Wahl der Verlustfunktion lässt sich rational begründen, da die (politischen) Kosten einer Überschätzung wesentlich höher sind als diejenigen einer Unterschätzung desselben Ausmaßes. Sie hängt daher also stark davon ab, ob die Prognose von den politisch verantwortlichen Personen (allein) durchgeführt wird oder ob politisch unabhängige Institutionen die Steuerschätzungen vornehmen. Ist es andererseits den prognostizierenden Personen möglich, das Ist-Ergebnis an den Soll-Wert heranzuführen, z. B. über eine Verringerung der Bearbeitungsverzögerung in der Finanzverwaltung, besteht darüber hinaus die Möglichkeit einer „selbsterfüllenden Prophezeiung“ (vgl. Leibrecht (2004), S. 46).

Zu den **organisatorischen Determinanten** gehören gemäß Leibrecht die Anzahl der unabhängig voneinander prognostizierenden Institutionen sowie die Dokumentation des Prognoseablaufs. Bretschneider et al. (1989) ermitteln in ihrer Studie zur Prognosegenauigkeit von Steuerschätzungen in den USA, dass die Präzision der Prognose steigt, wenn mehrere unabhängige Prognosen konkurrierender Institutionen der Exekutiv- und Legislativgewalt vorliegen. Sie kommen darüber hinaus zu dem Ergebnis, dass die Prognosegenauigkeit weiter zunimmt, wenn diese konkurrierenden Prognosen über eine formale Prozedur zu einer Gesamtprognose kombiniert werden (vgl. Bretschneider et al. (1989), S. 318). In ihrer Studie von 1987 identifizierten Bretschneider und Gorr bereits die Offenlegung des Prognoseablaufs als organisatorische Determinante, die die Prognosegenauigkeit erhöht (Bretschneider und Gorr (1987), S. 125 f.).

Sowohl die institutionellen als auch die organisatorischen Faktoren werden also direkt von den politischen Gegebenheiten beeinflusst. Wird die Steuerprognose allein von einer im Einflussbereich politischer Instanzen stehenden Institution erstellt, besteht zusätzlich die Möglichkeit einer **direkten politischen Einflussnahme** und damit verzerrten Prognose. Besonders in Wahljahren könnte eine bewusste Überschätzung des Steueraufkommens den Wählern den Eindruck vermitteln, dass genügend finanzielle Mittel für die Erfüllung der Wahlversprechen im Haushalt zur

Verfügung stehen. Deswegen wird eine möglichst hohe Unabhängigkeit der Steuerschätzung von den politischen Instanzen angestrebt.

Schließlich wirken sich der **Ablauf der Prognose** sowie die **Wahl der Prognosemethode** auf die Prognosegenauigkeit aus. Daher werden nun im folgenden Kapitel zunächst die zur Wahl stehenden Prognosemethoden gegenüber gestellt, bevor die Schwierigkeiten der Steuerschätzung in Bezug auf die vom AKS verwendeten Schätzverfahren näher betrachtet werden.

2.2.2. Überblick über die Methoden der Steueraufkommensprognose

Für die Prognose des Steueraufkommens ist eine Vielzahl von Vorgehensweisen und Verfahren möglich. Die verschiedenen Methoden lassen sich wie in Abbildung 2.2 dargestellt systematisieren.

Im ersten Schritt wird zwischen qualitativen und quantitativen Methoden unterschieden. Während letztere auf einer expliziten Datenanalyse basieren, sind **qualitative Methoden** subjektive Interpretationsmethoden, die auf dem Wissen und der Erfahrung der prognostizierenden Experten beruhen (siehe einführend z. B. Makridakis et al. (1998), Kapitel 10 und 11). Handelt es sich lediglich um eine einzelne prognostizierende Person, so spricht man von einer **subjektiven Prognose**. Wenn dagegen mehrere Experten eine gemeinsame Prognose erstellen, liegt eine so genannte **Konsensprognose** vor. Eine spezielle Form der Konsensprognose ist die Delphi-Methode, bei der gewöhnlich in einem anonymen mehrstufigen Meinungsbildungsprozess ein Konsens gefunden werden soll, der von allen Gruppenmitgliedern mitgetragen werden kann (für weitere Informationen zum Delphi-Verfahren siehe z. B. Fildes (1995), S. 5 f.).

Wenngleich qualitative Prognoseverfahren in der unternehmerischen Praxis des Öfteren zur Anwendung kommen, ist die Verwendung einer rein qualitativen Methode

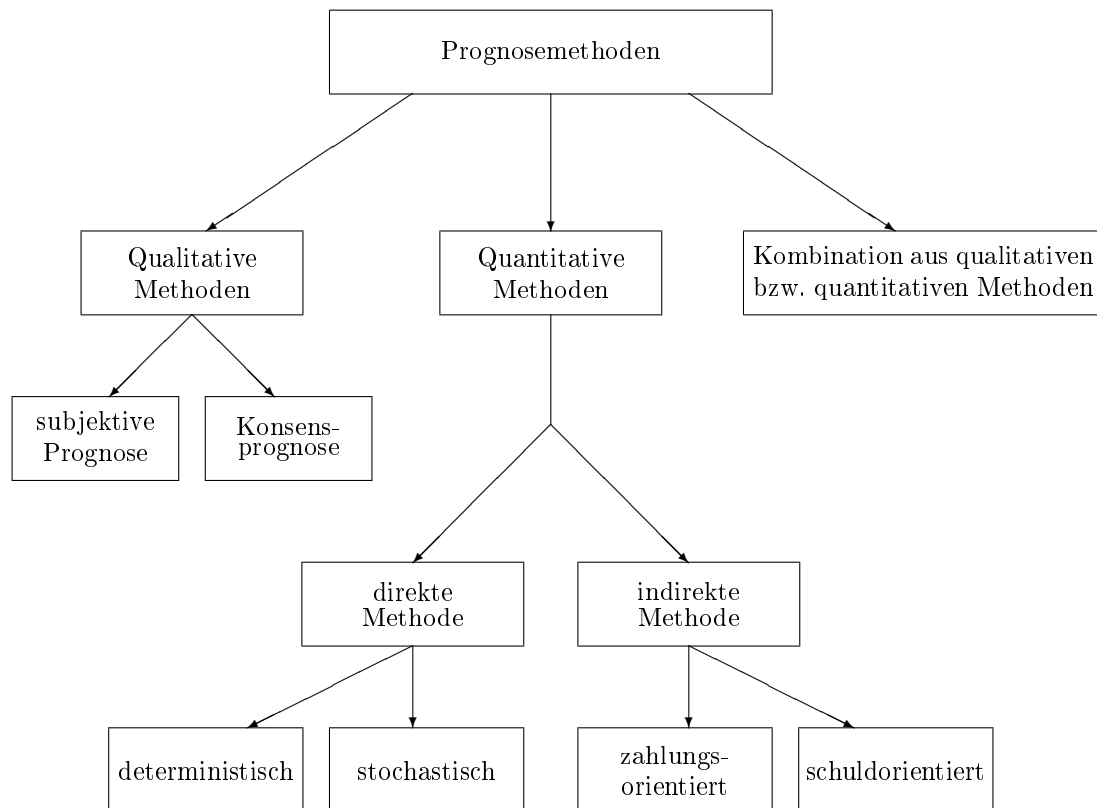


Abbildung 2.2.: Methoden der Steueraufkommensprognose (eigene Darstellung)

sehr kritisch zu betrachten, da eine Vielzahl subjektiver Verzerrungen auftreten können (eine umfassende Übersicht findet sich beispielsweise in Makridakis et al. (1998), S. 500 ff.). Darüber hinaus sind die zur Prognose getroffenen Annahmen meist nicht transparent und bei Konsensprognosen können weitere Probleme durch Gruppendynamik bzw. Gruppendruck entstehen.

Sofern die erforderlichen Daten zur Verfügung stehen, werden daher im Allgemeinen **quantitative Prognoseverfahren** bevorzugt. Diese lassen sich wiederum in direkte und indirekte Methoden untergliedern. Während direkte Prognoseverfahren ausschließlich auf der in der Vergangenheit beobachteten Entwicklung der zu prognostizierenden Variable beruhen, benötigen indirekte Verfahren geeignete erklärende Variable und deren Prognosewerte.

Die **direkten quantitativen Methoden** können schließlich in deterministische und stochastische Zeitreihenmethoden unterteilt werden, wobei nur die letzteren den der Zeitreihe unterliegenden stochastischen Prozess modellieren. Zu den **deterministischen Verfahren** zählen die mathematisch-statistischen Trendregressionen einerseits und die rein algebraischen Methoden andererseits (vgl. z. B. Pindyck und Rubinfeld (1998), Kapitel 15). Dabei beruhen Trendregressionen auf der regressionsanalytischen Anpassung einer Funktion der Zeit an die (Steuer-)Daten. Zur Modellierung von Sättigungstendenzen wird beispielsweise des Öfteren die logistische Funktion zur Trendregression und darauf basierender Prognosen herangezogen (vgl. beispielsweise Hartung (2009), S. 642 f.). Zu den wichtigsten Ansätzen der algebraischen Methoden gehören die so genannten Average-Ansätze sowie verschiedene Ansätze zur Glättung von Zeitreihen (eine ausführliche Darstellung dieser Methoden findet sich z. B. in Hartung (2009), Kapitel 1.3).

Zu den **stochastischen direkten Methoden** gehört die von George Box und Gwilym Jenkins im Jahr 1970 entwickelte Box-Jenkins-Methode (Box und Jenkins (1970)), die auf den so genannten ARIMA-Modellen basiert. Diese verwendet das Muster der seriellen Korrelation einer stationären Zeitreihe zur Modellierung des zugrunde liegenden stochastischen Prozesses. Eine ausführliche Darstellung der ARIMA-Modelle und die Anwendung des Box-Jenkins-Verfahrens bei der Steuerschätzung erfolgt im dritten Kapitel dieser Arbeit.

Im Gegensatz zu den direkten Methoden verwenden **indirekte quantitative Methoden** noch mindestens eine andere Variable zur Prognose des Steueraufkommens. So wäre als erklärende Variable für die Steuerschätzung beispielsweise die jeweilige steuerliche Bemessungsgrundlage wünschenswert, für die aber meistens die Datenbasis nicht ausreicht. Stattdessen werden dann so genannte Indikatorvariablen gesucht, die die fehlende Bemessungsgrundlage bestmöglich abbilden sollen und für welche auch Prognosewerte vorliegen, die dann als erklärende Variablen ins Modell eingehen (vgl. Leibrecht (2004), S. 63). Dies können zum Beispiel prognostizierte Werte der

Daten der Volkswirtschaftlichen Gesamtrechnung (VGR) sein, um die Abhängigkeit des Steueraufkommens von der konjunkturellen Entwicklung abzubilden.

Die indirekten Methoden lassen sich wiederum in zahlungsorientierte und steuerschuldorientierte Ansätze unterteilen. Dabei verwenden die **zahlungsorientierten Methoden** als abhängige Variable das kassenmäßige Steueraufkommen, welches unmittelbar unter Verwendung der Indikatorvariablen prognostiziert wird. Je nachdem ob Ein- oder Mehrgleichungsmodelle bzw. Ein- oder Mehrvariablenansätze gewählt werden, ergibt sich eine Vielzahl möglicher Schätzmethode, darunter auch die besonders häufig angewandte „Elastizitätenmethode“, die im folgenden Unterkapitel näher betrachtet wird.

Steuerschuldorientierte Methoden lassen sich dagegen durch ein zweistufiges Verfahren beschreiben. Im ersten Schritt wird die Steuerschuld abgeleitet, indem auf die prognostizierte Bemessungsgrundlage bzw. den Prognosewert einer geeigneten Indikatorvariable der geltende Steuertarif sowie die aktuellen Bestimmungen zu Frei- und Abzugsbeträgen angewandt werden. Im zweiten Schritt wird dann versucht, die Zahlungs- und Veranlagungsverzögerungen mit Hilfe so genannter Lag-funktionen zu berücksichtigen. Auf diese Weise wird die Steuerschuld auf verschiedene Zeiträume aufgeteilt, so dass schließlich das kassenmäßige Steueraufkommen ermittelt werden kann (vgl. Leibrecht (2004), S. 65). Zu den steuerschuldorientierten Ansätzen gehören auch die Mikrosimulationsmodelle, die sich besonders gut für die Quantifizierung der finanziellen Auswirkung geplanter Steuerrechtsänderungen eignen (vgl. beispielsweise Wagenhals (2011, 2004)).

Außerdem können verschiedene qualitative oder quantitative Methoden kombiniert werden. So ist die oben beschriebene Konsensprognose des AKS ebenfalls eine **Kombination mehrerer Prognosemethoden**, da die unabhängigen (meist quantitativen) Prognosen der verschiedenen Institutionen über einen qualitativen Diskussionsprozess, nämlich über das Konsensgespräch, zu einer Gesamtprognose zusammengeführt werden.

2.2.3. Schwierigkeiten bei der Steuerschätzung

Die bei der Steuerprognose auftretenden Probleme sind vielfältig und hängen stark von der Wahl der Prognosemethode und der zu prognostizierenden Steuerart ab.

Als wichtiges Hilfsmittel bei der Prognose des Steueraufkommens wird in der Literatur häufig der **Elastizitätenansatz** als indirekte zahlungsorientierte Methode aufgeführt (vgl. z. B. Rürup und Schachler (1983), S. 219). Allgemein gibt eine Elastizität das Verhältnis zweier Wachstumsraten zueinander wieder, genauer gesagt die relative Änderung einer Wirkungsgröße geteilt durch die relative Änderung der sie verursachenden Größe (vgl. Diller (1987), S. 294). Die Aufkommenselastizität (für das Steuersystem insgesamt bzw. einer Einzelsteuer) ist dementsprechend die Wachstumsrate des Steueraufkommens (der Steuereinnahmen insgesamt bzw. einer Einzelsteuer) dividiert durch die Wachstumsrate des nominalen Bruttoinlandsproduktes (bzw. bis 1993 des Bruttosozialproduktes). Elastizitäten sollen demnach als Maß für die Konjunktorempfindlichkeit einer Steuer dienen und das Ausmaß der Abhängigkeit zwischen dem Steueraufkommen und seinen Bestimmungsfaktoren abbilden (vgl. von der Lippe (1986), S. 336).

Dabei wird allerdings ein (einseitiger) Kausalzusammenhang zwischen der gesamtwirtschaftlichen Entwicklung als unabhängige Variable und der Entwicklung des Steueraufkommens als abhängige Variable unterstellt (vgl. Rürup und Schachler (1983), S. 219). Tatsächlich wird diese Annahme jedoch verletzt, da sich die konjunkturelle Entwicklung und das Steueraufkommen gegenseitig beeinflussen und damit interdependente Größen sind. Trotz dieses **Endogenitätsproblems** finden Steueraufkommenselastizitäten in der Literatur breite Anwendung (vgl. unter anderem Körner (1983), von der Lippe (1986), Schoof (1998), Gebhardt (2001)). Eine weitere Kenngröße zur Beschreibung der Steuereinnahmen ist die so genannte Steuerquote, die die Relation der Summe der kassenmäßigen Steuereinnahmen zum nominalen Bruttoinlandsprodukt bezeichnet (vgl. Schoof (1998), S. 210). Eine Elastizität größer als eins stellt also eine zunehmende Steuerquote dar, während eine Elastizität klei-

ner als eine abnehmende Steuerquote widerspiegelt (vgl. von der Lippe (1986), S. 336).

Bei der Schätzung der Elastizitäten besteht außerdem eine große Schwierigkeit dadurch, dass weder die Elastizität des gesamten Steueraufkommens noch die Aufkommenselastizitäten einzelner Steuerarten im Zeitablauf konstant sind. Die starken Schwankungen werden meist durch Durchschnittswerte für bestimmte Perioden gemittelt. **Verschiedene Stützzeiträume** beinhalten dabei allerdings unterschiedliche spezifische Gegebenheiten, wie das in der jeweiligen Zeitspanne geltende Steuerrecht, die Einkommensverteilung der Steuerpflichtigen, das Verfahren der Steuererhebung und die Struktur des Steueraufkommens, die sich im Zeitablauf ändern können. So lässt sich beispielsweise in den achtziger Jahren eine Zunahme des Gewichts der indirekten Steuern gegenüber den direkten in der Bundesrepublik Deutschland beobachten (vgl. von der Lippe (1986), S. 336).

Der Arbeitskreis „Steuerschätzungen“ verwendet vorwiegend eine **indirekte steuerschuldorientierte Methode** zur Schätzung der kassenmäßigen Einnahmen der einzelnen Steuern (vgl. von der Lippe (1986), S. 337). Ausgehend von den vorgegebenen Zielprojektionen zur gesamtwirtschaftlichen Entwicklung wird dabei im ersten Schritt die jeweilige Bemessungsgrundlage der einzelnen Steuerarten prognostiziert. Durch Anwendung des entsprechenden Steuertarifs erhält man im zweiten Schritt die geschätzte Steuerschuld, aus der gemäß prognostizierter Verzögerungen (Lags) im dritten Schritt die Steuerzahlung beziehungsweise das kassenmäßige Steueraufkommen abgeleitet wird. Außerdem berücksichtigt der AKS die für den Prognosezeitraum relevanten beschlossenen oder sehr wahrscheinlichen Steuerrechtsänderungen.

Dieses Vorgehen spiegelt sich in der rechnerischen Zerlegung der Aufkommenselastizität in die beiden Komponenten Besteuerungsmengenelastizität und Steuersatzelastizität wider. Die Besteuerungsmengenelastizität gibt die relative Änderung der Bemessungsgrundlage einer Steuer bezogen auf die relative Änderung des nominalen Bruttoinlandsproduktes an. Die Steuersatzelastizität berechnet sich als Wachstums-

rate des Steuertarifs dividiert durch die Wachstumsrate der Bemessungsgrundlage (vgl. Diller (1987), S. 294). Alternativ kann auch die Steuerschuldelastizität, nämlich das Verhältnis der Zuwachsrate des Steueraufkommens zur Wachstumsrate der Bemessungsgrundlage, gebildet werden. Werden diese Komponenten formal verknüpft, ergibt sich wieder die Aufkommenselastizität (für den genauen mathematischen Zusammenhang siehe beispielsweise Rürup und Schachler (1983), S. 220).

Jeder einzelne der oben beschriebenen Prognoseschritte birgt jedoch zahlreiche Probleme. So stellt bereits in vielen Fällen die **Vorausschätzung der Steuerbemessungsgrundlage** ein großes Problem dar. Der AKS bekommt die Konjunkturprognosen des Arbeitskreises „Gesamtwirtschaftliche Vorausschätzungen“ vorgegeben, die jedoch häufig selbst schon mit großen Prognosefehlern behaftet sind. Außerdem handelt es sich wie bereits erwähnt um Zielprojektionen der Regierung. Da die Bemessungsgrundlage vieler Steuerarten nominale Einkommensgrößen und in Geld bewertete Mengen sind, spielt dabei auch die Prognose der Preisentwicklung eine wichtige Rolle. Für die Bemessungsgrundlage einiger Steuern sind außerdem keine geeigneten Daten vorhanden, so dass auf **Ersatzgrößen** zurückgegriffen werden muss. So fehlt z. B. eine geeignete gesamtwirtschaftliche Gewinngröße (vgl. Schaft (1981), S. 490). Für die tatsächliche Bemessungsgrundlage der veranlagten Einkommensteuer werden als Ersatzgröße die Bruttoeinkommen aus Unternehmertätigkeit und Vermögen verwendet, die jedoch als Residualgröße mit entsprechendem Prognosefehler in der Volkswirtschaftlichen Gesamtrechnung ermittelt werden. Die Vorauszahlungen, die den Hauptteil des Aufkommens an veranlagter Einkommens- und Körperschaftssteuer ausmachen, stehen jedoch weder mit dieser Hilfsgröße noch mit den Gewinnen in einem engen Zusammenhang (vgl. von der Lippe (1986), S. 337). Außerdem liegen in der VGR und im Steuerrecht unterschiedliche Definitionen für Gewinne und Abschreibungen vor (vgl. Institut für Finanzen und Steuern (1988), S. 15).

Je größer die Aufkommenselastizität der jeweiligen Steuer ist, desto größer ist auch die Unsicherheit der Steuerschätzung, da sich Prognosefehler der Bemessungsgrund-

lage dann besonders stark niederschlagen. Vor allem bei den hinsichtlich ihres Aufkommens sehr gewichtigen Steuern, nämlich der Lohn- und Einkommensteuer sowie der Umsatzsteuer, werden in der Literatur die größten Elastizitäten angegeben (vgl. z. B. Rürup und Schachler (1983), S. 222).

Verbrauchssteuern weisen dagegen meist geringe Elastizitäten auf. Außerdem lässt sich bei den Verbrauchssteuern die Verbrauchsmenge meist relativ gut prognostizieren (vgl. von der Lippe (1986), S. 337). Häufig werden dazu direkte Prognosemethoden verwendet, wie z. B. die Trendregression anhand der logistischen Funktion zur Modellierung von Sättigungstendenzen. Einerseits kann allerdings der Wendepunkt der logistischen Funktion in den Prognosezeitraum fallen, andererseits können Steuerrechtsänderungen zu Verschiebungen des Sättigungsniveaus führen (vgl. Rürup und Schachler (1983), S. 222). Des Weiteren sorgt die Interdependenz zwischen Verbrauchsmenge und Steuersatz für Schwierigkeiten, ebenso wie die Schätzung der Preiskomponente, wenn die Verbrauchssteuer wertmäßig (und nicht mengenmäßig) erhoben wird. So stellt sich bei einer Steuererhöhung auch die Frage, wie stark die Verbrauchsmengen reagieren werden (vgl. von der Lippe (1986), S. 337). Dies ist wiederum stark von der Überwälzungstechnik der Anbieter und den Nachfrageelastizitäten abhängig (vgl. Körner (1983), S. 218).

Auch der nächste Schritt, die **Prognose der Steuerschuld durch Anwendung des Steuertarifs** auf die Bemessungsgrundlage, birgt weitere Probleme, sofern kein einheitlicher Steuersatz vorliegt. Denn vor allem die progressive Besteuerung bei der Einkommensteuer setzt Kenntnisse über die aktuelle Einkommensverteilung sowie deren Veränderungen im Prognosezeitraum voraus. Auch Freibeträge und Freigrenzen sowie differenzierte Steuersätze aufgrund anderer Merkmale als der Höhe der Bemessungsgrundlage erfordern die **Prognose einer disaggregierten Bemessungsgrundlage**. So reicht es nicht aus, die Höhe der steuerpflichtigen Tatbestände zu prognostizieren, auch deren Verteilung muss berücksichtigt und oftmals geschätzt werden (vgl. von der Lippe (1986), S. 337).

Schließlich stellt sich die **Prognose der zeitlichen Verzögerungen** (Lagstruktur) zwischen Entstehen der Steuerschuld und kassenmäßiger Einnahme als schwierig dar. Dabei sind Steuervorauszahlungen, -rückstände und -erlasse, Niederschlagungen und Steuererstattungen zu berücksichtigen (vgl. von der Lippe (1986), S. 338). Vor allem bei den Veranlagungssteuern spielt die komplizierte und instabile Lag-Struktur eine entscheidende Rolle bei der Prognose des Steueraufkommens, so dass die „Erfolge, diese diskontinuierlichen Prozesse in den Schätzergebnissen zu erfassen, bescheiden sind“ (Schoof (1998), S. 213).

Nicht zuletzt sorgen die häufigen **Steuerrechtsänderungen** für erhebliche Probleme bei der Vorausschätzung des Steueraufkommens, da sie einerseits zu vielen Strukturbrüchen in den Zeitreihen der Steuereinnahmen führen. Andererseits sind die zukünftigen Auswirkungen der für den Prognosezeitraum beschlossenen und geplanten Steuerrechtsänderungen extrem schwer abzuschätzen, da sie meist stark vom Verhalten der Zensiten abhängen. Auch die zeitliche Wirkung einer Steuerrechtsänderung ist vor allem von den schwer vorhersehbaren Reaktionen der Steuerpflichtigen abhängig, was zu zusätzlichen Schwierigkeiten bei der periodengerechten Zuordnung der prognostizierten Effekte führt (vgl. Gebhardt (2001), S. 132). So kann es durch die Ankündigung von Steuerrechtsänderungen bereits zu Vorzieh- oder Verzögerungseffekten kommen, die sich auf die Aufkommensentwicklung auswirken. Andererseits besteht auch die Möglichkeit, dass die Zensiten z. B. aufgrund vertraglicher Bindungen erst verzögert auf Steuerrechtsänderungen reagieren und sich die verschobenen Verhaltensänderungen dementsprechend auch erst stark verzögert auf das Steueraufkommen auswirken, was noch durch die steuerrechtlichen Veranlagungs- und Zahlungslags verstärkt wird (vgl. Leibrecht (2004), S. 51). Außerdem entsteht eine weitere Schwierigkeit dadurch, dass sich die Auswirkungen verschiedener Steuerrechtsänderungen oftmals überlagern. Ein zusätzliches **Identifikationsproblem** ergibt sich durch etwaige gleichzeitige geldpolitische oder andere fiskalpolitische Maßnahmen, so dass sich der Effekt einer Steuerrechtsänderung noch schwerer prognostizieren lässt (vgl. Leibrecht (2004), S. 50).

Der Arbeitskreis „Steuerschätzungen“ versucht in seinen Steuerprognosen die Primärwirkungen der bereits beschlossenen Steuerrechtsänderungen (statische Effekte) zu berücksichtigen, wogegen die dynamischen makroökonomischen Rückwirkungen, so genannte **Feedback-Effekte**, bei der Quantifizierung der finanziellen Auswirkungen außer Acht bleiben (vgl. Schoof (1998), S. 211). Zwar versucht der Arbeitskreis „Gesamtwirtschaftliche Vorausschätzungen“ bei der Erstellung der konjunkturellen Eckdaten die makroökonomischen Impulse der Steuerrechtsänderungen zu berücksichtigen, ihre finanzwirtschaftlichen Wirkungen werden aber nicht explizit vom BMF ausgewiesen (vgl. Gebhardt (2001), S. 132). Zusätzlich führen zwar **geplante, aber noch nicht verabschiedete Steuerrechtsänderungen** zu weiteren Schätzrisiken, da sie nicht ins Prognosekalkül eingehen (vgl. z. B. Diller (1987), S. 295). Denn die Gesetzentwürfe unterliegen häufig umfangreichen Änderungen im Laufe des Gesetzgebungsverfahrens, so dass eine Berücksichtigung geplanter Steuerrechtsänderungen eine laufende Überarbeitung der Steuerschätzung erforderlich machen würde (vgl. Bundesregierung (1996), S. 4). Ihre finanziellen Auswirkungen werden allerdings laut BMF gesondert in der Haushalts- und Finanzplanung berücksichtigt (vgl. Bundesministerium der Finanzen (2005), S. 10).

Bei diesem indirekten (steuerschuldorientierten) Schätzverfahren, das auf makroökonomischer Ebene also die Steuerveranlagung eines Durchschnitts-Zensiten simuliert, werden viele **Input-Daten** benötigt, die alle selbst Prognosen für zukünftige Zeitpunkte darstellen, und damit **mit Prognosefehlern** behaftet sind. Sie eignen sich vor allem zur Vorhersage der Wirkung von Steuerrechtsänderungen, da hierfür die Kausalzusammenhänge unabdingbar sind. Ob allerdings die reine Prognose des Steueraufkommens tatsächlich mit dieser Methode treffsicherer ist als mit indirekten Verfahren, bleibt fraglich und soll im Folgenden untersucht werden.

2.3. Bisherige empirische Untersuchungen zur Qualität der AKS-Prognosen

Obwohl die Steuerschätzungen des AKS die Grundlage der Aufstellung des Bundeshaushalts sowie der Haushalte der Länder sind, finden sich in der Literatur nur vergleichsweise wenige empirische Untersuchungen zu deren Prognosegüte. Auffällig ist dabei besonders, dass es einige Veröffentlichungen Anfang bis Mitte der 80er Jahre gab (s. hauptsächlich Körner (1983), Rürup und Schachler (1983), Flascha (1985) und von der Lippe (1986)), das Thema dann in der wissenschaftlichen Diskussion weitestgehend verschwand und erst in den letzten zehn Jahren wieder von ein paar wenigen Autoren aufgegriffen wurde (vgl. Gebhardt (2001), Becker und Buettner (2007) sowie Büttner und Kauder (2008b) und Lehmann (2010)).

Körner (1983) führt sowohl eine horizontale als auch vertikale Treffsicherheitsanalyse durch. Bei der horizontalen Treffsicherheitsanalyse wird die Entwicklung der Prognosegenauigkeit bei zeitlich zunehmender Schätzdistanz im Rahmen von Prognosen für verschiedene Prognoseziele untersucht (**fixed event**). Die vertikale Treffsicherheitsanalyse betrachtet dagegen die Prognosegüte der Prognosen, denen ein bestimmter gleichbleibender Prognosehorizont zugrunde liegt (**fixed horizon**) (vgl. z. B. Osterloh (2008), S. 28).

Ausgangspunkt für Körners **horizontale Treffsicherheitsanalyse** ist ein einfacher absoluter Soll-Ist-Vergleich der vom AKS für die Jahre 1969 bis 1980 prognostizierten Werte für das Gesamtsteueraufkommen der BRD und der tatsächlich realisierten Werte. Allerdings führt Körner eine teilweise **Bereinigung** des Steueraufkommens durch, indem er all diejenigen Steuerrechtsänderungen mit Rückwirkung auf das Aufkommen im mittelfristigen Schätzzeitraum eliminiert, die in der auf die Prognose folgenden Periode in Kraft getreten sind (vgl. Körner (1983), S. 240). Als relative Fehlermaße werden dann der relative gewogene Schätzfehler und die Abweichungselastizität sowie ein modifizierter Theilscher Ungleichheitskoeffizient betrachtet (für die genauen Formeln siehe Körner (1983), S. 240). Ein Ergebnis der horizontalen

Treffsicherheitsanalyse ist, dass die Prognosen bis 1976 von Unterschätzungen geprägt waren und danach mehr Überschätzungen zu beobachten sind. Außerdem stellt Körner fest, dass sich die Prognosefehler des Bruttosozialproduktes stark auf die Prognosefehler des Steueraufkommens auswirken (vgl. Körner (1983), S. 240).

Die **vertikale Treffsicherheitsanalyse** führt Körner mit Hilfe eines linearen Regressionsansatzes durch, der als abhängige Variable das realisierte Steueraufkommen verwendet, das durch das (als unabhängig angenommene) prognostizierte Steueraufkommen erklärt werden soll. Dabei wird für die Prognosehorizonte $h = 1$ bis $h = 5$ jeweils eine separate Gleichung geschätzt (allerdings mit einer jeweils relativ geringen Anzahl an Beobachtungen) und das nicht besonders überraschende Ergebnis erzielt, dass die Prognosefehler mit wachsender Schätzdistanz stark zunehmen (vgl. Körner (1983), S. 244).

Horizontale Treffsicherheitsanalysen von einigen Einzelsteuern ergeben wesentlich schlechtere Ergebnisse als für das gesamte Steueraufkommen. Dies liegt an der kompensierenden Wirkung von Über- und Unterschätzungen einzelner Steuerarten. Dabei weisen die Verbrauchssteuern eine erheblich bessere Prognosegüte auf als die Lohn- und Umsatzsteuer, deren Treffsicherheit Körner noch mit „befriedigend“ bezeichnet, während die Steuerarten, die im Veranlagungsverfahren erhoben werden, so große Prognosefehler verzeichnen, dass ihre Treffsicherheit vom Autor als „ungenügend“ eingestuft wird (vgl. Körner (1983), S. 248).

Auch Rürup und Schachler (1983) untersuchen die absoluten und relativen Abweichungen der Schätzergebnisse des AKS von den tatsächlichen Werten des Gesamteueraufkommens für die Jahre 1975 bis 1981. Für diesen Zeitraum beobachten die Autoren, dass „in Zeiten einer anziehenden Inflationierung des Preisniveaus das Aufkommen unter-, in Zeiten einer deflatorischen Unterauslastung das Aufkommen tendenziell überschätzt wird“ (Rürup und Schachler (1983), S. 219).

Zu ähnlichen Ergebnissen kommt auch Flascha (1985). Sie schließt aufgrund ihrer Untersuchung der Prognosegüte der offiziellen Steuerschätzungen des AKS für

die Einzelsteuern, dass „die 30% des Steueraufkommens, die mit einem Schätzfehler $> 8\%$ behaftet sind, eine deutliche Sprache für die notwendige Verbesserung der Prognosequalität“ sprechen (Flascha (1985), S. 198). Zum Vergleich erstellt Flascha eigene Prognosen anhand von saisonalen ARIMA-Schätzungen sowohl für das Gesamtsteueraufkommen als auch für ausgewählte Einzelsteuern. Ihren Schätzungen liegen Quartalsdaten vom ersten Quartal 1969 bis zum ersten Quartal 1984 des unbeeinigten kassenmäßigen Gesamtsteueraufkommens sowie Monatsdaten von Januar 1976 bis April 1984 des Aufkommens einiger ausgewählter Einzelsteuern zu Grunde. Allerdings hatte Flascha im Jahre 1985 noch große **Probleme bei der elektronischen Datenverarbeitung**. So scheiterten beispielsweise laut Flascha (1985) gewünschte Modellspezifikationen anhand des von ihr verwendeten Programmpaketes SPSS (vgl. Flascha (1985), S. 165 und 170). Außerdem hatte sie bei den teilweise sehr rechenintensiven ARIMA-Schätzungen zur damaligen Zeit noch mit großen Schwierigkeiten aufgrund der eingeschränkten Rechnerkapazität zu kämpfen.

Um so beachtlicher ist daher, dass Flaschas kurz- und mittelfristige Prognosen für das Gesamtsteueraufkommen für den untersuchten Zeitraum eine größere Treffsicherheit als die offiziellen AKS-Schätzungen aufweisen (s. Flascha (1985), S. 169). Zur Beurteilung der Prognosegüte vergleicht Flascha jeweils die durchschnittlichen relativen Prognosefehler der offiziellen AKS-Schätzungen mit denen ihrer eigenen Prognosen. Auch für die Lohnsteuer als Einzelsteuer ergibt sich laut Flascha eine „deutliche Überlegenheit der univariaten Schätzmethode“ (Flascha (1985), S. 175). Weiterhin kann Flascha durch Anwendung der Box-Jenkins-Methode für die veranlagte Einkommensteuer und die Körperschaftssteuer kleinere durchschnittliche relative Prognosefehler erzielen als diejenigen der AKS-Schätzungen, obwohl ihre ARIMA-Schätzungen für sich genommen eine eher schwache Prognosegüte aufweisen (vgl. Flascha (1985), S. 186, 188). Für die Umsatz- und die Vermögenssteuer erreicht die Zeitreihenmethode immerhin noch gleichwertige Ergebnisse im Vergleich zu den AKS-Schätzungen (vgl. Flascha (1985), S. 178 und 191). Einzig die Prognose der Mineralölsteuer schneidet deutlich schlechter als die offizielle Schätzung ab (vgl. Flascha (1985), S. 181). Zusammenfassend schließt Flascha damit, dass „für

das reine Prognoseziel oder als Kontrollschätzung univariate Methoden in der Regel zu empfehlen“ sind (Flascha (1985), S. 199).

Einen weiteren Beitrag zur Untersuchung der Treffsicherheit der Prognosen des Arbeitskreises „Steuerschätzungen“ in den achtziger Jahren liefert von der Lippe (1986). Im von ihm analysierten Beobachtungszeitraum von 1960 bis 1981 kommt von der Lippe zu dem Ergebnis, dass 85% des Steueraufkommens nur mit einem durchschnittlichen Prognosefehler von mindestens $\pm 5\%$ vom AKS geschätzt werden können (vgl. von der Lippe (1986), S. 339). Seine Schlussfolgerung lautet, dass sich „für fast alle Steuern die Notwendigkeit [ergibt], die Prognosequalität durch verbesserte Schätzmethode zu steigern“ (von der Lippe (1986), S. 340).

15 Jahre später untersucht Gebhardt (2001) erneut die Treffsicherheit der mittelfristigen AKS-Schätzungen für das jeweils laufende Jahr, nun aber für den wesentlich größeren Beobachtungszeitraum von 1970 bis 2000. Das Jahr 1990 wurde wegen der Wiedervereinigung Deutschlands eliminiert (vgl. Gebhardt (2001), S. 133). Auch Gebhardt betrachtet hauptsächlich die absoluten und relativen Abweichungen der prognostizierten von den beobachteten Werten für das Steueraufkommen. Allerdings liegen dieser Analyse der Prognosegenauigkeit jeweils die ex post **um Steuerrechtsänderungen bereinigten Prognosen** zu Grunde, obwohl die Haushaltsplanungen der Regierung immer auf den unbereinigten AKS-Schätzungen basieren (müssen). In den von Gebhardt untersuchten drei Jahrzehnten ist die durchschnittliche Überschätzung des Gesamtsteueraufkommens größer als die durchschnittliche Unterschätzung (vgl. Gebhardt (2001), S. 133). Er kommt ebenfalls zu dem nicht überraschenden Ergebnis, dass die Prognosefehler des Bruttoinlandsproduktes entsprechende Schätzfehler des Steueraufkommens nach sich ziehen.

Besonders hohe Abweichungen der Prognose von den tatsächlichen Steuereinnahmen lassen sich für die Jahre 1995 bis 1997 beobachten. Die „Schätzungen mussten wiederholt in zweistelliger Milliardenhöhe nach unten korrigiert werden“ (Gebhardt (2001), S. 137). Außerdem war damals erstmalig eine Korrekturphase von über drei

Jahren notwendig. Insgesamt bescheinigt Gebhardt den offiziellen AKS-Schätzungen aber eine hohe Treffsicherheit. Die hohen Prognosefehler Mitte der neunziger Jahre rechnet er hauptsächlich temporären Sonderfaktoren zu (für Details dieser Sonderfaktoren siehe Gebhardt (2001), S. 137-145). Trotzdem sind laut Gebhardt bei der Prognose des Steueraufkommens „weitere Verbesserungen angezeigt“ (Gebhardt (2001), S. 145).

Der AKS selbst ist sich seiner Daueraufgabe bewusst, die verwendete Schätzmethode laufend zu verbessern, und hält deshalb zusätzliche **Methodensitzungen** ab. So nennt das BMF im Jahr 2005 als aktuelle Projekte den Einsatz eines Mikrosimulationsmodells bei der Schätzung der Lohnsteuer sowie die Nutzung ökonomischer Ansätze bei der Schätzung der Umsatzsteuer (vgl. Bundesministerium der Finanzen (2005), S. 14). Außerdem ist eine Verbesserung der Datengrundlage im Zeitablauf zu beobachten. So erwähnt das BMF 2005 „neue“ Daten zur Zahlungsstruktur und eine „neue“ Steuerdatenbank, die die Schätzungen im BMF unterstützt (vgl. Bundesministerium der Finanzen (2005), S. 14).

Zur Beurteilung der Schätzqualität des AKS veröffentlichte das BMF im Jahr 2005 ebenfalls einen Vergleich der Ergebnisse der November-Steuerschätzung zum jeweiligen Ist-Ergebnis des folgenden Haushaltsjahres für die Jahre von 1994 bis 2003 (vgl. Bundesministerium der Finanzen (2005), S. 17). Allerdings ist auch bei dieser Analyse der Prognosegenauigkeit zu beachten, dass die **Schätzwerte ex post um Steuerrechtsänderungen bereinigt** wurden, die nach dem Schätzzeitpunkt für das zu schätzende Jahr in Kraft traten. In den betrachteten zehn Jahren fallen wiederum die hohen Überschätzungen der Jahre 1995 bis 1997 auf. Von 1998 bis 2000 unterschätzte der AKS die Gesamtsteuereinnahmen leicht, von 2001 bis 2003 kam es allerdings erneut zu sehr hohen Überschätzungen. Insgesamt musste der Bund in diesen drei Jahren mit 55 Milliarden Euro weniger als erwartet auskommen (vgl. Bundesministerium der Finanzen (2005), S. 17). Dabei bemerkte beispielsweise Karrenberg bereits in einem seiner Artikel in der Kommunalen Steuer-Zeitschrift im Juli

2001, dass der Deutsche Städtetag skeptisch gegenüber den Schätzungen des AKS sei (vgl. Karrenberg (2001), S. 125).

Becker und Buettner (2007) testen die Prognosefehler der AKS-Schätzungen erstmalig auf ihre **Prognosefunktionalität**, das heißt auf Unverzerrtheit und Informationseffizienz. **Unverzerrtheit** liegt dann vor, wenn der Erwartungswert des Prognosefehlers Null ist. **Effizienz** bedeutet, dass alle Informationen, die zum Zeitpunkt der Prognose zur Verfügung stehen, vollständig für die Prognose genutzt werden. Becker und Buettner (2007) untersuchen sowohl die Mai-Prognosen für das laufende Jahr sowie die November-Prognosen für das folgende Jahr im Prognosezeitraum 1971 bis 2005. Sie kommen zu dem Ergebnis, dass die **AKS-Schätzungen unverzerrt** sind und ein Großteil der Prognosefehler auf die Unsicherheit der Konjunkturprognose zurückzuführen ist (vgl. Becker und Buettner (2007), S. 8 ff.).

Um auf Informationseffizienz zu testen, nehmen Becker und Buettner (2007) folgende Größen in ihr Regressionsmodell auf, die als Informationen zum Zeitpunkt der Prognose zur Verfügung stehen. Einerseits sind den Schätzern die jeweiligen Prognosefehler des Vorjahres bekannt. Zum anderen wird die Abweichung der von der Regierung vorgegebenen Prognose des Bruttoinlandsproduktes von der Gemeinschaftsdiagnose der führenden wirtschaftswissenschaftlichen Forschungsinstitute in die Schätzgleichung aufgenommen und auf Signifikanz getestet (vgl. Becker und Buettner (2007), S. 6 f.). Außerdem wird mit Hilfe einer Dummy-Variable untersucht, ob es in den Jahren, in denen eine Bundestagswahl stattfand, ein größerer bzw. kleinerer Prognosefehler zu erwarten ist, als in den Jahren, in denen keine Wahl stattfand. Letztere Dummy-Variable für Wahljahre erweist sich als einzige der getesteten Größen als statistisch signifikant von Null verschieden (vgl. Becker und Buettner (2007), S. 11 ff.). Gemäß der Schätzgleichung des Prognosefehlers für das laufende Jahr lässt sich demnach erwarten, dass der AKS in Wahljahren das Gesamtsteueraufkommen tendenziell unterschätzt, was dadurch erklärt werden könnte, dass der AKS bei seinen Steuerschätzungen die Möglichkeit nicht in Betracht zieht, dass die Bundesregierung in Wahljahren die Konjunktur möglichst positiv stimuliert

(vgl. Becker und Buettner (2007), S. 13). Alle anderen getesteten Informationen werden jedoch laut den Autoren bei der offiziellen Steuerschätzung **effizient** genutzt, so dass Becker und Buettner (2007) insgesamt zum Ergebnis kommen, dass im untersuchten Zeitraum kein signifikanter systematischer Fehler in den Steuerprognosen des AKS erkennbar ist.

Büttner und Kauder (2008a) untersuchen erstmalig im Auftrag des BMF die Prognosequalität offizieller Steuerschätzungen im internationalen Vergleich. Bei der Gegenüberstellung der relativen Vorhersagefehler für das Gesamtsteueraufkommen von zwölf OECD-Ländern im Prognosezeitraum 1998 bis 2008 schneidet die deutsche November-Schätzung vergleichsweise gut ab (vgl. Büttner und Kauder (2008a), S. 60). Die Mai-Schätzung ist naturgemäß mit einer größeren Unsicherheit behaftet, so dass die relativen Prognosefehler der betrachteten zehn Jahre im Vergleich zu den anderen Ländern eine eher große Spannweite aufweisen (vgl. Büttner und Kauder (2008a), S. 60). Auch Büttner und Kauder (2008b) können in ihrer Analyse die Nullhypothese unverzerrter Vorhersagefehler sowohl für die November- als auch Mai-Schätzungen nicht verwerfen (Büttner und Kauder (2008b), S. 32).

Auch Lehmann (2010) untersucht die AKS-Prognosen anhand verschiedener deskriptiver Gütemaße auf ihre Treffsicherheit. Er analysiert dabei die Mai-Schätzungen des Gesamtsteueraufkommens für das laufende sowie das folgende Jahr zwischen 1973 und 2008. Aufgrund der deutschen Wiedervereinigung wurden von ihm die Jahre 1989 und 1990 eliminiert (vgl. Lehmann (2010), S. 34 f.). Es zeigt sich, dass der AKS im von ihm betrachteten Zeitraum ebenfalls häufiger über- als unterschätzt. Neben der Berechnung einiger deskriptiver Maße führt Lehmann außerdem eine traditionelle Mincer-Zarnowitz-Regression¹ durch, um auf Unverzerrtheit der Prognosen zu testen. Wie auch schon Becker und Buettner (2007) und Büttner und Kauder (2008a) kommt Lehmann (2010) ebenfalls zu dem Ergebnis, dass die AKS-Schätzungen unverzerrt sind (vgl. Lehmann (2010), S. 37). Dabei gilt allerdings zu

¹Der traditionelle Test auf Unverzerrtheit anhand der Mincer-Zarnowitz-Tradition wird ausführlich in Kapitel 4.1.1 erläutert.

beachten, dass auch er separate Tests für die beiden von ihm betrachteten Prognosehorizonte durchführt.

Zusammenfassend kann festgehalten werden, dass die AKS-Steuerschätzungen gemäß bisheriger empirischer Untersuchungen als unverzerrt und in den meisten der untersuchten Fälle auch als effizient eingestuft werden können. Außerdem weisen sie eine relativ hohe Treffsicherheit auf. Dennoch kommt es immer wieder zu Jahren mit sehr starken Überschätzungen des Steueraufkommens, was enorme Auswirkungen auf die Finanzhaushalte aller Ebenen hat und politische Maßnahmen bis hin zu Haushaltssperren im laufenden Fiskaljahr notwendig macht. Im Folgenden soll nun untersucht werden, inwieweit saisonale ARIMA-Modelle mindestens vergleichbar treffsichere Steuerschätzungen liefern können. Des Weiteren stellt sich die Frage, ob die AKS-Prognosen auch dann noch als rational bezeichnet werden können, wenn gemäß eines neuen Ansatzes für alle Prognosehorizonte gemeinsam auf Unverzerrtheit und Informationseffizienz getestet wird.

3. Eigene Schätzung des Steueraufkommens mit (S)ARIMA-Modellen

Die Steuerschätzungen des AKS basieren hauptsächlich auf indirekten Schätzmethoden, für die Prognosewerte der erklärenden Input-Variablen benötigt werden, die zwangsläufig schon mit Prognosefehlern behaftet sind. Vor allem die gesamtwirtschaftlichen Vorausschätzungen bringen bereits ein hohes Maß an Unsicherheit in die Steuerschätzungen und sind zudem von der Regierung als Zielgröße vorgegeben. Um diese potentiellen Quellen für Fehlschätzungen zu umgehen, wird in diesem Kapitel das Gesamtsteueraufkommen der Bundesrepublik Deutschland mit einer **direkten stochastischen Methode** geschätzt. Die univariate Zeitreihenanalyse versucht nämlich eine Variable allein durch die Gesetzmäßigkeiten ihrer in der Vergangenheit beobachteten Realisationen zu erklären und zukünftige Werte dieser Variablen zu prognostizieren. Dadurch werden weder vergangene noch für die Zukunft prognostizierte Werte erklärender Variablen benötigt. Außerdem besteht keine Notwendigkeit, die komplexe Steuerstruktur, sich ändernde steuerrechtliche Gegebenheiten bzw. die vielfältigen Interdependenzen mit den konjunkturellen Größen in einem ökonometrischen Modell abzubilden.

3.1. Saisonale ARIMA-Modelle

George Box und Gwilym Jenkins entwickelten bereits im Jahr 1970 die Box-Jenkins-Methode, um den stochastischen Prozess, der einer univariaten Zeitreihe zugrunde liegt, durch „**Autoregressive Integrated Moving Average**“-Modelle (**ARIMA-Modelle**) abzubilden (Box und Jenkins (1970)). Es zeigte sich im Laufe der Zeit, dass diese Modelle in besonderem Maße für die Prognose von Zeitreihen mit ausgeprägtem saisonalen Muster geeignet sind.

Die in dieser Arbeit verwendete Notation ist an Hamilton (1994) angelehnt. Der zugrunde liegende stochastische Prozess beziehungsweise die Folge der Zufallsvariable über die (als diskret betrachtete) Zeit wird mit $\{Y_t\}$ bezeichnet, während $\{y_t\}$ für die tatsächliche Realisation des stochastischen Prozesses bzw. die Folge der Beobachtungswerte steht. Demnach bezeichnet (y_1, y_2, \dots, y_T) also eine konkrete Stichprobenrealisation. Aufgrund der Abhängigkeiten zwischen den Zufallsvariablen $\{\dots Y_{t-2}, Y_{t-1}, Y_t \dots\}$ und da nur eine einzige Realisation des stochastischen Prozesses beobachtet werden kann, muss eine Zeitreihe zwei Kriterien erfüllen, damit die Koeffizienten des gewünschten ökonomischen Modells konsistent geschätzt werden können. Diese beiden Kriterien sind die (schwache) Stationarität und die Ergodizität.

Ein stochastischer Prozess Y_t ist **schwach bzw. kovarianzen-stationär**, wenn (vgl. z. B. Lütkepohl (2004), S. 11)

$$\begin{aligned} E(Y_t) &= \mu \quad \forall t, \\ \text{Var}(Y_t) &= \sigma^2 \quad \forall t, \\ \text{Cov}(Y_t, Y_{t-j}) &= \gamma_j \quad \forall t \quad \text{und} \quad \forall j \geq 1. \end{aligned} \tag{3.1}$$

Der Erwartungswert und die Varianz sind also über die Zeit hinweg konstant. Die Autokovarianzen hängen lediglich von der Zeitspanne j zwischen zwei Zufallsvariablen ab, nicht aber vom Zeitpunkt t .

Dagegen spricht man von **striker Stationarität**, wenn für den stochastischen Prozess Y_t

$$F_{Y_{t_1}, \dots, Y_{t_T}}(y_1, \dots, y_T) = F_{Y_{t_1+j}, \dots, Y_{t_T+j}}(y_1, \dots, y_T) \quad \forall j \in \mathbb{Z}$$

gilt, also die gemeinsame Verteilungsfunktion zweier oder mehrerer Zufallsvariablen nur von j , nicht aber von t abhängt. Ist Y_t strikt stationär und besitzt endliche erste und zweite Momente, so ist der Prozess auch schwach stationär (vgl. z. B. Hamilton (1994), S. 46). Für den besonderen Fall, dass Y_t normalverteilt ist, impliziert schwache Stationarität auch strikte Stationarität. Im Folgenden wird der Begriff „stationär“ verkürzt für „schwach stationär“ verwendet.

Außerdem wird ein stochastischer Prozess **trend-stationär** genannt, wenn er nach Subtraktion einer (oft linearen) Funktion der Zeit t , dem so genannten deterministischen Trend, stationär ist (vgl. z. B. Lütkepohl (2004), S. 12). Dagegen spricht man von **Differenzen-Stationarität**, wenn Y_t aufgrund einer Einheitswurzel bzw. eines stochastischen Trends nicht stationär ist, aber die erste Differenz $\Delta Y_t = Y_t - Y_{t-1}$ einen stationären Prozess erzeugt. Der stochastische Prozess wird dann auch integriert der Ordnung 1 oder $I(1)$ -Prozess genannt. Ist eine d -fache Differenzenbildung nötig, um einen stationären Prozess zu erzeugen, heißt Y_t integriert der Ordnung d bzw. $I(d)$ -Prozess.

Das zweite Kriterium zur Schätzung von Zeitreihenmodellen ist die Ergodizität. Ein stationärer stochastischer Prozess Y_t ist **ergodisch**, wenn die Abhängigkeit zwischen den Zufallsvariablen Y_t und Y_{t-j} mit wachsender Zeitspanne j zwischen den Variablen immer mehr abnimmt. Hayashi (2000) definiert die Ergodizität eines stationären Prozesses dementsprechend als asymptotische Unabhängigkeit, das heißt, dass zwei Zufallsvariablen, die zeitlich weit genug entfernt sind, fast unabhängig

voneinander sind. Gilt für zwei beliebige beschränkte Funktionen $f : \mathbb{R}^k \rightarrow \mathbb{R}$ und $g : \mathbb{R}^l \rightarrow \mathbb{R}$ also die Grenzwertbeziehung

$$\lim_{n \rightarrow \infty} \mathbb{E} [f(z_i, \dots, z_{i+k})g(z_{i+n}, \dots, z_{i+n+l})] = \mathbb{E} [f(z_i, \dots, z_{i+k})] \mathbb{E} [g(z_{i+n}, \dots, z_{i+n+l})],$$

ist der stationäre stochastische Prozess Y_t ergodisch (vgl. Hayashi (2000), S. 101).

Hamilton (1994) schränkt seine Definition der Ergodizität auf stationäre normalverteilte stochastische Prozesse ein und erhält damit eine leichter prüfbare Bedingung für Ergodizität. Einen stationären normalverteilten stochastischen Prozess Y_t nennt man ihm zufolge nämlich ergodisch, wenn die Summe der Varianz und aller Autokovarianzen endlich ist, also

$$\sum_{j=0}^{\infty} |\gamma_j| < \infty \quad (3.2)$$

mit $\gamma_0 = \text{Var}(Y_t)$ und $\gamma_j = \text{Cov}(Y_t, Y_{t-j})$ für alle $j \geq 1$ gilt (vgl. Hamilton (1994), S. 47).

Ist ein stochastischer Prozess stationär und ergodisch, besagt das **Theorem der Ergodizität**, dass die Momente dieses Prozesses (sofern sie existieren und endlich sind) konsistent durch die Stichprobenmomente geschätzt werden können (vgl. beispielsweise Hayashi (2000), S. 102).

Um eine stationäre Zeitreihe zu modellieren, wird zunächst das ARMA-Modell betrachtet und dann schrittweise zum saisonalen ARIMA-Modell erweitert, welches in dieser Arbeit verwendet wird. Das allgemeine **ARMA(p, q)-Modell** setzt sich aus einer autoregressiven (AR) Komponente der Ordnung p und einer so genannten Moving Average (MA) Komponente der Ordnung q zusammen. Der AR-Teil bildet die Abhängigkeit der Variable von ihrer eigenen Vergangenheit ab, während der MA-Teil als gleitender Durchschnitt den Einfluss vergangener Schocks erfasst. Das

ARMA(p, q)-Modell lässt sich somit durch folgenden stochastischen Prozess $\{Y_t\}$ abbilden (vgl. z. B. Hamilton (1994), S. 59 oder Lütkepohl (2004), S. 27)

$$Y_t = c + \underbrace{\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p}}_{\text{AR-Teil}} + \underbrace{\theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}}_{\text{MA-Teil}} + \varepsilon_t. \quad (3.3)$$

Dabei stellt der Störterm $\{\varepsilon_t\}$ normalverteiltes weißes Rauschen dar, das heißt $\{\varepsilon_t\}$ erfüllt folgende Eigenschaften

$$\begin{aligned} E(\varepsilon_t) &= 0, \\ \text{Var}(\varepsilon_t) &= E(\varepsilon_t^2) = \sigma^2 \quad \forall t, \\ \text{Cov}(\varepsilon_t, \varepsilon_{t-j}) &= E(\varepsilon_t \varepsilon_{t-j}) = 0 \quad \forall j \neq 0, \\ \text{und } \varepsilon_t &\sim N(0; \sigma^2). \end{aligned} \quad (3.4)$$

Die zu schätzenden Koeffizienten werden mit $c, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$ bezeichnet. Die Zufallsvariable Y_t wird also durch die Linearkombination einer Konstanten, vergangener Werte der Zufallsvariablen und einem gegenwärtigen Schock sowie vergangener Schocks erklärt.

Als nützliche Notationshilfe wird der **Lag-Operator** L eingeführt. Dieser ist definiert als

$$L^k Y_t = Y_{t-k} \quad (3.5)$$

und erfüllt sowohl die Eigenschaft der Kommutativität, Assoziativität als auch der Distributivität (vgl. z. B. Hamilton (1994), S. 26 f.). Subtrahiert man den AR-Teil in Gleichung (3.3) und setzt den Lag-Operator aus Gleichung (3.5) ein, erhält man das ARMA(p, q)-Modell als

$$\underbrace{(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)}_{\text{AR Lag-Polynom}} Y_t = c + \underbrace{(1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q)}_{\text{MA Lag-Polynom}} \varepsilon_t. \quad (3.6)$$

Ersetzt man nun das **AR Lag-Polynom** durch

$$\Phi_p(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p \quad (3.7)$$

und das **MA Lag-Polynom** durch

$$\Theta_q(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q, \quad (3.8)$$

ergibt sich für das ARMA(p, q)-Modell die übersichtlichere Schreibweise

$$\Phi_p(L)Y_t = c + \Theta_q(L)\varepsilon_t. \quad (3.9)$$

Es kann gezeigt werden, dass jeder endliche Moving Average Prozess stationär und ergodisch ist (vgl. z. B. Hamilton (1994), S. 51). Ein Autoregressiver Prozess ist dagegen nur dann stationär und ergodisch, wenn die Lösungen des AR Lag-Polynoms aus Gleichung (3.7) außerhalb des Einheitskreises liegen bzw. betragsmäßig größer als eins sind (vgl. z. B. Kirchgässner und Wolters (2007), S. 49). Für den gemischten ARMA(p, q)-Prozess entscheiden folglich auch die Koeffizienten des AR-Teils, ob der Prozess stationär und ergodisch ist.

Außerdem kann jeder stationäre ARMA(p, q)-Prozess in einen unendlichen Moving Average-Prozess, kurz MA(∞)-Prozess, umgeformt werden (vgl. z. B. Hamilton (1994), S. 60). Da Herman Wold bereits 1938 zeigte, dass (unter bestimmten Regularitätsbedingungen) jede kovarianzen-stationäre Zeitreihe als ein MA(∞)-Prozess dargestellt werden kann (vgl. **Wold'sches Zerlegungstheorem** aus Wold (1938)), lässt sich die Schlussfolgerung ziehen, dass jede stationäre Zeitreihe durch einen ARMA(p, q)-Prozess modelliert werden kann. Wie man eine gegebene Zeitreihe auf Stationarität testet, wird im Abschnitt 3.3 im Zuge der Box-Jenkins-Methode direkt anhand der in dieser Arbeit untersuchten Steuerdaten erläutert.

Ist eine d -fache Differenzenbildung nötig, um aus dem nicht-stationären stochastischen Prozess Y_t einen stationären Prozess zu erzeugen, so wird der $I(d)$ -Prozess Y_t durch ein **ARIMA**(p, d, q)-Modell der Form

$$\underbrace{(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)}_{\text{Autoregressive}} \underbrace{(1 - L)^d}_{\text{Integrated}} Y_t = c + \underbrace{(1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q)}_{\text{Moving Average}} \varepsilon_t \quad (3.10)$$

dargestellt. In der verkürzten Schreibweise ergibt sich entsprechend

$$\Phi_p(L)(1 - L)^d Y_t = c + \Theta_q(L)\varepsilon_t. \quad (3.11)$$

Die durch d -fache Differenzenbildung gefilterte Zeitreihe $\tilde{Y}_t = (1 - L)^d Y_t$ hat dann wiederum die ARMA(p, q)-Repräsentation

$$\Phi_p(L)\tilde{Y}_t = c + \Theta_q(L)\varepsilon_t \quad (3.12)$$

(vgl. z. B. Lütkepohl (2004), S. 27).

Liegt zusätzlich ein ausgeprägtes **saisonales Muster** vor, eignen sich die saisonalen ARIMA-Modelle in besonderem Maße zur Prognose der Zeitreihe mit Saisonfigur. Der stochastische Prozess Y_t mit der saisonalen Periodizität s – beispielweise $s = 4$ für Quartalsdaten oder $s = 12$ für Monatsdaten – wird allgemein als multiplikatives **SARIMA**(p, d, q)(P, D, Q) $_s$ -Modell bezeichnet und lässt sich mit Hilfe des saisonalen AR Lag-Polynoms

$$\Phi_P(L^s) = 1 - \phi_{s1} L^s - \phi_{s2} L^{2s} - \dots - \phi_{sP} L^{sP} \quad (3.13)$$

und des saisonalen MA Lag-Polynoms

$$\Theta_q(L^s) = 1 + \theta_{s1} L^s + \theta_{s2} L^{2s} + \dots + \theta_{sQ} L^{sQ}, \quad (3.14)$$

als

$$\Phi_p(L)\Phi_P(L^s)(1-L)^d(1-L^s)^DY_t = c + \Theta_q(L)\Theta_Q(L^s)\varepsilon_t \quad (3.15)$$

darstellen. Zusätzlich zu den nicht-saisonalen AR und MA Operatoren sowie nicht-saisonalen Differenzen werden also auch noch saisonale AR und MA Operatoren sowie saisonale Differenzen in Betracht gezogen. Durch die multiplikative Verknüpfung werden auch die Interaktionen der nicht-saisonalen und saisonalen Effekte berücksichtigt (vgl. z. B. Enders (1995), S. 113) und es sind meist weniger Parameter zu schätzen, als es bei additiven saisonalen Komponenten der Fall wäre (vgl. z. B. Lütkepohl (2004), S. 27). Dies soll am Beispiel eines SARIMA(1,0,1)(1,0,1)₄-Modells für Quartalsdaten näher erläutert werden. Durch Ausmultiplizieren der multiplikativen Verknüpfungen erhält man nämlich

$$\begin{aligned} (1 - \phi_1 L)(1 - \phi_4 L^4)Y_t &= (1 - \theta_1 L)(1 - \theta_4 L^4)\varepsilon_t \\ (1 - \phi_1 L - \phi_4 L^4 + \phi_1 \phi_4 L^5)Y_t &= (1 - \theta_1 L - \theta_4 L^4 + \theta_1 \theta_4 L^5)\varepsilon_t \\ Y_t - \phi_1 Y_{t-1} - \phi_4 Y_{t-4} + \phi_5 Y_{t-5} &= \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_4 \varepsilon_{t-4} + \theta_5 \varepsilon_{t-5}. \end{aligned}$$

Während im multiplikativen SARIMA(1,0,1)(1,0,1)₄-Modell also nur 4 Koeffizienten geschätzt werden müssen, beinhaltet es trotzdem die durch Interaktion hervorgerufenen AR- und MA-Effekte des fünften Lags. Demgegenüber würde das entsprechende additive Modell die Schätzung zweier zusätzlicher Koeffizienten erfordern. Aus diesem Grund werden die multiplikativen SARIMA-Modelle den additiven meist vorgezogen.

Die hier vorgestellten SARIMA-Modelle werden im Folgenden verwendet, um das Gesamtsteueraufkommen der BRD zu schätzen und zu prognostizieren. Zuerst wird im nächsten Unterkapitel die Datengrundlage und Datenaufbereitung beschrieben. Anschließend werden dann die einzelnen Schritte der Schätzung und Prognose anhand der Box-Jenkins-Methode erläutert.

3.2. Datengrundlage und Datenaufbereitung

Die Datengrundlage dieser Arbeit bilden Quartalsdaten der kassenmäßigen Steuereinnahmen der Bundesrepublik Deutschland für die Jahre 1961 bis 2010. Als Datenquelle dienen die publizierten Tabellen des Statistischen Bundesamtes, bereitgestellt von Klaus-Jürgen Hammer und Bettina Bronz. Daraus wird zunächst ein Zeitreihendatensatz aus den jeweiligen Tabellen für die einzelnen Quartale der Jahre 1961 bis 2010 erstellt, so dass die Zeitreihe des Gesamtsteueraufkommens 200 Beobachtungswerte aufweist. In den betrachteten Zeitraum fallen die Wiedervereinigung sowie die Währungsumstellung von DM auf Euro. Bis einschließlich 1990 liegen nur die Daten für Westdeutschland vor, ab 1991 wird das gesamte Bundesgebiet betrachtet. Um diesen Strukturbruch zu modellieren, wird eine Dummyvariable „Wiedervereinigung“ generiert, die für alle Quartale ab dem Jahr 1991 den Wert eins annimmt. Alle Werte bis 2002 werden mit dem in der Verordnung (EG) Nr. 2866/98 des Rates der EU vom 31. Dezember 1998 fixierten Wechselkurs¹ in Euro umgerechnet, so dass alle Werte des Gesamtsteueraufkommens der BRD von 1961 bis 2010 zur weiteren Analyse in Millionen Euro vorliegen.

¹1 Euro = 1,95583 DM.

3.3. Schätzung des Steueraufkommens anhand der Box-Jenkins-Methode

Die von George Box und Gwilym Jenkins im Jahr 1970 entwickelte Box-Jenkins-Methode lässt sich in die folgenden drei Phasen gliedern (vgl. auch Makridakis et al. (1998), S. 314).

Phase I: Identifikation

- Datentransformation, um eine stationäre Zeitreihe zu erzeugen:
 - eventuelles Logarithmieren zur Stabilisierung der Varianz
 - Tests auf Einheitswurzeln und eventuelles Differenzieren bzw. Abzug eines deterministischen Trends
- Modellspezifikation:
 - Untersuchung der Autokorrelationsfunktion (ACF) und partiellen Autokorrelationsfunktion (PACF), um mögliche Modellspezifikationen zu identifizieren

Phase II: Schätzen und Testen

- Modellschätzungen und -auswahl:
 - Maximum-Likelihood-Schätzungen der möglichen Modellvarianten
 - Modellauswahl anhand verschiedener (Informations-)Kriterien
- Überprüfung der Residuen auf weißes Rauschen:
 - Überprüfung der Autokorrelationsfunktion und partiellen Autokorrelationsfunktion der Residuen
 - Portmanteau-Tests

Phase III: Prognose

- Punktschätzer und Konfidenzintervalle

Die einzelnen Schritte werden im weiteren Verlauf für die Zeitreihe des Gesamtsteueraufkommens der BRD detailliert erläutert.

3.3.1. Phase I: Identifikation

Datentransformationen

Im ersten Schritt des Box-Jenkins-Verfahrens wird die Zeitreihe zunächst betrachtet, um erste Vermutungen bezüglich Ausreißer, Strukturbrüche und Stationarität anzustellen. Abbildung 3.1 zeigt das Gesamtsteueraufkommen der BRD vom ersten Quartal 1961 bis zum vierten Quartal 2010 in Millionen Euro.

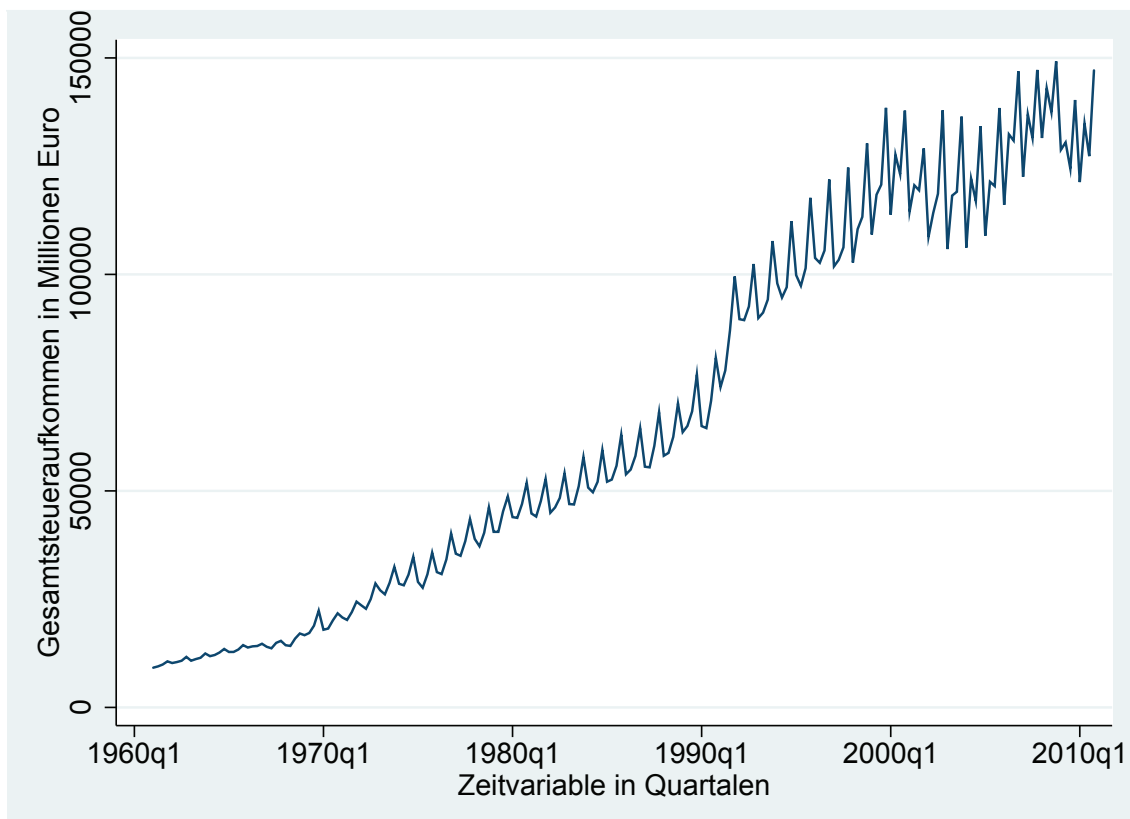


Abbildung 3.1.: Abbildung des Gesamtsteueraufkommens der BRD vom ersten Quartal 1961 bis zum vierten Quartal 2010 in Millionen Euro

Beim Betrachten der Zeitreihe fallen keine Ausreißer auf. Allerdings ist im Jahr 1991 ein Sprung des Gesamtsteueraufkommens zu beobachten, der auf einen durch die Wiedervereinigung bedingten Strukturbruch hindeutet. Außerdem ist klar zu erkennen, dass die Zeitreihe höchstwahrscheinlich nicht stationär ist. Die Varianz scheint im Zeitverlauf zuzunehmen und es ist ein deutlicher positiver Trend erkennbar. Zusätzlich sticht die stark ausgeprägte Saisonfigur ins Auge. Auf Grundlage dieser ersten Beobachtungen werden die einzelnen Vermutungen im Folgenden anhand statistischer Methoden und Tests untersucht.²

Als Erstes wird die scheinbar im Zeitablauf zunehmende **Varianz** des Gesamtsteueraufkommens analysiert. Die üblichen Tests auf Homoskedastizität, nämlich der Breusch-Pagan-Test, der White-Test, der Goldfeld-Quandt-Test sowie der Cook-Weisberg-Test³, können alle die Nullhypothese der Homoskedastizität auf dem 1%-Signifikanzniveau eindeutig verwerfen. Um eine im Zeitverlauf konstante Varianz zu erzeugen, wird der natürliche Logarithmus der Zeitreihe gebildet, was einer Box-Cox-Transformation mit dem Parameter $\lambda = 0$ entspricht (vgl. Box und Cox (1982), S. 209).

Abbildung 3.2 zeigt das **logarithmierte Gesamtsteueraufkommen** der BRD vom ersten Quartal 1961 bis zum vierten Quartal 2010. Die Varianz scheint nun annähernd konstant im Zeitverlauf zu sein und bei einem weiteren Breusch-Pagan- bzw. Cook-Weisberg-Test kann die Nullhypothese der Homoskedastizität auf dem 5% Signifikanzniveau nun nicht mehr verworfen werden.

Da die Zeitreihe weiterhin nicht stationär im Mittelwert ist, werden im nächsten Schritt verschiedene **Tests auf Einheitswurzeln** durchgeführt. Aufgrund der stark ausgeprägten Saisonfigur könnten **saisonale Einheitswurzeln** vorliegen. Deshalb wird zunächst ein so genannter **HEGY-Test** für Quartalsdaten gemäß Hyllenberg et al. (1990) durchgeführt. Im geschätzten Modell (s. Hyllenberg et al. (1990),

²Einzelheiten zu den im Folgenden durchgeführten Tests sind der Stata do- bzw. log-Datei zu entnehmen, die bei Interesse direkt von der Autorin erhältlich ist.

³Eine detaillierte Beschreibung dieser Tests auf Homoskedastizität findet sich beispielsweise in Greene (2003), S. 222 ff.

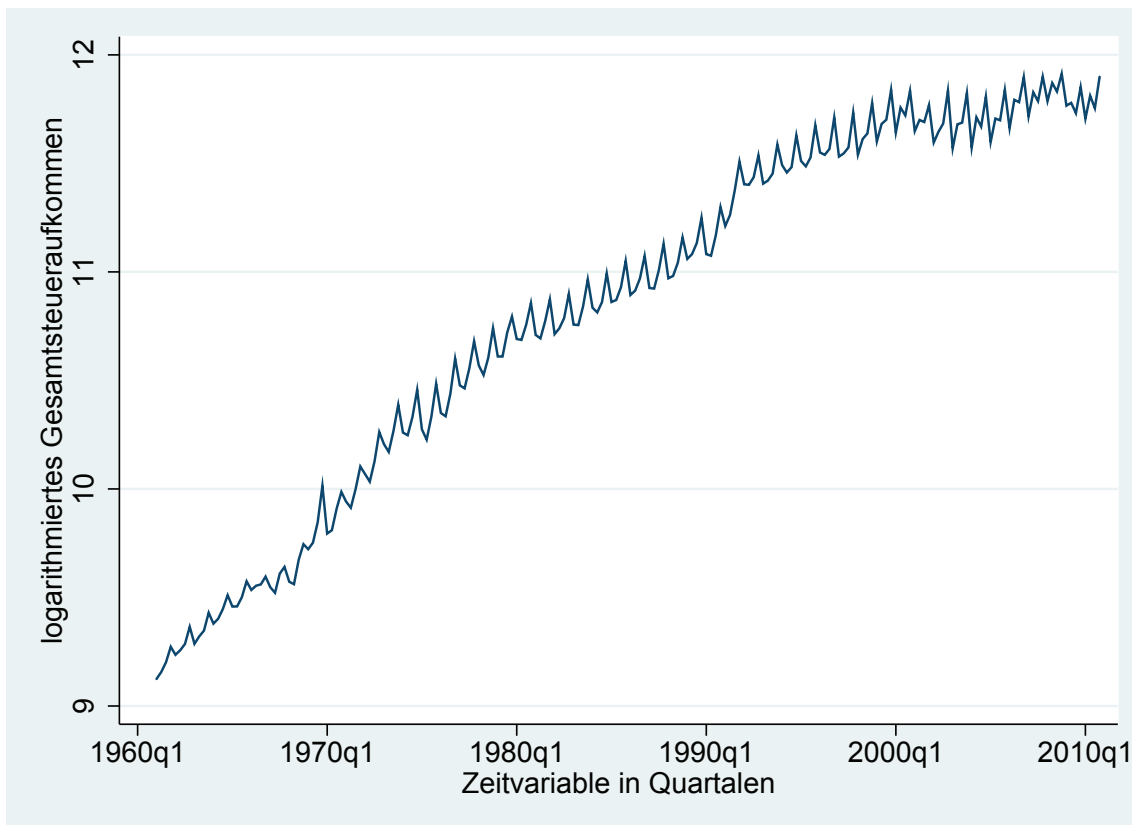


Abbildung 3.2.: Abbildung des logarithmierten Gesamteueraufkommens der BRD vom ersten Quartal 1961 bis zum vierten Quartal 2010

S. 223 f.) werden zusätzlich eine Konstante, drei Saisondummies und ein linearer Trend als deterministische Variablen aufgenommen. Die Nullhypothese saisonaler Einheitswurzeln ($H_0 : \pi_3 = \pi_4 = 0$) kann durch den F -Test allerdings klar verworfen werden. Das Testergebnis ist außerdem robust gegenüber der Wahl verschiedener Lag-Längen sowie der Berücksichtigung des linearen Trends. Des Weiteren deutet der zugehörige t -test auf eine nicht-saisonale Einheitswurzel hin, da die Nullhypothese $H_0 : \pi_1 = 0$ nicht verworfen werden kann. Zusätzlich wird der alternative „root“-Test von Depalo (2009) durchgeführt, der zum gleichen Testergebnis kommt. Auch bei diesem Test auf saisonale Einheitswurzeln werden in der Regressionsgleichung eine Konstante, drei Saisondummies und ein linearer Trend berücksichtigt.

Um den möglichen Strukturbruch durch die Wiedervereinigung im Jahr 1990 zu beachten, wird der HEGY-Test auf saisonale Einheitswurzeln ebenfalls für die beiden Teilperioden vor und nach der Wiedervereinigung durchgeführt. Für die Zeit vor dem ersten Quartal 1990 werden die bisherigen Testergebnisse bestätigt. Für die zweite Teilperiode nach der Wiedervereinigung kann die Nullhypothese saisonaler Einheitswurzeln zwar nicht mehr verworfen werden. Allerdings muss aufgrund der relativ geringen Anzahl an Beobachtungen (78 Quartale) davon ausgegangen werden, dass die Güte des sequentiellen HEGY-Tests zu gering ist (vgl. Depalo (2009), S. 427). Aufgrund der weitgehend übereinstimmenden Testergebnisse wurde **keine saisonale Differenzierung** vorgenommen, sondern die Vermutung einer nicht-saisonalen stochastischen Einheitswurzel beibehalten und weiter untersucht.

Im nächsten Schritt wird getestet, ob ein **stochastischer oder deterministischer Trend** vorliegt. Dabei spielt allerdings der mögliche Strukturbruch durch die Wiedervereinigung eine wichtige Rolle. Denn die üblichen **Augmented-Dickey-Fuller-Tests**⁴ (ADF-Tests) auf eine Einheitswurzel erweisen sich bei einem Strukturbruch als wenig sinnvoll (vgl. Perron (1989), S. 1368 ff.), da dieser dann fälschlicherweise als Testergebnis das Vorliegen einer Einheitswurzel ergeben kann. So zeigte Perron (1989) in seiner Monte-Carlo-Simulationsstudie auch, dass der ADF-Test inkonsistent ist, wenn sich die Steigung des deterministischen Trends durch einen Strukturbruch ändert (vgl. Perron (1989), S. 1372). Allerdings kann der ADF-Test für die beiden Teilperioden vor und nach der Wiedervereinigung getrennt angewandt werden (vgl. auch Kirchgässner und Wolters (2007), S. 176). Dabei wird jeweils die Nullhypothese des stochastischen Trends, also einer Einheitswurzel, gegen die Alternativhypothese eines deterministischen Trends getestet. Für die Zeitreihe bis zur Wiedervereinigung kann die Nullhypothese des stochastischen Trends nicht verworfen werden. Dieses Ergebnis erwies sich auch robust bezüglich der Wahl der Lag-Länge. Für die Teilperiode nach der Wiedervereinigung kann die Nullhypothese zwar auf dem 5%-Signifikanzniveau verworfen werden. Aufgrund der relativ gerin-

⁴Vgl. Dickey und Fuller (1979) oder Standardlehrbücher wie zum Beispiel Lütkepohl (2004), S. 54 ff.

gen Anzahl an Beobachtungen verliert der ADF-Test jedoch an Macht und dadurch hängt das Testergebnis entscheidend von der Wahl der Lag-Länge ab.

Wenn die Zeitreihe einen unbekanntem Mittelwert oder Trend aufweist, besitzt der modifizierte Dickey-Fuller-Test, bei dem gemäß Elliott et al. (1996) die Zeitreihe zuerst über eine verallgemeinerte Kleinstquadrate Regression (**GLS**) transformiert wird, eine höhere Macht. Aufgrund des Strukturbruchs wird auch dieser Test getrennt für die beiden Teilperioden angewandt. Die Nullhypothese der Einheitswurzel kann dabei sowohl für die Periode vor der Wiedervereinigung als auch für den Zeitraum danach auf allen Signifikanzniveaus nicht verworfen werden.

Eine Alternative zu den Tests auf eine Einheitswurzel bietet der **KPSS-Test** auf Trendstationarität nach Kwiatkowski et al. (1992). Bei diesem Test wird die Nullhypothese des deterministischen Trends gegen die Alternativhypothese des stochastischen Trends getestet. Für beide Teilperioden kann die Nullhypothese eindeutig verworfen werden, so dass sich die bisherigen Ergebnisse bestätigen.

Allerdings haben alle bisher durchgeführten Tests den Nachteil, dass die Zeitreihe aufgrund des Strukturbruchs in die beiden Teilperioden zerlegt werden muss und der Zeitraum nach der Wiedervereinigung eine verhältnismäßig geringe Anzahl an Beobachtungen aufweist. Pierre Perron entwickelte 1989 einen Test auf eine Einheitswurzel, bei dem ein **exogener Strukturbruch** zu einem im Voraus bekannten Zeitpunkt berücksichtigt werden kann (vgl. Perron (1989) und die Weiterentwicklungen Perron (1994, 1997)). Die Nullhypothese dieses Tests schließt einen Random Walk mit Drift sowie einen möglichen exogenen Strukturbruch ein. Als Alternativhypothese wird ein trendstationärer stochastischer Prozess angenommen, bei dem ebenfalls ein exogener Strukturbruch auftreten kann.

Bei diesem Test sind verschiedene Modellvarianten möglich (vgl. Perron (1989), S. 1364). Für die Zeitreihe des logarithmierten Gesamtsteueraufkommens der BRD (vgl. Abbildung 3.2) kommt einerseits Perrons Modell A in Betracht, das lediglich eine Niveaushift der Trendgeraden ab 1991 zulässt. Alternativ kommt auch

Perrons Modell C infrage, welches zusätzlich zur Niveauverschiebung auch eine Änderung der Steigung der Trendgeraden erlaubt (vgl. Perron (1989), S. 1364). Eine weitere Variationsmöglichkeit besteht darin, zu unterscheiden, ob von einer sofortigen Veränderung ohne entsprechende Anpassungsreaktionen ausgegangen wird – das so genannte additive outlier (AO) Modell – oder aber Anpassungen in den dem Strukturbruch folgenden Perioden gestattet werden – das innovational outlier (IO) Modell (vgl. Perron (1994), S. 118 f.). Bei allen Modellvarianten kann die Nullhypothese eindeutig nicht verworfen werden. Das Testergebnis bestätigt sich ebenfalls unter Berücksichtigung anderer deterministischer Wachstumspfade und ist robust gegenüber der Wahl der Lag-Länge.

Zivot und Andrews (1992) modifizierten den Test von Perron (Modell C), indem sie nicht mehr von einem exogenen Strukturbruch zu einem bekannten Zeitpunkt ausgehen, sondern in ihrer Alternativhypothese einen **endogenen Strukturbruch** im Niveau und in der Steigung des deterministischen Trends erlauben (vgl. Zivot und Andrews (1992), S. 254). Inwieweit die Hypothese eines endogenen Strukturbruchs für das hier betrachtete logarithmierte Gesamtsteueraufkommen der BRD aufgrund der historischen Gegebenheit der Wiedervereinigung allerdings sinnvoll ist, bleibt fraglich. Nichtsdestotrotz kann auch bei diesem Test die Nullhypothese einer Einheitswurzel nicht verworfen werden.⁵

Des Weiteren stellt sich die Frage, ob sich das Testergebnis ändert, wenn mehr als ein endogener Strukturbruch berücksichtigt wird. Der entsprechende Test von Clemente et al. (1998) ermöglicht zwar **zwei endogene Strukturbrüche** in Form von Niveauverschiebungen (Modell A), sowohl für das additive outlier Modell als auch das innovational outlier Modell, lässt allerdings in der Alternativhypothese keinen linearen Trend mehr zu. Insofern ist es nicht sehr verwunderlich, dass auch bei allen Modellvarianten dieses Tests, die Nullhypothese des stochastischen Trends nicht verworfen werden kann.

⁵Im Gegensatz zu Perrons Test beinhaltet die Nullhypothese des Tests von Zivot und Andrews (1992) jedoch nicht die Möglichkeit eines Strukturbruchs.

Da nahezu alle durchgeführten Tests auf eine **Einheitswurzel** hindeuten, wird die **erste Differenz** des logarithmierten Gesamtsteueraufkommens gebildet, um eine stationäre Zeitreihe zu erzeugen (Differenzenstationarität). Da nun also sowohl der natürliche Logarithmus als auch die erste zeitliche Differenz des Gesamtsteueraufkommens gebildet wurde, ergibt sich näherungsweise die Zeitreihe der **Wachstumsraten** des Gesamtsteueraufkommens der BRD vom ersten Quartal 1961 bis zum vierten Quartal 2010, die in Abbildung 3.3 dargestellt ist.

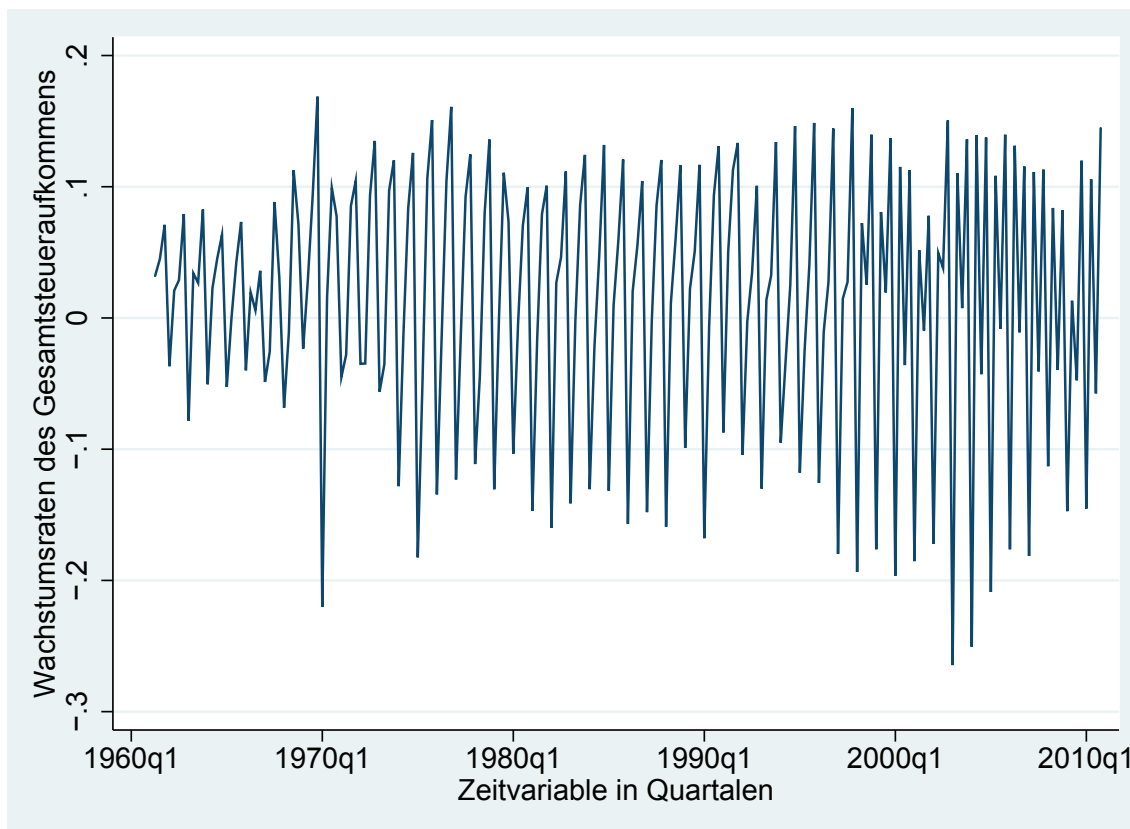


Abbildung 3.3.: Abbildung der Wachstumsraten des Gesamtsteueraufkommens der BRD vom ersten Quartal 1961 bis zum vierten Quartal 2010

Um zu überprüfen, ob die Zeitreihe nunmehr tatsächlich kovarianz-stationär ist, wird erneut ein ADF-Test durchgeführt. Die Nullhypothese der Einheitswurzel kann jetzt eindeutig verworfen werden.

Modellspezifikation

Damit ist die erforderliche Datentransformation abgeschlossen und es folgt der zweite Schritt der ersten Phase des Box-Jenkins-Verfahrens, nämlich die Modellspezifikation. Um mögliche SARIMA-Modelle zu identifizieren, die die beobachtete Zeitreihe beschreiben könnten, wird zunächst ein Vergleich der theoretischen und empirischen Autokorrelationsfunktion (ACF) und partiellen Autokorrelationsfunktion (PACF) durchgeführt.

Die **theoretische Autokorrelationsfunktion** φ_j eines stationären stochastischen Prozesses $\{Y_t\}$ berechnet sich als

$$\varphi_j = \frac{\gamma_j}{\gamma_0} = \frac{\text{Cov}(Y_t, Y_{t-j})}{\text{Var}(Y_t)}; \quad j = 0, 1, 2, \dots \quad (3.16)$$

mit $-1 \leq \varphi_j \leq 1$ (vgl. z. B. Hamilton (1994), S. 49). Als graphische Darstellung dieser ACF dient das so genannte Korrelogramm, das die Autokorrelationen φ_j in Abhängigkeit der verschiedenen Zeitabstände $j = 0, 1, 2, \dots$ abbildet. Für einfache AR(p)- und MA(q)-Prozesse lässt sich die theoretische ACF noch relativ leicht beschreiben (vgl. beispielsweise Makridakis et al. (1998), S. 342):

- AR(1)-Prozess: $\varphi_j = \phi^j$, d.h. exponentielle Abnahme mit positivem Vorzeichen für $\phi > 0$ und alternierendem Vorzeichen für $\phi < 0$.
- AR(p)-Prozess: exponentielle Abnahme oder gedämpfte Sinus-Welle, abhängig von Vorzeichen und Größe der Koeffizienten ϕ_1, \dots, ϕ_p .
- MA(1)-Prozess: $\varphi_1 = \frac{\theta}{1+\theta^2}$ und $\varphi_j = 0$ für $j > 1$, d.h. Ausschlag nur beim ersten Lag (positiv für $\theta > 0$ und negativ für $\theta < 0$).
- MA(q)-Prozess: Ausschläge nur für die Lags 1 bis q , da $\gamma_j = 0$ für $j > q$.

Für gemischte Modelle werden die theoretischen Autokorrelationen jedoch immer komplizierter.

Die **empirischen Autokorrelationen** werden anhand der tatsächlichen Beobachtungen (y_1, \dots, y_T) berechnet:

$$\hat{\varphi}_j = \frac{\hat{\gamma}_j}{\hat{\gamma}_0} = \frac{\frac{1}{T} \sum_{t=j+1}^T (y_t - \bar{y})(y_{t-j} - \bar{y})}{\frac{1}{T} \sum_{t=1}^T (y_t - \bar{y})^2} \quad \text{mit} \quad \bar{y} = \frac{1}{T} \sum_{t=1}^T y_t. \quad (3.17)$$

Für normalverteiltes weißes Rauschen gilt⁶

$$\hat{\varphi}_j \stackrel{a}{\sim} N\left(0; \frac{1}{T}\right), \quad (3.18)$$

so dass mit Hilfe eines einfachen t -Tests die Nullhypothese $H_0 : \varphi_j = 0$ gegen die Alternativhypothese $H_1 : \varphi_j \neq 0$ getestet werden kann. Statistikprogramme geben typischerweise im Korrelogramm die empirische ACF und das dazugehörige 95%-Konfidenzintervall (meist nach Bartlett) an. So können die signifikanten Ausschläge mit den theoretischen Autokorrelationen bestimmter (einfacher) AR- oder MA-Prozesse verglichen werden, um mögliche Modellspezifikationen abzuleiten. Außerdem lassen sich Saisonmuster durch entsprechende stark von Null abweichende empirische Autokorrelationen für $j = s, 2s, 3s, \dots$ erkennen.

Für die (stationären) Wachstumsraten des Gesamtsteueraufkommens der BRD ergibt sich zwischen dem ersten Quartal 1961 und dem vierten Quartal 2010 die in Abbildung 3.4 dargestellte empirische Autokorrelationsfunktion. Die starken Ausschläge für $j = 4, 8, 12, \dots$ spiegeln die ausgeprägte Saisonfigur der Zeitreihe wider. Der langsam abnehmende Verlauf lässt einen saisonalen autoregressiven Term vermuten (eventuell $P = 1$). Aber auch ein nicht-saisonaler AR-Term ist aufgrund der empirischen Autokorrelationen nicht ausgeschlossen. Nichtsdestotrotz ist keine eindeutige Modellspezifikation aus der empirischen ACF erkennbar.

⁶Siehe unter anderem Anderson (1942), Bartlett (1946) und Quenouille (1949).

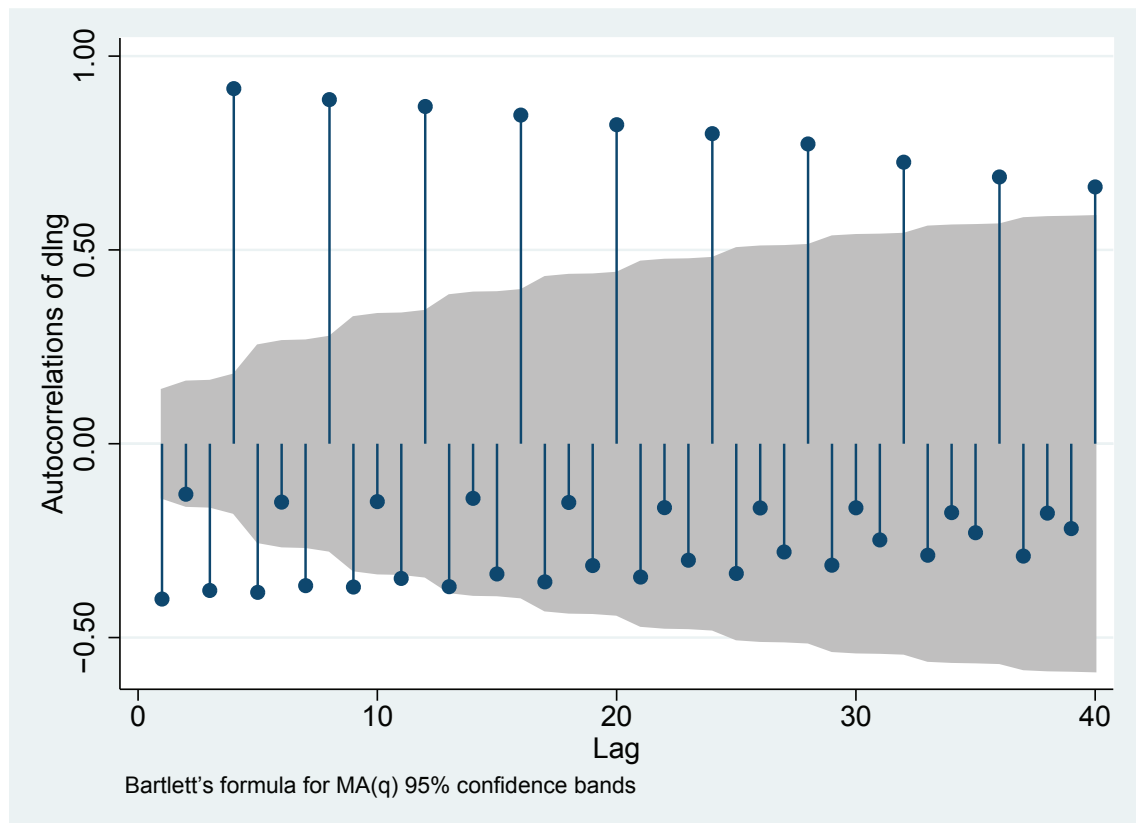


Abbildung 3.4.: Abbildung der empirischen Autokorrelationsfunktion der Wachstumsraten des Gesamteueraufkommens

Als zusätzliche Hilfe werden die **partiellen Autokorrelationen** untersucht. Sie messen den linearen Zusammenhang zwischen Y_t und Y_{t-k} , wobei der Einfluss der dazwischen liegenden Variablen ausgeschaltet wird. Die theoretische partielle Autokorrelation zwischen Y_t und Y_{t-k} ist also ihre bedingte Korrelation, gegeben der Werte für die dazwischenliegenden Zufallsvariablen $Y_{t-1}, \dots, Y_{t-k+1}$ (Lütkepohl (2004), S. 36):

$$\beta_k = \text{Corr}(Y_t, Y_{t-k} | Y_{t-1}, \dots, Y_{t-k+1}). \quad (3.19)$$

Der **empirische** partielle Autokorrelationskoeffizient $\hat{\beta}_k$ kann durch die OLS-Regression des entsprechenden AR(k)-Modells

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_k Y_{t-k} + \varepsilon_t$$

geschätzt werden (vgl. z. B. Hamilton (1994), S. 11). Box und Jenkins (1970) haben gezeigt, dass

$$\hat{\beta}_k \stackrel{a}{\sim} N\left(0; \frac{1}{T}\right), \quad (3.20)$$

so dass anhand eines t -Tests die Nullhypothese $\beta_k = 0$ gegen die Alternativhypothese $\beta_k \neq 0$ getestet werden kann. Die (signifikanten) empirischen partiellen Autokorrelationen $\hat{\beta}_k$ werden dann auch wiederum mit den theoretischen β_k einfacher AR- und MA-Prozesse verglichen. Dabei sind die Eigenschaften gerade spiegelbildlich zu denen der ACF (eine Übersicht findet sich beispielsweise bei Makridakis et al. (1998), S. 342):

- AR(1)-Prozess: $\beta_1 \neq 0$, $\beta_k = 0$ für $k > 1$, d.h. Ausschlag nur beim ersten Lag (positiv für $\phi > 0$ und negativ für $\phi < 0$).
- AR(p)-Prozess: Ausschläge nur für die Lags 1 bis p , da $\beta_k = 0$ für $k > p$.
- MA(1)-Prozess: exponentielle Abnahme, mit positivem Vorzeichen für $\theta > 0$ und alternierendem Vorzeichen für $\theta < 0$.
- MA(q)-Prozess: exponentielle Abnahme oder gedämpfte Sinus-Welle, abhängig von Vorzeichen und Größe der Koeffizienten $\theta_1, \dots, \theta_q$.

Für die Wachstumsraten des Gesamtsteueraufkommens der BRD ergibt sich zwischen dem ersten Quartal 1961 und dem vierten Quartal 2010 die in Abbildung 3.5 dargestellte empirische partielle Autokorrelationsfunktion. Auch bei der PACF spiegeln die signifikanten Ausschläge für $j = 4, 8, 12, 16$ die ausgeprägte Saisonfigur der Zeitreihe wider und deuten auf einen möglichen saisonalen MA-Term hin (eventuell $Q = 1$). Aber auch ein nicht-saisonaler MA-Term könnte aufgrund der empirischen partiellen Autokorrelationen vorliegen. Allerdings lässt sich auch mit Hilfe der PACF keine eindeutige Modellspezifikation ableiten.

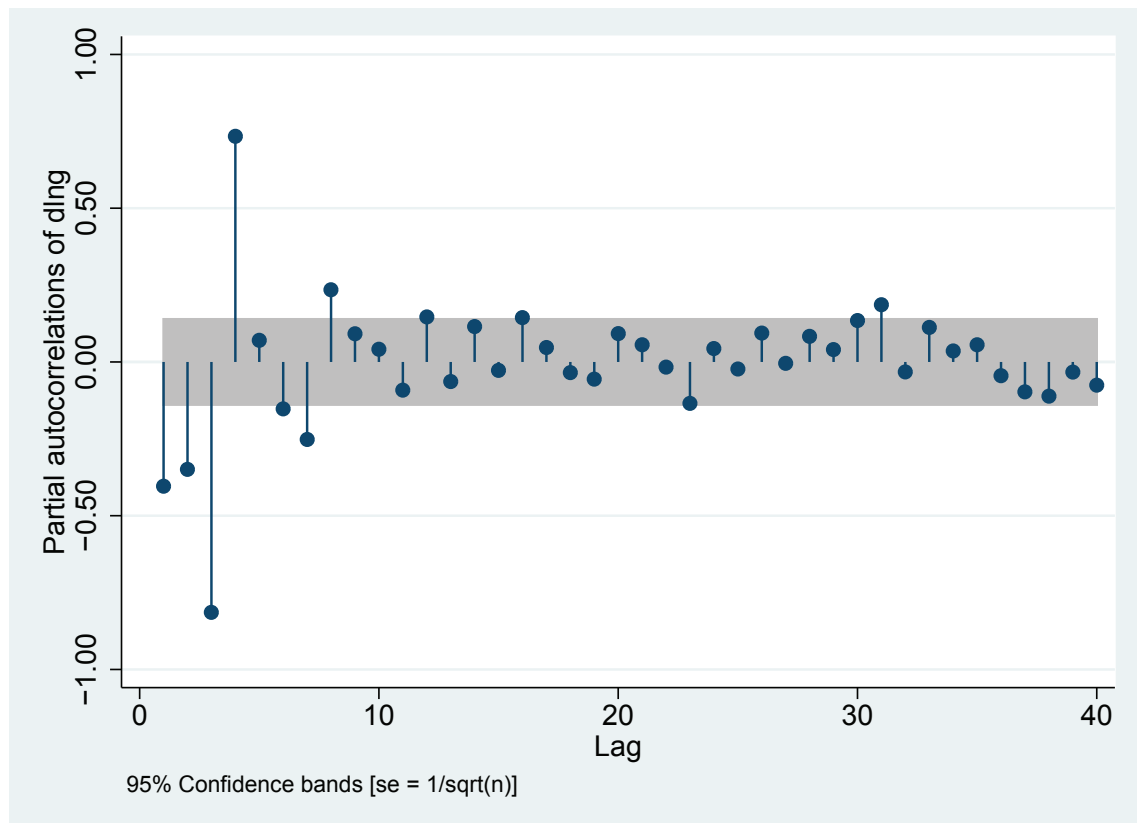


Abbildung 3.5.: Abbildung der empirischen partiellen Autokorrelationsfunktion der Wachstumsraten des Gesamteueraufkommens

3.3.2. Phase II: Schätzen und Testen

Modellschätzungen und -auswahl

Da viele unterschiedliche Modellspezifikationen die beobachteten ACF und PACF der Wachstumsraten des Gesamteueraufkommens erzeugt haben könnten, werden **zahlreiche verschiedene Modelle** geschätzt, um anschließend anhand geeigneter Kriterien eine Auswahl zu treffen. Die den folgenden Schätzungen zugrunde liegende stationäre Zeitreihe ist die bereits einmal differenzierte ($d = 1$) Zeitreihe des logarithmierten Gesamteueraufkommens der BRD, also näherungsweise die Wachstumsraten des Gesamteueraufkommens. Wie oben bereits beschrieben, ist keine saisonale Differenzierung nötig ($D = 0$), wengleich eine starke Saisonfigur mit Periodizität $s = 4$ (Quartalsdaten) vorliegt. Für die saisonalen Parameter P und Q

sind die Werte $P = 1, 2$ für den saisonalen AR-Teil und die Werte $Q = 0, 1, 2$ für den saisonalen MA-Teil denkbar. Für die nicht-saisonalen Parameter werden ebenfalls jeweils drei mögliche Werte, nämlich $p = 0, 1, 2$ und $q = 0, 1, 2$ zugelassen.

Die 54 möglichen Modellspezifikationen, die sich aus den Kombinationen dieser Parameter ergeben, werden mit der Statistik-Software Stata geschätzt. Da der Störterm $\{\varepsilon_t\}$ annahmegemäß identisch und unabhängig normalverteiltes weißes Rauschen darstellt⁷, liegt eine **Maximum-Likelihood-Schätzung (ML-Schätzung)** zugrunde. Stata führt diese anhand des **Kalman-Filters**⁸ durch.

Der ML-Schätzer $\hat{\kappa}$ des zu schätzenden Parametervektors

$$\kappa = (c, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \phi_{s1}, \dots, \phi_{sP}, \theta_{s1}, \dots, \theta_{sQ}, \sigma^2)$$

ist konsistent, asymptotisch effizient und asymptotisch normalverteilt (vgl. Greene (2003), S. 473). Einzelne Parameter können demnach anhand gängiger t -Tests, und mehrere Parameterrestriktionen mit Hilfe der üblichen Likelihood-Ratio, Lagrange-Multiplikator und Wald Tests getestet werden.⁹

Damit das Modell aber nicht überidentifiziert wird, erfolgt die Modellauswahl hauptsächlich anhand so genannter **Informationskriterien** (vgl. Enders (1995), S. 97). Diese sorgen für eine sparsame Parametrisierung, indem das Hinzunehmen weiterer Terme bzw. zu schätzender Koeffizienten bestraft wird. Üblicherweise werden das **Akaike Information Criterion (AIC)** nach Akaike (1974) und das **Bayesian Information Criterion (BIC)** nach Schwarz (1978) zur Modellauswahl herangezogen. Für beide Informationskriterien gilt, dass jeweils das Modell bevorzugt wird, das den kleinsten Wert des AIC bzw. BIC erreicht. Denn bei beiden Maßen geht der (logarithmierte) Wert der Likelihood am Maximum L (der möglichst groß sein

⁷Um diese Annahme zu testen, werden die Residuen der ausgewählten Modellvarianten im nächsten Unterkapitel anhand verschiedener statistischer Tests auf Normalverteilung getestet.

⁸Für eine genaue Beschreibung der ML-Schätzung anhand des Kalman-Filters siehe Kapitel 13 in Hamilton (1994), Kapitel 15.2 in Gourieroux und Monfort (1997) oder sehr ausführlich bei Harvey (1989).

⁹Für eine Beschreibung dieser Testverfahren bei ML-Schätzungen siehe beispielsweise Greene (2003), Kapitel 17.5.

sollte) mit negativem Vorzeichen ein, wogegen die Anzahl der zu schätzenden Parameter $n = 1 + p + q + P + Q$ im so genannten Strafterm mit positivem Vorzeichen entgegensteuert. Während sich das AIC als

$$\text{AIC} = -2 \ln L + 2n \quad (3.21)$$

berechnet (vgl. Akaike (1974), S. 716), bestraft das BIC die Anzahl der zu schätzenden Parameter n mit steigender Anzahl an Beobachtungen T stärker¹⁰:

$$\text{BIC} = -2 \ln L + \ln(T)n. \quad (3.22)$$

Im Optimalfall implizieren beide Kriterien dieselbe Modellauswahl. Ansonsten wird dem BIC tendenziell mehr Gewicht beigemessen, da es die Anzahl der unterschiedlichen Lags konsistent schätzt, wohingegen das AIC zur Überidentifikation neigt (vgl. beispielsweise Hayashi (2000), S. 397).

Um auch den möglichen Strukturbruch aufgrund der deutschen Wiedervereinigung zu berücksichtigen, werden alle 54 Modellvarianten nochmals zusätzlich mit der Dummyvariablen für die Wiedervereinigung geschätzt. Der Koeffizient dieser Dummyvariablen ist jedoch in den meisten Fällen auf dem 5%-Signifikanzniveau nicht signifikant von Null verschieden. Bei manchen der insgesamt 108 geschätzten Modellvarianten erweist sich außerdem die Maximierung als problematisch. Zum Teil werden sehr viele Iterationen benötigt und der Algorithmus muss dabei oft gewechselt werden, um das Maximum zu bestimmen. In Einzelfällen kann die Software das Maximum gar nicht finden. Bei manchen Schätzungen kann außerdem der Standardfehler für einzelne geschätzte Koeffizienten nicht ermittelt werden.

Anhand der Informationskriterien der 108 durchgeführten Schätzungen der Wachstumsraten des Steueraufkommens Y_t bestätigen sich für die **saisonalen Parameter** die aufgrund der ACF und PACF vermuteten Werte $P = 1$ **und** $Q = 1$ eindeutig. Für die nicht-saisonalen Parameter unterscheiden sich jedoch die Implikationen des

¹⁰Für $T \geq 8$ ist $\ln(T) \geq 2$.

AIC und BIC. Den kleinsten Wert des BIC erreicht das SARIMA(1,0,0)(1,0,1)₄-Modell ohne Wiedervereinigungsdummy

$$(1 - \phi_1 L)(1 - \phi_4 L^4)Y_t = c + (1 - \theta_4 L^4)\varepsilon_t, \quad (3.23)$$

und als das zweitbeste Modell gemäß BIC erweist sich das SARIMA(0,0,1)(1,0,1)₄-Modell (ebenfalls ohne Dummyvariable für die Wiedervereinigung)

$$(1 - \phi_4 L^4)Y_t = c + (1 - \theta_1 L)(1 - \theta_4 L^4)\varepsilon_t. \quad (3.24)$$

Die **nicht-saisonale Struktur** kann also ähnlich gut über einen AR(1)- oder einen MA(1)-Term modelliert werden. Die Hinzunahme der Dummyvariable für die Wiedervereinigung, um einen möglichen Strukturbruch zu modellieren, ergibt bei beiden Modellen keine Verbesserung, und der geschätzte Koeffizient dieser Dummyvariablen ist nicht signifikant von Null verschieden.

Die drittbeste Modellvariante gemäß BIC entspricht auch dem Modell, das den kleinsten Wert des AIC aufweist, nämlich das SARIMA(1,0,2)(1,0,1)₄-Modell mit Wiedervereinigungs-Dummy. Bei dieser Schätzung kann allerdings für den geschätzten nicht-saisonalen MA(1)-Koeffizienten aufgrund von Iterationsproblemen kein Standardfehler ermittelt werden. Außerdem ist der Wiedervereinigungs-Dummy auf dem 5%-Signifikanzniveau ebenfalls nicht signifikant von Null verschieden, so dass von dieser Modellvariante abgesehen wird. Auch das zweitbeste Modell gemäß AIC ist aufgrund von Schätzproblemem nicht aussagekräftig.

Als **Vergleichsmodell** zu den beiden bisherigen sparsam parametrisierten Modellen wird noch das SARIMA(2,0,1)(1,0,1)₄-Modell (wiederum ohne Dummyvariable für die Wiedervereinigung)

$$(1 - \phi_1 L)(1 - \phi_2 L)(1 - \phi_4 L^4)Y_t = c + (1 - \theta_1 L)(1 - \theta_4 L^4)\varepsilon_t \quad (3.25)$$

betrachtet, das gemäß AIC als drittbestes und gemäß BIC als viertbestes Modell eingeordnet wird.

Alle drei ausgewählten Modellschätzungen sind basierend auf dem jeweiligen Wald-Test insgesamt signifikant. Auch die geschätzten Koeffizienten sind jeweils einzeln – mit Ausnahme der Konstanten – mindestens auf dem 5%-Signifikanzniveau signifikant von Null verschieden. Außerdem ergibt sich bei allen drei Varianten ein hoher positiver Schätzwert für den saisonalen AR(1)-Koeffizienten $\hat{\phi}_4 \approx 0,98$, und ein negativer Schätzwert von knapp $\hat{\theta}_4 \approx -0,6$ für den saisonalen MA(1)-Koeffizienten. Die detaillierten Schätzergebnisse sind in den Tabellen A.1 bis A.3 im Anhang aufgeführt.

Überprüfung der Residuen auf weißes Rauschen

Auch wenn die ausgewählten Modelle gemäß der Informationskriterien zwar die besten der betrachteten Modellvarianten darstellen, so sollten trotzdem die jeweiligen Residuen der Modellschätzungen untersucht werden. Denn wenn alle Muster beziehungsweise Regelmäßigkeiten der Zeitreihe (der Wachstumsraten des Gesamtsteueraufkommens) bereits modelliert werden, dann sollten die Residuen nur noch weißes Rauschen beinhalten. Dementsprechend sollten sich die (empirischen) Autokorrelationen $\hat{\varphi}_k$ sowie die (empirischen) partiellen Autokorrelationen (der Zeitreihe der Residuen) nicht signifikant von Null unterscheiden (vgl. beispielsweise Enders (1995), S. 98).

Abbildung 3.6 zeigt beispielhaft die empirische ACF und PACF der Residuen des SARIMA(1,0,0)(1,0,1)₄-Modells. Die entsprechenden ACF und PACF der beiden anderen ausgewählten Modelle sind in den Abbildungen A.1 und A.2 im Anhang dargestellt. Für alle drei Modellvarianten gibt es nur wenige signifikante Ausschläge der ACF beziehungsweise der PACF.

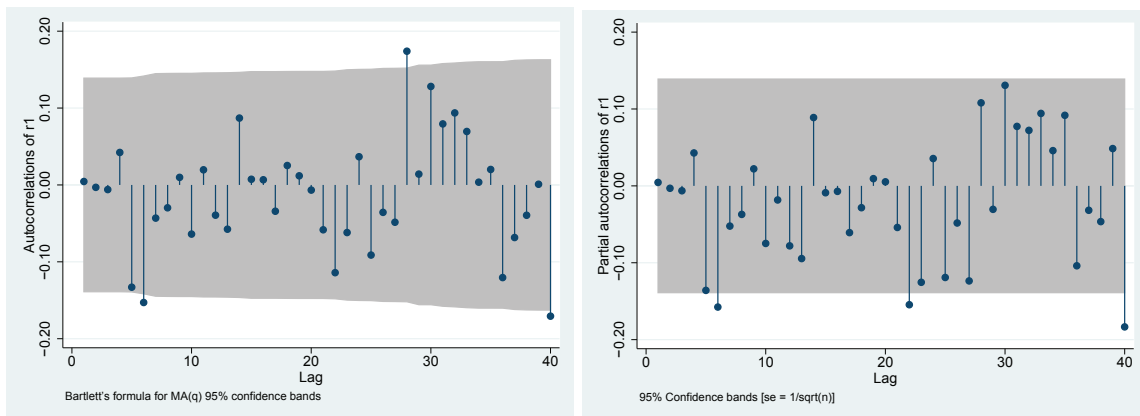


Abbildung 3.6.: Abbildung der empirischen ACF sowie der empirischen PACF des SARIMA(1,0,0)(1,0,1)₄-Modells

Des Weiteren bieten die so genannten **Portmanteau-Tests** die Möglichkeit, die Nullhypothese, dass die Residuen weißes Rauschen sind, statistisch zu testen. Der gängigste Portmanteau-Test ist der **Box-Pierce-Test**¹¹, der auf der Teststatistik

Q_{BP}

$$Q_{BP} = T \cdot \sum_{k=1}^l \hat{\varphi}_k^2 \quad (3.26)$$

basiert, wobei l hier für die Anzahl der geschätzten Autokorrelationen $\hat{\varphi}_k$ steht (vgl. beispielsweise Enders (1995), S. 87). Unter der Nullhypothese $H_0 : Q_{BP} = 0$ (Residuen sind weißes Rauschen) ist die Teststatistik Q_{BP} asymptotisch χ^2 -verteilt mit l Freiheitsgraden.

Eine für kleine Stichproben modifizierte Variante des Box-Pierce-Tests ist der **Ljung-Box-Test**¹² mit der Teststatistik Q_{LB}

$$Q_{LB} = T(T+2) \sum_{k=1}^l (T-k)^{-1} \hat{\varphi}_k^2 \quad (3.27)$$

und ebenfalls $Q_{LB}|H_0 \stackrel{a}{\sim} \chi^2(l)$ (vgl. z. B. Kirchgässner und Wolters (2007), S. 17).

¹¹Box und Pierce (1970).

¹²Ljung und Box (1978).

Für alle drei vorgeschlagenen Modellvarianten kann die Nullhypothese, dass die Residuen der Schätzung der Wachstumsraten des Gesamtsteueraufkommens weißes Rauschen sind, eindeutig nicht verworfen werden. Damit scheinen alle wesentlichen Regelmäßigkeiten beziehungsweise Muster der Zeitreihe in den betrachteten Modellen abgebildet zu sein.

Bevor die ermittelten Modellspezifikationen zur Prognose herangezogen werden, soll anhand der Residuen noch die **Annahme der Normalverteilung des Störterms** überprüft werden. Die Residuen werden daher den üblichen Tests auf Normalverteilung unterzogen. Allerdings kann sowohl für den Jarque-Bera-Test¹³ als auch für den Shapiro-Wilk-Test¹⁴ und für den weiterentwickelten Omnibustest gemäß Doornik und Hansen¹⁵ die Nullhypothese normalverteilter Residuen für alle ausgewählten Modellspezifikationen klar verworfen werden. Lediglich der Kolmogorov-Smirnov-Test¹⁶ bestätigt die getroffene Annahme. Da dieser Test jedoch dazu neigt, die Nullhypothese selten verwerfen zu können, ist eine genauere Untersuchung der Residuen notwendig (vgl. Hartung (2009), S. 522). Abbildung 3.7 zeigt daher beispielhaft die Residuen des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells im zeitlichen Verlauf.

Beim Betrachten der Zeitreihe der Residuen fallen zwei Ausreißer auf, nämlich der überdurchschnittlich große (positive) Wert des vierten Quartals 1969 sowie der besonders kleine (negative) Wert im ersten Quartal 1970. Vergleichbare Extremwerte sind auch für die beiden anderen Modellspezifikationen in diesen beiden Quartalen zu beobachten. Diese Ausreißer könnten an der Vielzahl der Steuerrechtsänderungen in den Jahren 1967 bis 1970 liegen.¹⁷ So trat im Dezember 1967 im Rahmen des zweiten Steueränderungsgesetzes das bis 1974 gültige Gesetz zur Einführung der dreiprozentigen Ergänzungsabgabe der Einkommensteuer bei höheren Einkommen in Kraft. Außerdem umfasste dieses zweite Steueränderungsgesetz den Abbau von

¹³Jarque und Bera (1980) und Jarque und Bera (1987).

¹⁴Shapiro und Wilk (1965).

¹⁵Doornik und Hansen (2008).

¹⁶Kolmogorov (1933), Smirnov (1939) sowie die Weiterentwicklung gemäß Conover (1999).

¹⁷Eine Übersicht aller Steuerrechtsänderungen ab dem Jahr 1964 findet sich auf der Internetseite des BMF.

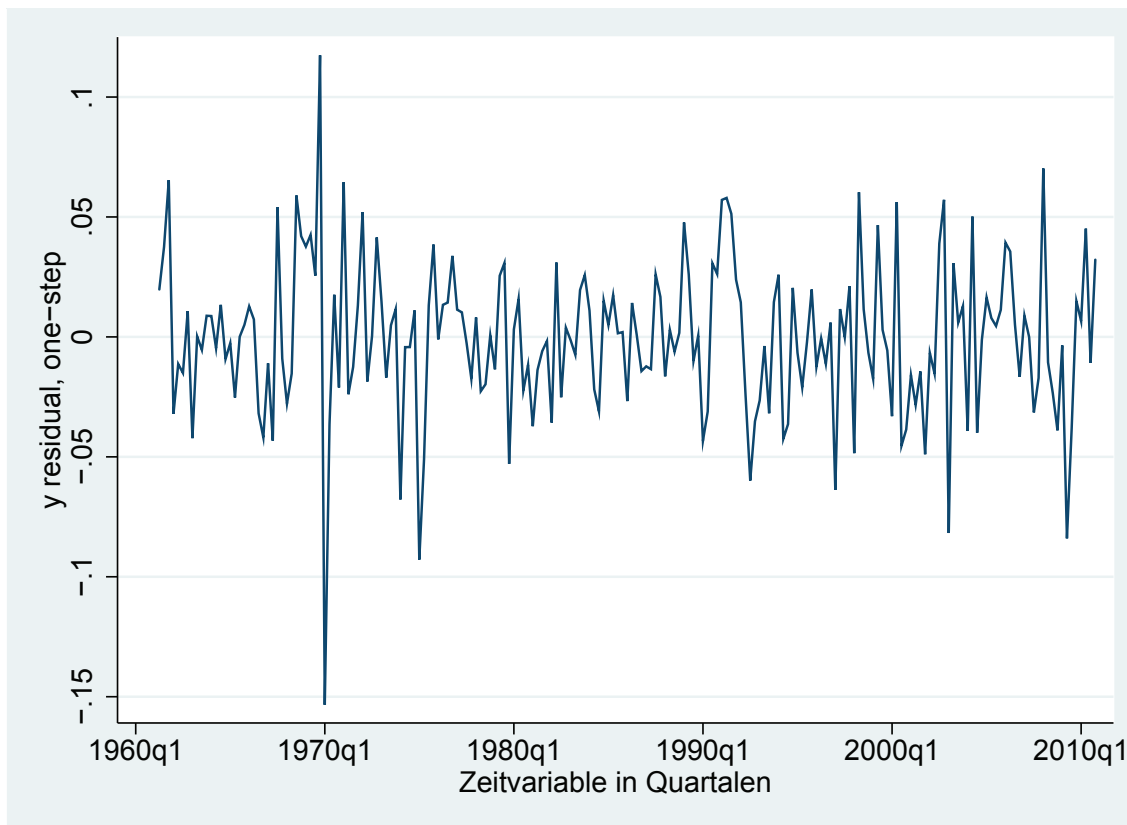


Abbildung 3.7.: Abbildung der Residuen des SARIMA(1,0,0)(1,0,1)₄-Modells

Steuervergünstigungen im Kreditgewerbe. Beide Maßnahmen könnten unter Umständen zum besonders großen (positiven) Wert des vierten Quartals 1969 beigetragen haben. Das Steueränderungsgesetz vom August 1969 umfasst dagegen neben anderen Maßnahmen die Gewährung von Investitionszulagen sowie die Verlängerung der Ermächtigung nach Paragraph 51 des Einkommensteuergesetzes. Zusätzlich trat Ende Dezember 1969 das Aufwertungsausgleichsgesetz in Kraft, das einen Einkommensausgleich für die Landwirtschaft im Rahmen der Umsatzsteuer vorsah. Diese drei Maßnahmen könnten zum besonders kleinen (negativen) Wert im ersten Quartal 1970 beigetragen haben. Eine genaue Zuordnung ist allerdings nicht möglich.

Sobald diese beiden Ausreißerwerte eliminiert werden, kann für alle oben aufgeführten Tests die Nullhypothese normalverteilter Residuen eindeutig nicht mehr verworfen werden. Daher werden die drei ausgewählten Modellspezifikationen als geeignet

betrachtet, um im nächsten Schritt des Box-Jenkins-Verfahrens zur Erstellung eigener Prognosen des Steueraufkommens verwendet werden zu können.

3.3.3. Phase III: Prognose als Anwendung

Für die Prognose der Variablen Y_{t+1} stehen die bis zum Zeitpunkt t beobachteten Werte

$$x_t = (y_t, y_{t-1}, \dots, y_1)' \quad (3.28)$$

zur Verfügung und der gesuchte Prognosewert wird mit $Y_{t+1}^*|t$ bezeichnet.

Welcher Prognosewert allerdings optimal ist, hängt von der zu Grunde liegenden **Verlustfunktion** ab. Auch wenn aus politischer Sicht Überschätzungen des Steueraufkommens aufgrund der dadurch auftretenden Steuerausfälle schwerwiegender sind als Unterschätzungen, so wird in der Literatur üblicherweise dennoch eine quadratische Verlustfunktion bevorzugt (vgl. z. B. Hamilton (1994), S. 72). $Y_{t+1}^*|t$ wird demnach so gewählt, dass der durchschnittliche quadrierte Prognosefehler – der **Mean Squared Error (MSE)**

$$\text{MSE}(Y_{t+1}^*|t) = \text{E}[\underbrace{(Y_{t+1} - Y_{t+1}^*|t)^2}_{\text{Prognosefehler}}] \quad (3.29)$$

minimiert wird. Dadurch gehen also Über- und Unterschätzungen gleichermaßen in die Berechnung ein und große Abweichungen werden stärker gewichtet als kleine.

Es kann gezeigt werden, dass die Prognose mit dem kleinsten MSE der bedingte Erwartungswert von Y_{t+1} gegeben x_t ist (Beweis siehe Hamilton (1994), S. 73). Werden nur lineare Funktionen von x_t als Prognosemodell zugelassen, gilt

$$Y_{t+1}^*|t = \text{E}[Y_{t+1}|x_t] = \alpha'x_t, \quad (3.30)$$

wobei α den gesuchten (optimalen) Parametervektor bezeichnet. Den kleinsten MSE hat dann die so genannte **lineare Projektion** $\hat{P}(Y_{t+1}|x_t) = Y_{t+1}^*|t = \alpha'x_t$ (Beweis

siehe Hamilton (1994), S.74), für die gilt, dass der Prognosefehler mit x_t unkorreliert ist

$$\underbrace{E[(Y_{t+1} - \alpha'x_t) x_t']}_{\text{Prognosefehler}} \stackrel{!}{=} \mathbf{0}. \quad (3.31)$$

Sofern auch eine Konstante in die Prognose eingeht, wird die lineare Projektion mit $\hat{E}(Y_{t+1}|x_t)$ (für $1 \in x_t$) bezeichnet. Wenn der stochastische Prozess $\{Y_t\}$ (**kovarianz**)-**stationär und ergodisch** ist, kann der gesuchte optimale Parametervektor α dieser linearen Projektion aufgrund des Theorems der Ergodizität durch den OLS-Schätzer einer Regression von Y_{t+1} auf x_t konsistent geschätzt werden (Beweis siehe Hamilton (1994), S. 74 ff.).

Da alle stationären (S)ARMA-Prozesse eine Wold-Repräsentation besitzen (vgl. z. B. Kirchgässner und Wolters (2007), S. 76), lässt sich jeder stationäre stochastische Prozess $\{Y_t\}$ als MA(∞)-Prozess darstellen:

$$Y_t = \mu + \Psi(L)\varepsilon_t = \mu + \varepsilon_t + \Psi_1\varepsilon_{t-1} + \Psi_2\varepsilon_{t-2} + \dots \quad (3.32)$$

Demnach ergibt sich für Y_{t+s} :

$$Y_{t+s} = \mu + \varepsilon_{t+s} + \Psi_1\varepsilon_{t+s-1} + \Psi_2\varepsilon_{t+s-2} + \dots + \Psi_s\varepsilon_t + \Psi_{s+1}\varepsilon_{t-1} + \dots \quad (3.33)$$

Da alle unbekanntem zukünftigen Schocks im Erwartungswert Null sind

$$E[\varepsilon_{t+s}|\varepsilon_t, \varepsilon_{t-1}, \dots] = 0; \quad \text{für } s \geq 0 \quad (3.34)$$

und unter der Annahme, dass alle bis zum Zeitpunkt t bereits realisierten Schocks $(\varepsilon_t, \varepsilon_{t-1}, \dots)$ sowie der Parametervektor $(\mu, \Psi_1, \Psi_2, \dots)$ bekannt sind, ergibt sich als optimale lineare Projektion bzw. **beste lineare unverzerrte Prognose (BLUP)** für Y_{t+s}

$$Y_{t+s}^*|t = \hat{E}[Y_{t+s}|\varepsilon_t, \varepsilon_{t-1}, \dots] = \mu + \Psi_s\varepsilon_t + \Psi_{s+1}\varepsilon_{t-1} + \Psi_{s+2}\varepsilon_{t-2} + \dots \quad (3.35)$$

(vgl. Hamilton (1994), S.77 oder auch Kirchgässner und Wolters (2007), S.78). Der dazugehörige Prognosefehler u_{t+s} berechnet sich demzufolge als

$$u_{t+s} = Y_{t+s} - \hat{E}[Y_{t+s} | \varepsilon_t, \varepsilon_{t-1}, \dots] = \varepsilon_{t+s} + \Psi_1 \varepsilon_{t+s-1} + \dots + \Psi_{s-1} \varepsilon_{t+1}, \quad (3.36)$$

so dass sich für den entsprechenden MSE

$$\text{MSE}(Y_{t+s}^* | t) = E[u_{t+s}^2] = (1 + \Psi_1^2 + \Psi_2^2 + \dots + \Psi_{s-1}^2) \sigma^2 \quad (3.37)$$

ergibt. Da ε_t annahmegemäß normalverteiltes weißes Rauschen darstellt, können somit für die Punktschätzer die dazugehörigen **Konfidenzintervalle** durch

$$\left[Y_{t+s}^* | t - k \sqrt{\text{MSE}(Y_{t+s}^* | t)}; Y_{t+s}^* | t + k \sqrt{\text{MSE}(Y_{t+s}^* | t)} \right] \quad (3.38)$$

angegeben werden, wobei k das entsprechende Quantil der Standardnormalverteilung bezeichnet. Da sich die Prognosefehler mit zunehmendem Prognosehorizont akkumulieren, wird auch der MSE der Prognose im Zeitablauf immer größer, was zu breiter werdenden Konfidenzintervallen führt.

Allerdings sind die bisherigen Annahmen in zwei Punkten problematisch. Einerseits sind nicht alle vergangenen Schocks bekannt. Stattdessen stehen nur die Beobachtungswerte (y_1, \dots, y_t) zur Verfügung. Unter der Annahme, dass alle Schocks ε_t für $t < 1$ vernachlässigt werden können, lassen sich die Innovationen $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_t)$ rekursiv aus den Beobachtungswerten (y_1, \dots, y_t) berechnen, sofern die Invertibilitätsbedingungen für den stationären (S)ARMA-Prozess erfüllt sind. Somit ergibt sich unter Verwendung der **rekursiven Berechnung**

$$\hat{\varepsilon}_t = Y_t - \hat{Y}_{t|t-1} \quad (3.39)$$

als **Annäherung für die optimale lineare Prognose** (vgl. beispielsweise auch Hamilton (1994), S.85)

$$\begin{aligned}\hat{E}(Y_{t+s}|y_1, \dots, y_t) &\cong \hat{E}(Y_{t+s}|y_1, \dots, y_t, \varepsilon_0 = \varepsilon_{-1} = \dots = 0) \\ &= \mu + \Psi_s \hat{\varepsilon}_t + \Psi_{s+1} \hat{\varepsilon}_{t-1} + \dots + \Psi_{t+s-1} \hat{\varepsilon}_1.\end{aligned}\tag{3.40}$$

Diese Annahme wird auch von Stata als Standardeinstellung für die Prognose von (S)ARIMA-Modellen verwendet (vgl. StataCorp (2009), S. 73). Die **exakte optimale lineare Projektion** bei endlichem Stichprobenumfang kann wiederum mit Hilfe des Kalman Filters oder der Faktorisierung über Dreiecksmatrizen¹⁸ berechnet werden. Da allerdings die **wahren Parameter** $(\mu, \Psi_1, \Psi_2, \dots)$ **unbekannt** sind, können für die Prognose lediglich die entsprechenden ML-Schätzer verwendet werden. Diese bringen aber ihrerseits zusätzliche Unsicherheit mit sich, die nicht durch die oben beschriebenen Konfidenzintervalle abgedeckt wird (vgl. beispielsweise Kirchgässner und Wolters (2007), S. 79). Als Faustregel empfiehlt Enders (1995), keiner (S)ARIMA-Prognose zu trauen, wenn weniger als 50 Beobachtungswerte zur Schätzung der Koeffizienten zur Verfügung stehen (s. Enders (1995), S. 105). Da für die hier durchgeführten Schätzungen und Prognosen der Wachstumsraten des Gesamtsteueraufkommens der BRD jeweils mindestens 144 Beobachtungswerte vorliegen, kann davon ausgegangen werden, dass die ML-Schätzer (aufgrund der Konsistenz-Eigenschaft) hinreichend genau sind.

Zunächst wird für jede Modellvariante eine Prognose unter Verwendung eines so genannten „**Holdout-Set**“ erstellt, das heißt es gehen nicht alle Beobachtungswerte bis zum aktuellen Rand in die Schätzung der vorgeschlagenen SARIMA-Modelle ein. Stattdessen werden für die erste Prognose (zum Prognosezeitpunkt $t =$ erstes Quartal 1997) die Werte des Gesamtsteueraufkommens der letzten verfügbaren 14 Jahre, also alle Werte ab dem ersten Quartal 1997, zur Modellschätzung weggelassen und ab diesem Zeitpunkt bis zum aktuellen Rand (viertes Quartal 2010) dynamisch prognostiziert. Dabei wird wie oben beschrieben die optimale lineare Prognose nä-

¹⁸Siehe z. B. Kapitel 4.4 in Hamilton (1994).

herungsweise durch die Annahme bestimmt, dass alle Schocks ε_t für $t < 1961$ (also vor dem ersten Beobachtungswert) vernachlässigt werden können.

Da es sich bei der betrachteten Zeitreihe des logarithmierten Gesamtsteueraufkommens um einen I(1)-Prozess handelt, werden die Modellschätzungen und Prognosen jeweils für die einmal differenzierte Zeitreihe, also die Wachstumsraten, erstellt. Die Prognosewerte für das logarithmierte Steueraufkommen im Niveau für den Zeitpunkt t erhält man dann dadurch, dass die prognostizierten Differenzen zum Niveau des (ab dem ersten Quartal 1997 ebenfalls prognostizierten) logarithmierten Steueraufkommens zum Zeitpunkt $t - 1$ addiert werden (vgl. z. B. Wooldridge (2009), S. 671). Anschließend werden die Prognosewerte für das logarithmierte Steueraufkommen exponenziert und das resultierende **prognostizierte Gesamtsteueraufkommen der BRD in Milliarden Euro** umgerechnet.

Analog zum beschriebenen Vorgehen wird hierauf der **Prognosezeitpunkt** immer um ein Quartal **erhöht**, das Holdout-Set also jeweils um ein Quartal verkleinert, so dass die dynamische Prognose stets ein Quartal später beginnt. Die dadurch entstehende Prognosematrix umfasst die Prognosen für $T = 40$ **Prognoseziele** für jeweils $h = 1$ bis $h = 16$ **Prognosehorizonte**. Dies entspricht also Prognosewerten des Gesamtsteueraufkommens (in Milliarden Euro) der BRD für 40 Quartale bzw. 10 Jahre, nämlich vom ersten Quartal 2001 bis zum vierten Quartal 2010 bei einer wachsenden Schätzdistanz von bis zu 16 Quartalen beziehungsweise 4 Jahren. Außerdem werden auf die gleiche Art und Weise Prognosematrizen für einen maximalen Prognosehorizont von $H = 12$ (Drei-Jahres-Prognosen) sowie für $H = 8$ (Zwei-Jahres-Prognosen) und für einen maximalen Prognosehorizont von $H = 4$ (Ein-Jahres-Prognosen) erstellt, um im folgenden Kapitel detaillierte Untersuchungen der Prognosequalität vornehmen zu können.

Allerdings treten für das dritte Vergleichsmodell, nämlich das höher parametrisierte SARIMA(2,0,1)(1,0,1)₄-Modell, beim Erstellen der Prognosematrix Schätzprobleme auf. Trotz sehr vieler Iterationen und Wechsel des Maximierungsalgorithmus kann

für etliche Prognosen kein Maximum gefunden werden, so dass diese Modellvariante im Weiteren verworfen wird. Im folgenden Kapitel wird nun die Prognosequalität der beiden anderen vorgeschlagenen SARIMA-Modelle gründlich untersucht und mit derjenigen der offiziellen AKS-Schätzungen verglichen.

4. Analyse der Prognosegüte

Prognostizieren die im vorhergehenden Kapitel ermittelten SARIMA-Modelle das Gesamtsteueraufkommen der BRD besser oder mindestens gleich gut wie der Arbeitskreis „Steuerschätzungen“? Um eine Antwort auf diese Frage zu bekommen, wird zunächst der gepoolte Modellansatz gemäß Clements et al. (2007) zur Analyse der Prognoserationalität, das heißt der Unverzerrtheit und Informationseffizienz, vorgestellt. Im Anschluss wird einerseits überprüft, ob die eigenen SARIMA-Prognosen dementsprechend als rational bezeichnet werden können. Außerdem werden die offiziellen AKS-Prognosen erstmalig gemäß dieses gepoolten Ansatzes auf Unverzerrtheit und Informationseffizienz untersucht. Abschließend wird zusätzlich die Prognosequalität der SARIMA-Schätzungen anhand deskriptiver Gütemaße sowohl bezüglich ihrer horizontalen als auch vertikalen Treffsicherheit mit derjenigen der AKS-Prognosen verglichen.

4.1. Modellansatz zur Untersuchung der Prognoserationalität

Eine Prognose gilt dann als **rational**, wenn sie keine systematische Verzerrung aufweist, und wenn alle Informationen, die zum Zeitpunkt der Prognose zugänglich sind, effizient genutzt werden (vgl. u. a. Becker und Buettner (2007), S. 3).

In Anlehnung an die Notation von Clements et al. (2007) bezeichnet im Folgenden A_t den tatsächlichen Wert des Steueraufkommens zum Zeitpunkt t , wogegen F_{th} für den Prognosewert des Steueraufkommens für das Prognoseziel t mit dem Prognosehorizont h steht. Die Prognose wurde dementsprechend zum Zeitpunkt $t - h$ erstellt, und der Prognosefehler e_{th} ergibt sich als Differenz des prognostizierten vom tatsächlichen Wert.

Die **traditionelle Analyse** der Prognoserationalität nach Mincer und Zarnowitz (1969) sowie Holden und Peel (1990) untersucht typischerweise nur Prognosen mit demselben Prognosehorizont (fixed horizon) für unterschiedliche Prognoseziele. Dabei wird zunächst die Regressionsgleichung

$$A_t = \alpha + \beta F_{th} + \varepsilon_t \quad (4.1)$$

separat für jeden Prognosehorizont h geschätzt, wobei für den Störterm ε_{th} ein MA(h-1)-Prozess zugelassen wird. Um auf Unverzerrtheit der Prognosen zu testen, wird gemäß Mincer und Zarnowitz (1969) die gemeinsame Nullhypothese $\alpha = 0$ und $\beta = 1$ getestet (vgl. Mincer und Zarnowitz (1969), S. 10). Holden und Peel (1990) wenden allerdings ein, dass diese Nullhypothese zwar hinreichend aber nicht nötig ist, um auf Unverzerrtheit zu testen. Denn die betrachtete Prognose F_{th} ist in Gleichung (4.1) auch dann unverzerrt, wenn $\alpha = (1 - \beta)F_{th}$ gilt. Aus diesem Grund schlagen Holden und Peel (1990) vor, stattdessen die Regressionsgleichung

$$e_{th} = A_t - F_{th} = \tau + \nu_t \quad (4.2)$$

zu schätzen, um die Nullhypothese der Unverzerrtheit $\tau = 0$ zu testen.

Dieser traditionelle Ansatz des Tests auf Unverzerrtheit findet beispielsweise bei Osterloh (2008) Anwendung, der damit die deutschen Konjunkturprognosen untersucht. Auch Becker und Buettner (2007) testen die Prognosen des Arbeitskreises „Steuerschätzungen“ separat für verschiedene Prognosehorizonte. Clements et al. (2007) stellen hingegen einen Ansatz vor, der die Prognosen für unterschiedliche

Prognoseziele über alle Prognosehorizonte zusammenführt und gemeinsam untersucht (**pooled approach**). Dieser Ansatz baut einerseits auf Keane und Runkle (1990), Davies und Lahiri (1995) sowie Davies und Lahiri (1999) auf, und erweitert andererseits Nordhaus (1987), Clements (1995) sowie Clements (1997).

Dieses „gepoolte“ Vorgehen hat mehrere **Vorteile**. Zunächst ist es intuitiv sinnvoll, dass rationale Prognosen für jeden untersuchten Prognosehorizont unverzerrt sein sollten. Denn stellt man sich die Frage der Unverzerrtheit separat für jede Schätzdistanz, sind die Ergebnisse möglicherweise schwer einzuordnen. Findet man beispielsweise heraus, dass die Prognosen mit einem Prognosehorizont von $h = j$ verzerrt, diejenigen mit einem Prognosehorizont von $h = j + 1$ unverzerrt sind, während die Prognosen mit $h = j + 2$ wieder eine systematische Verzerrung aufweisen, stellt sich eine logische Schlussfolgerung daraus doch als sehr schwierig dar (vgl. auch Clements et al. (2007), S. 122).

Darüber hinaus ergibt sich ein Unterschied für die inhaltliche Aussagekraft der Ergebnisse, je nachdem, ob man ein bestimmtes Signifikanzniveau für den Test der Unverzerrtheit der Prognosen für alle Prognosehorizonte zugrunde legt, oder ob man dasselbe Signifikanzniveau für die separaten Tests der Unverzerrtheit der Prognosen der verschiedenen Prognosehorizonte unterstellt. Denn testet man mehrere Prognosereihen für unterschiedliche Prognosehorizonte einzeln, dann ist die Wahrscheinlichkeit, die Nullhypothese für mindestens eine Prognosereihe irrtümlich zu verwerfen, größer als das jeweils unterstellte Signifikanzniveau des Einzeltests. Außerdem ist dieser t -Test, wenn er für alle Prognosereihen zusammen verwendet wird, aufgrund der größeren Anzahl an Beobachtungen trennschärfer als die einzelnen t -Tests für jeden Prognosehorizont.

4.1.1. Test auf Unverzerrtheit

Zunächst wird beschrieben, wie die Prognosereihen über alle Prognosehorizonte gepoolt werden. Darauf aufbauend wird das Modell gemäß Clements et al. (2007) vorgestellt, mit dem die Nullhypothese der Unverzerrtheit der Prognosen für alle Prognosehorizonte zusammen getestet wird, bevor anschließend ihre Informations-effizienz untersucht wird.

Beim „Poolen“ werden die Prognosewerte F_{th} zunächst zum TH -dimensionalen Spaltenvektor F

$$F = (F_{1H}, \dots, F_{11}, F_{2H}, \dots, F_{21}, \dots, F_{TH}, \dots, F_{T1})' \quad (4.3)$$

zusammengesetzt. Die tatsächlichen Werte A_t bilden den T -dimensionalen Spaltenvektor $A^* = (A_1, A_2, \dots, A_T)'$, und i_H ist ein H -dimensionaler Spaltenvektor aus Einsen. Außerdem wird der TH -dimensionale Spaltenvektor A durch das Kronecker-Produkt¹

$$\begin{aligned} A &= A^* \otimes i_H \\ &= (A_1, A_2, \dots, A_T)' \otimes (1, \dots, 1)' \\ &= (A_1, \dots, A_1, A_2, \dots, A_2, \dots, A_T, \dots, A_T)' \end{aligned} \quad (4.4)$$

definiert, so dass sich der TH -dimensionale Spaltenvektor e der Prognosefehler als

$$\begin{aligned} e &= A - F \\ &= (A_1, \dots, A_1, \dots, A_T, \dots, A_T)' - (F_{1H}, \dots, F_{11}, \dots, F_{TH}, \dots, F_{T1})' \\ &= (A_1 - F_{1H}, \dots, A_1 - F_{11}, \dots, A_T - F_{TH}, \dots, A_T - F_{T1})' \\ &= (e_{1H}, \dots, e_{11}, \dots, e_{TH}, \dots, e_{T1})' \end{aligned} \quad (4.5)$$

ergibt.

Wenn man über die verschiedenen Prognosehorizonte poolt, spielt einerseits die Annahme bezüglich des Störterms und andererseits die Modellierung der Korrelationsstruktur der Prognosefehler für unterschiedliche Prognoseziele und -horizonte eine entscheidende Rolle.

¹Das Kronecker-Produkt $C = A \otimes B$ berechnet sich, indem man jedes Element der Matrix A mit der Matrix B multipliziert, also $C = a_{ij}B$ (vgl. z. B. Greene (2003), S. 824 f.).

Ausgangspunkt im Modell von Clements et al. (2007) ist die **Zerlegung des Prognosefehlers** gemäß Davies und Lahiri (1995). In deren Modell erstellen allerdings mehrere Individuen verschiedene Prognosen für dasselbe Prognoseziel und denselben Prognosehorizont. Diese Arbeit betrachtet hingegen lediglich Prognosen eines einzigen Individuums bzw. einer einzigen Institution (AKS) für $t = 1, \dots, T$ Prognoseziele und $h = 1, \dots, H$ Prognosehorizonte. Der Prognosefehler e_{th} wird demzufolge für ein einziges Individuum bzw. eine einzige Institution zerlegt in

$$\begin{aligned} e_{th} = A_t - F_{th} &= \alpha + \lambda_{th} + \varepsilon_{th} \\ &= \alpha + v_{th}. \end{aligned} \quad (4.6)$$

Dabei bezeichnet α die zu testende Verzerrung und ε_{th} den **idiosynkratischen Störterm**, für den $E[\varepsilon_{th}] = 0 \forall t, h$ gilt. Indes steht λ_{th} für die **aggregierten makro-ökonomischen Schocks**, die sich als Summe aller makro-ökonomischer Schocks u_{th} mit $E[u_{th}] = 0 \forall t, h$ ergeben, die zwischen $t - h$ und t auftraten, also

$$\lambda_{th} = \sum_{j=1}^h u_{tj} \quad (4.7)$$

(vgl. Davies und Lahiri (1995), S. 208 bzw. Clements et al. (2007), S. 123). Der Störterm ε_{th} beschreibt den spezifischen Fehler bei der Prognose zum Zeitpunkt $t - h$ mit dem Prognoseziel t und Prognosehorizont h . Daher wird die Varianz des idiosynkratischen Störterms $Var[\varepsilon_{th}] = E[\varepsilon_{th}^2] = \sigma_\varepsilon^2$ als konstant für alle Prognosehorizonte h (und für alle t) angenommen. Im Gegensatz dazu nimmt die Varianz des aggregierten makro-ökonomischen Schocks $Var[\lambda_{th}]$ mit zunehmendem Prognosehorizont h zu, da λ_{th} alle Schocks zwischen $t - h$ und t kumuliert (vgl. Clements et al. (2007), S. 123 f.).

In Matrixschreibweise kann die Testgleichung (4.6) als

$$\mathbf{Variante I:} \quad e = i_{TH} \alpha + v \quad (4.8)$$

geschrieben werden, wobei i_{TH} ein TH -dimensionaler Spaltenvektor aus Einsen ist, und v den TH -dimensionalen Spaltenvektor des zusammengefassten Störterms darstellt (vgl. Clements et al. (2007), S. 124). Die zu testende **Verzerrung** α unterliegt in Gleichung (4.8) also der Bedingung, dass sie **für alle Prognosehorizonte gleich groß** sein muss.

Diese Annahme kann dadurch gelockert werden, dass eine **gesonderte Verzerrung** α_h **für jeden Prognosehorizont** h zugelassen wird. Diese horizont-spezifischen Verzerrungen bilden den H -dimensionalen Spaltenvektor $\alpha^H = (\alpha_H, \dots, \alpha_1)'$. Anstelle von Gleichung (4.8) erhält man dann die Testgleichung

$$\text{Variante II:} \quad e = (i_T \otimes I_H) \alpha^H + v, \quad (4.9)$$

wobei I_H die H -dimensionale Einheitsmatrix definiert, so dass die Regressormatrix $X = i_T \otimes I_H$ in (4.9) die erforderliche Dimension ($TH \times H$) besitzt (vgl. Clements et al. (2007), S. 124).

In beiden Varianten wird die Nullhypothese getestet, dass die Prognosen für alle Prognosehorizonte unverzerrt sind. Bei Variante I wird also die Nullhypothese $\alpha = 0$ durch einen t -Test getestet, während für den Test der Nullhypothese $\alpha^H = 0$ in Variante II ein F -Test durchzuführen ist. Auch wenn sich dadurch für beide Varianten dasselbe Modell unter H_0 ergibt, so **unterscheiden sich die Alternativhypothesen** doch maßgeblich. Denn wenn die Verzerrungen tatsächlich über die Prognosehorizonte systematisch variieren, dann könnte die einschränkende Annahme einer für alle Prognosehorizonte gleich großen Verzerrung (Variante I) irrtümlicherweise dazu führen, dass die Nullhypothese nicht abgelehnt werden kann (vgl. auch Clements et al. (2007), S. 124). Im Folgenden werden beide Varianten berücksichtigt, um die Nullhypothese der Unverzerrtheit zu testen.

Außerdem gilt es zu beachten, dass der Test in Variante II **nicht** dem traditionellen Vorgehen entspricht, bei dem für jeden Prognosehorizont eine eigene Schätzung

vorgenommen wird, und dann die horizont-spezifischen Verzerrungen jeweils anhand eines t -Tests auf Signifikanz getestet werden.

Des Weiteren werden im Folgenden nochmals zwei Modellalternativen unterschieden je nachdem, welche Annahme bezüglich des idiosynkratischen Störterms ε_{th} getroffen wird. Im ursprünglichen Modell von Davies und Lahiri (1995) charakterisiert ε_{ith} den „individuen-spezifischen“ idiosynkratischen Fehler, also den Fehler, den ein bestimmtes Individuum i beispielsweise durch ineffiziente Nutzung von (privaten) Informationen oder subjektive Eingriffe beim Prognostizieren macht (vgl. Davies und Lahiri (1995), S. 206 ff.). Da in dieser Arbeit jedoch die zu untersuchenden Prognosen nur von einem einzigen Individuum bzw. einer einzigen Institution erstellt werden, erscheint zunächst die Zerlegung des Störterms v_{th} in den idiosynkratischen Fehler ε_{th} und den aggregierten makro-ökonomischen Schock λ_{th} fraglich.

Daraus ergeben sich **zwei mögliche Vorgehensweisen**. Entweder wird der idiosynkratische Störterm ε_{th} nicht mehr individuen-spezifisch interpretiert (Annahme A), oder komplett weggelassen (Annahme B).

Die Modellierung unter **Annahme A** interpretiert den Störterm ε_{th} als idiosynkratischen Fehler, der zum Zeitpunkt t für den Prognosehorizont h beispielsweise durch ineffiziente Nutzung von Informationen, Messfehler oder Fehlspezifikationen im Prognosemodell hervorgerufen wird (vgl. Clements et al. (2007), S. 124). Somit würde ε_{th} als „**prognose-spezifischer**“ **idiosynkratischer Fehler** mit konstanter Varianz für alle h betrachtet werden.

Wenn es aber beispielsweise nur ein prognostizierendes Individuum gibt, das außerdem über private Informationen verfügt, die im Sinne rationaler Erwartungen optimal genutzt werden, würde ε_{th} durch $\eta_{th} = \sum_{j=1}^h \varepsilon_{tj}$ ersetzt², so dass die Varianz $\text{Var}(\eta_{th})$ mit wachsendem Prognosehorizont ebenfalls zunimmt. Doch dann könnten

²Für eine ausführliche Herleitung siehe Davies und Lahiri (1995), S.219, Fussnote 13 und Clements et al. (2007), S. 124.

die beiden inhaltlich zwar verschiedenen Störgrößen ε_{th} und λ_{th} formal nicht mehr getrennt werden, da demzufolge

$$\begin{aligned} e_{th} = A_t - F_{th} &= \alpha + (\lambda_{th} + \eta_{th}) \\ &= \alpha + v_{th} \end{aligned} \quad (4.10)$$

mit

$$v_{th} = \sum_{j=1}^h (u_{tj} + \varepsilon_{tj}) = \sum_{j=1}^h u_{tj}^* \quad (4.11)$$

gilt (vgl. Clements et al. (2007), S. 124). In diesem Fall wäre die Modellierung **ohne idiosynkratischen Störterm**, also unter **Annahme B**, zutreffend.

Die Nullhypothese der Unverzerrtheit soll im Folgenden unter beiden Annahmen getestet werden, um die Robustheit des Testergebnisses bezüglich der Annahme der Zusammensetzung des Störterms zu überprüfen.

Sowohl unter Annahme A als auch unter Annahme B gilt zwar $E[v] = 0$. Die Varianz-Kovarianz-Matrix $\Sigma = E[vv']$ des zusammengesetzten Störterms v ist allerdings nicht proportional zur Einheitsmatrix und hängt davon ab, welche Annahme bezüglich des Störterms getroffen wird. Wird ein idiosynkratischer Störterm mit konstanter Varianz σ_ε^2 zugelassen (Annahme A), für den außerdem $E[\varepsilon_{th}\varepsilon_{t+i,h+j}] = 0$ $\forall i, j \neq 0$ gilt, erhält man

$$\Sigma = \sigma_\varepsilon^2 I_{TH} + \Psi, \quad (4.12)$$

wobei Ψ die Korrelationsstruktur der Prognosen bzw. Prognosefehler beschreibt, die dadurch hervorgerufen wird, dass sie von denselben makro-ökonomischen Schocks beeinflusst werden (vgl. Clements et al. (2007), S. 125). Unter Annahme B ist $\sigma_\varepsilon^2 = 0$, so dass $\Sigma = \Psi$ gilt.

Abbildung 4.1 veranschaulicht den Zusammenhang der Prognosen und ihrer (aggregierten makro-ökonomischen) Störterme anhand eines Zeitstrahls. Exemplarisch wird der Zeitraum von vier Jahren vom ersten Quartal des Jahres 2008 ($2008 Q_1$) bis zum ersten Quartal des Jahres 2012 ($2012 Q_1$) in Quartalsabständen abgebil-

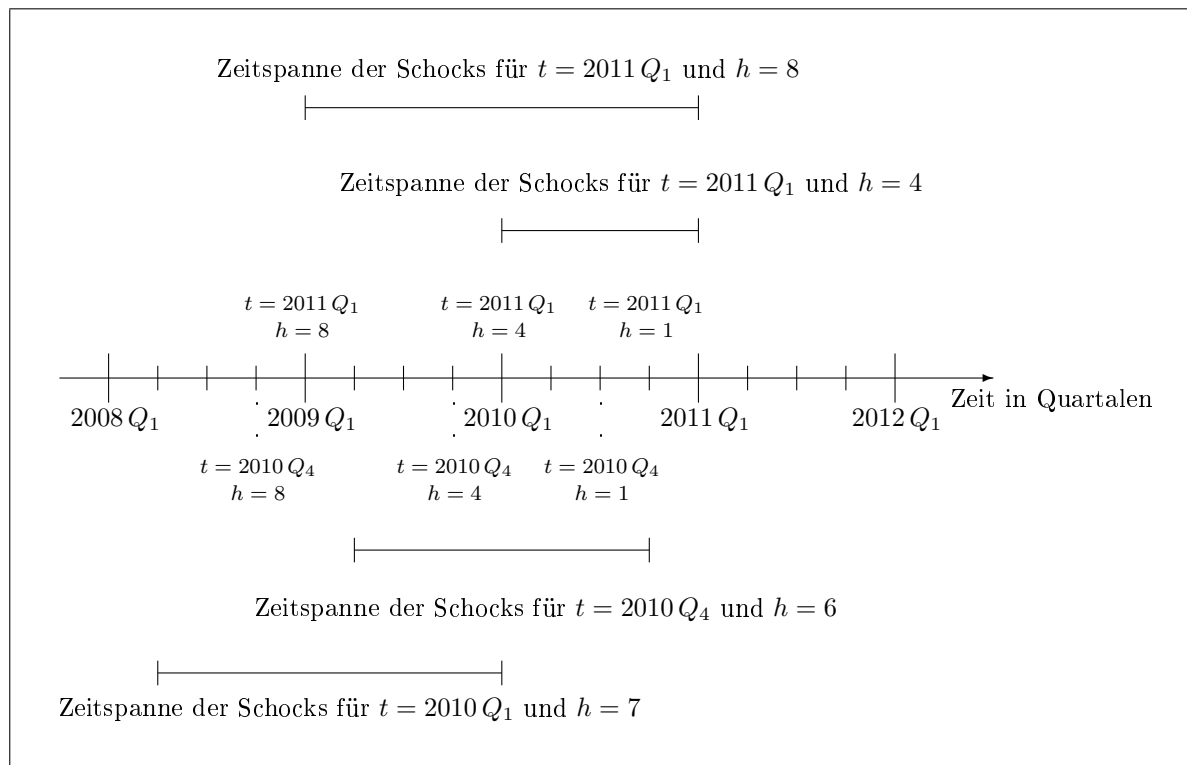


Abbildung 4.1.: Schematische Darstellung der Struktur der Prognosefehler (eigene Darstellung)

det. Die beiden oberen Intervalle in Abbildung 4.1 umfassen jeweils die Zeitspanne, in der Schocks auftreten können, die die Prognosen mit demselben Prognoseziel $t = 2011 Q_1$, aber unterschiedlichen Prognosehorizonten $h = 4$ beziehungsweise $h = 8$, beeinflussen. Der Teilbereich, den beide Zeitspannen gemeinsam haben, enthält dementsprechend die Ursache **serieller Korrelation** zwischen den Prognosen bzw. Prognosefehlern **verschiedener Prognosehorizonte** (bei gleichem Prognoseziel).

Das Intervall direkt unter dem Zeitstrahl umfasst den Zeitraum, in dem Schocks vorkommen können, die sich auf die Prognose mit dem Prognoseziel $t = 2010 Q_4$ und beispielsweise dem Prognosehorizont $h = 6$ auswirken. Der Teilbereich, den diese Zeitspanne und eine der beiden oberen Zeitspannen gemeinsam haben, enthält folglich die Ursache **serieller Korrelation** zwischen Prognosen bzw. Prognosefehlern

aufeinanderfolgender Prognoseziele (in Abhängigkeit der betrachteten Prognosehorizonte).

Die genaue Form von Ψ hängt maßgeblich vom minimalen und maximalen Prognosehorizont ab und soll im Folgenden exemplarisch für Prognosen mit $h = 1, \dots, 8$ hergeleitet werden. Dabei ist zunächst zu beachten, dass von homoskedastischen makro-ökonomischen Schocks u_{th} ausgegangen wird, so dass Ψ nur noch von einem einzigen unbekanntem Parameter, nämlich der Varianz $\sigma_u^2 = Var[u_{th}] = E[u_{th}^2]$ abhängt (vgl. Clements et al. (2007), S. 125). Unter Annahme B steht σ_u^2 aus Vereinfachungsgründen für die Varianz $Var[u_{th}^*] = E[u_{th}^{*2}]$.

Somit nimmt die $(TH \times TH)$ -dimensionale Matrix Ψ für $T = 40$ Prognoseziele und für exemplarisch bis zu $H = 8$ Prognosehorizonte (Zwei-Jahres-Prognosen) die Form

$$\Psi = \sigma_u^2 \Theta = \sigma_u^2 \begin{bmatrix} A & B & C & D & E & F & G & H & 0 & \dots & & & 0 \\ B' & A & B & C & D & E & F & G & H & \ddots & & & \\ C' & B' & A & B & C & D & E & F & G & \ddots & & & \\ D' & C' & B' & A & B & C & D & E & F & \ddots & & & \\ E' & D' & C' & B' & A & B & C & D & E & \ddots & & & \\ F' & E' & D' & C' & B' & A & B & C & D & \ddots & & & \\ G' & F' & E' & D' & C' & B' & A & B & C & \ddots & & & \\ H' & G' & F' & E' & D' & C' & B' & A & B & \ddots & & & \\ 0 & H' & G' & F' & E' & D' & C' & B' & A & \ddots & & & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & & \\ & & & & & & & & & & & & C \\ & & & & & & & & & & & & A & B \\ 0 & & & & & & & & & & & & C' & B' & A \end{bmatrix} \quad (4.13)$$

an, wobei Θ die folgenden $(H \times H)$ -dimensionalen Komponentenmatrizen enthält

$$A = \begin{bmatrix} 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 7 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 6 & 6 & 6 & 5 & 4 & 3 & 2 & 1 \\ 5 & 5 & 5 & 5 & 4 & 3 & 2 & 1 \\ 4 & 4 & 4 & 4 & 4 & 3 & 2 & 1 \\ 3 & 3 & 3 & 3 & 3 & 3 & 2 & 1 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 \\ 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 \\ 6 & 6 & 5 & 4 & 3 & 2 & 1 & 0 \\ 5 & 5 & 5 & 4 & 3 & 2 & 1 & 0 \\ 4 & 4 & 4 & 4 & 3 & 2 & 1 & 0 \\ 3 & 3 & 3 & 3 & 3 & 2 & 1 & 0 \\ 2 & 2 & 2 & 2 & 2 & 2 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 6 & 5 & 4 & 3 & 2 & 1 & 0 & 0 \\ 6 & 5 & 4 & 3 & 2 & 1 & 0 & 0 \\ 6 & 5 & 4 & 3 & 2 & 1 & 0 & 0 \\ 5 & 5 & 4 & 3 & 2 & 1 & 0 & 0 \\ 4 & 4 & 4 & 3 & 2 & 1 & 0 & 0 \\ 3 & 3 & 3 & 3 & 2 & 1 & 0 & 0 \\ 2 & 2 & 2 & 2 & 2 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \end{bmatrix}$$

$$D = \begin{bmatrix} 5 & 4 & 3 & 2 & 1 & 0 & 0 & 0 \\ 5 & 4 & 3 & 2 & 1 & 0 & 0 & 0 \\ 5 & 4 & 3 & 2 & 1 & 0 & 0 & 0 \\ 5 & 4 & 3 & 2 & 1 & 0 & 0 & 0 \\ 4 & 4 & 3 & 2 & 1 & 0 & 0 & 0 \\ 3 & 3 & 3 & 2 & 1 & 0 & 0 & 0 \\ 2 & 2 & 2 & 2 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}, \quad E = \begin{bmatrix} 4 & 3 & 2 & 1 & 0 & 0 & 0 & 0 \\ 4 & 3 & 2 & 1 & 0 & 0 & 0 & 0 \\ 4 & 3 & 2 & 1 & 0 & 0 & 0 & 0 \\ 4 & 3 & 2 & 1 & 0 & 0 & 0 & 0 \\ 4 & 3 & 2 & 1 & 0 & 0 & 0 & 0 \\ 3 & 3 & 2 & 1 & 0 & 0 & 0 & 0 \\ 2 & 2 & 2 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad F = \begin{bmatrix} 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$G = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{und} \quad H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Die Zusammensetzung von Θ lässt sich am einfachsten anhand der drei Beispiele aus Abbildung 4.1 nachvollziehen. Betrachtet man als Erstes die beiden **Prognosen mit demselben Prognoseziel**, also die Prognosen für das erste Quartal 2011, mit den unterschiedlichen Prognosehorizonten $h = 8$ bzw. $h = 4$, können in vier Quartalen für beide Prognosen dieselben makro-ökonomischen Schocks auftreten (2010 Q_1 bis 2010 Q_4). Dementsprechend steht sowohl in der ersten Zeile ($h=8$) und fünften Spalte ($h=4$) der symmetrischen Matrix A (für Prognosen mit demselben Prognoseziel t) der Wert 4, als auch in der fünften Zeile und ersten Spalte.

Für **Prognosen mit aufeinanderfolgenden Prognosezielen**, ist die Komponentenmatrix B relevant. Als Beispiel wird die Prognose für $t = 2010 Q_4$ und $h = 6$ mit derjenigen für das darauffolgende Quartal $t = 2011 Q_1$ und $h = 4$ verglichen. Gemäß Abbildung 4.1 gibt es drei überlappende Quartale (2010 Q_1 bis einschließlich 2010 Q_3), so dass in der Komponentenmatrix B (für Prognosen mit aufeinanderfolgenden Prognosezielen t und $t + 1$) in der dritten Zeile ($h = 6$) und fünften Spalte ($h = 4$) die 3 steht.

Abschließend werden noch die Prognose für $t = 2010 Q_1$ und $h = 7$ sowie diejenige für $t = 2011 Q_1$ und $h = 8$ betrachtet. Zwischen diesen beiden Prognosezielen liegen vier Quartale, so dass die Komponentenmatrix E entscheidend ist. In der zweiten Zeile ($h = 7$) und ersten Spalte ($h = 8$) kann man ablesen, dass in 4 Quartalen dieselben makro-ökonomischen Schocks auftreten können. In Abbildung 4.1 überlappen die beiden Zeitspannen dementsprechend vom ersten Quartal 2009 bis zum Ende des vierten Quartals 2009.

Da Ψ hier für einen maximalen Prognosehorizont von $H = 8$ dargestellt wurde, sind alle Matrizen in Θ für Prognosen mit mehr als sieben Quartalen Abstand zwischen den Prognosezielen (8×8)-dimensionale Nullmatrizen. Untersucht man die Prognosen der mittelfristigen Finanzplanung, also Prognosen mit einem maximalen Prognosehorizont von drei oder vier Jahren bzw. $H = 12$ oder $H = 16$ Quartalen, muss die Zusammensetzung von Θ jeweils entsprechend angepasst und erweitert werden. Dementsprechend ist eine Anpassung der Zusammensetzung von Θ auch für die kurzfristigen Ein-Jahres-Prognosen mit maximalem Prognosehorizont von $H = 4$ erforderlich.

Um die Nullhypothese der Unverzerrtheit zu testen, wird für **Variante I** die Verzerrung α in Gleichung (4.8)

$$e = i_{TH} \alpha + v$$

durch den Kleinstquadrateschätzer $\hat{\alpha} = (X'X)^{-1}X'e$ mit dem konsistenten Huber-White-Sandwich-Schätzer der Varianz-Kovarianz-Matrix

$$Cov[\hat{\alpha}] = (X'X)^{-1}X'\Sigma X(X'X)^{-1} \quad (4.14)$$

geschätzt, wobei $X = i_{TH}$ gilt.

Für **Variante II** werden die horizont-spezifischen Verzerrungen α^H in Gleichung (4.9)

$$e = (i_T \otimes I_H) \alpha^H + v$$

analog durch $\hat{\alpha}^H = (X'X)^{-1}X'e$ mit der Varianz-Kovarianz-Matrix

$$Cov[\hat{\alpha}^H] = (X'X)^{-1}X'\Sigma X(X'X)^{-1} \quad (4.15)$$

geschätzt, wobei die Regressormatrix $X = i_T \otimes I_H$ verwendet wird (vgl. auch Clements et al. (2007), S. 125).

Um für $\hat{\alpha}$ bzw. $\hat{\alpha}^H$ **konsistente Standardfehler** zu erhalten, muss Σ in (4.14) bzw. (4.15) durch den Schätzer

$$\hat{\Sigma} = \hat{\sigma}_\varepsilon^2 I_{TH} + \hat{\sigma}_u^2 \Theta \quad (4.16)$$

ersetzt werden. Somit müssen also zunächst Schätzwerte für σ_ε^2 und σ_u^2 ermittelt werden. Dazu berechnet man die entsprechenden Residuen als $\hat{v} = e - i_{TH} \hat{\alpha}$ für Variante I beziehungsweise als $\hat{v} = e - (i_T \otimes I_H) \hat{\alpha}^H$ für Variante II. Die Varianz des zusammengesetzten Störterms v_{th} ergibt sich unter den obigen Annahmen als

$$Var[v_{th}] = E[v_{th}^2] = E[(\varepsilon_{th} + \lambda_{th})^2] = \sigma_\varepsilon^2 + h \sigma_u^2, \quad (4.17)$$

so dass unter **Annahme A** die geschätzten Koeffizienten $\hat{\rho}_0$ und $\hat{\rho}_1$ der Regressionsgleichung

$$\hat{v} \odot \hat{v} = \rho_0 i_{TH} + \rho_1 \tau + \zeta \quad (4.18)$$

den Schätzern für $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ entsprechen (vgl. Clements et al. (2007), S.125). Dabei bezeichnet \odot das Hadamard-Produkt, das für elementweise Multiplikation steht, $\tau = i_T \otimes \tau_H$ mit $\tau'_H = (H, H - 1, \dots, 1)$ und ζ bezeichnet den Störterm (vgl. auch Clements et al. (2007), S. 126). Unter **Annahme B** ist $\sigma_\varepsilon^2 = 0$ und $\sigma_{u^*}^2$ wird in der Regressionsgleichung

$$\hat{v} \odot \hat{v} = \rho_1^* \tau + \zeta^* \quad (4.19)$$

durch $\hat{\rho}_1^*$ geschätzt.

Somit können die Schätzer für σ_ε^2 und σ_u^2 in Gleichung (4.16) verwendet werden, so dass anschließend $\hat{\Sigma}$ zur Berechnung der Varianz-Kovarianz-Matrix von $\hat{\alpha}$ (Gleichung (4.14)) beziehungsweise von $\hat{\alpha}^H$ (Gleichung (4.15)) herangezogen werden kann.

Dieser gepoolte Ansatz wird in Kapitel 4.2 verwendet, um die eigenen SARIMA-Prognosen auf Unverzerrtheit zu testen. Außerdem wird in Kapitel 4.3 erstmalig untersucht, ob die offiziellen AKS-Prognosen auch für alle Prognosehorizonte zusammen unverzerrt sind. Zunächst wird allerdings erst der entsprechende Test auf Informationseffizienz vorgestellt.

4.1.2. Test auf Informationseffizienz

Nach dem Test auf Unverzerrtheit schließt sich die Frage an, ob die Prognosefehler mit Hilfe der Informationen, die zum Zeitpunkt $t-h$ der Erstellung der Prognose verfügbar sind, vorhersagbar sind oder nicht. Dementsprechend könnte eine Regression mit dem Prognosefehler $e_{t,h}$ als abhängige Variable und den verfügbaren Informationen, beispielsweise den vergangenen Prognosefehlern $e_{t,h-1}$, als Regressor geschätzt werden, um zu testen, ob der Koeffizient der verfügbaren Informationen signifikant von Null verschieden ist. Allerdings wäre in diesem Beispiel aufgrund der kumulierten makro-ökonomischen Schocks λ_{th} im zusammengesetzten Störterm v_{th} (vgl. Gleichung (4.6) sowie (4.7)) der Regressor $e_{t,h-1}$ bei vorliegen vergangener makro-ökonomischer Schocks u_{tj} mit $j = 1, \dots, h - 1$ mit dem Störterm dieser Schätzung

korreliert. Um dieses Endogenitätsproblem zu umgehen, schlagen Davies und Lahiri (1995) vor, nicht die Prognosefehler selbst, sondern die Änderungen der Prognosewerte beziehungsweise damit die Änderungen der Prognosefehler auf Effizienz zu testen (vgl. Davies und Lahiri (1995), S. 217).

Die Änderung zwischen der Prognose für das Prognoseziel t mit dem Prognosehorizont $h - 1$ und derjenigen für dasselbe Prognoseziel mit dem darauf folgenden Prognosehorizont h wird als

$$r_{t,h-1;h} = F_{t,h-1} - F_{t,h} = e_{t,h} - e_{t,h-1} \quad (4.20)$$

definiert (vgl. Clements et al. (2007), S. 126). Diese Änderung sollte mit Hilfe aller zum Zeitpunkt $t - h$ verfügbaren Informationen nicht vorhersagbar sein.

Um diese Hypothese der Informationseffizienz testen zu können, muss die betrachtete **Menge verfügbarer Informationen** allerdings eingeschränkt werden. Im Folgenden soll einerseits die Informationsmenge der Änderungen gleichbleibender aufeinanderfolgender Prognosehorizonte für verschiedene vergangene Prognoseziele $\{r_{t-1;h-1;h}, r_{t-2;h-1;h}, \dots\}$ (Fall 1) sowie andererseits die Informationsmenge der vergangenen Änderungen der Prognosen desselben Prognoseziels $\{r_{t;h-1+i;h+i}\}$ (Fall 2) für $i = 1, 2, \dots$ berücksichtigt werden. Die entsprechenden Tests auf Rationalität für diese beiden alternativen Informationsmengen werden im weiteren Verlauf getrennt voneinander in Erwägung gezogen, da bei einer simultanen Betrachtung die Kovarianz-Matrix des Störtermvektors zu komplex wäre (vgl. auch Clements et al. (2007), S. 126).

Zunächst wird Fall 1, also die **Informationsmenge der Änderungen der Prognosen für vergangene Prognoseziele**, betrachtet. Um die Korrelationsstruktur dieser Änderungen herzuleiten, wird die im Modell angenommene Zerlegung des

Prognosefehlers aus Gleichung (4.6) und Gleichung (4.7) mit horizont-spezifischer Verzerrung, also

$$e_{t,h} = \alpha_h + \varepsilon_{t,h} + \sum_{j=1}^h u_{tj}, \quad (4.21)$$

in Gleichung (4.20) eingesetzt, so dass die Änderung $r_{t;h-1;h}$ als

$$r_{t;h-1;h} = \alpha_h - \alpha_{h-1} + \varepsilon_{t,h} - \varepsilon_{t,h-1} + u_{t,h} \quad (4.22)$$

dargestellt werden kann. Unter den bisher getroffen Annahmen bezüglich der Störterme $\varepsilon_{t,h}$ und $u_{t,h}$ folgt daraus, dass die Kovarianz $Cov[r_{t;h-1;h}, r_{t-s;h-1;h}] = 0$ ist für alle $s > 0$ (vgl. auch Clements et al. (2007), S. 126). Somit sind die Änderungen rationaler Prognosen für aufeinanderfolgende Prognoseziele unkorreliert, egal ob der idiosynkratische Schock $\varepsilon_{t,h}$ im Modell gelassen wird (Annahme A), oder nicht (Annahme B).

Um die Prognosen auf Informationseffizienz zu testen, kann **einzeln für jeden Prognosehorizont** h die Nullhypothese $\gamma = 0$ in der Regressionsgleichung

$$r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t \quad \forall \quad t = 2, \dots, T \quad (4.23)$$

getestet werden, wobei der Koeffizient δ für den Achsenabschnitt, also für die Differenz der horizont-spezifischen Verzerrungen, steht, und ω_t den Störterm bezeichnet (vgl. Clements et al. (2007), S. 127). Unter der Nullhypothese ist demzufolge

$$\omega_t = r_{t;h-1;h} - \delta = \alpha_h - \alpha_{h-1} + \varepsilon_{t,h} - \varepsilon_{t,h-1} + u_{t,h} - \delta, \quad (4.24)$$

so dass die Varianz des Störterms $Var[\omega_t] = 2\sigma_\varepsilon^2 + \sigma_u^2$ für alle t beträgt und die Kovarianz $Cov[\omega_t \omega_{t-s}] = 0$ für alle $s \neq 0$ ist. Der Vektor ω der $T - 1$ Störterme ω_t hat folglich die diagonale Varianz-Kovarianz-Matrix

$$Cov[\omega \omega'] = (2\sigma_\varepsilon^2 + \sigma_u^2) I_{T-1}. \quad (4.25)$$

Im Modell ohne $\varepsilon_{t,h}$ (Annahme B), ist $\sigma_\varepsilon^2 = 0$ und σ_u^2 bezeichnet vereinfachend $\sigma_{u^*}^2$, also die Varianz des Störterms $u_{t,h}^*$, der sowohl die makro-ökonomischen Schocks als auch die Schocks, die aus der rationalen Nutzung privater Informationen resultieren, umfasst (s. Gleichung (4.11)). Die Analyse beschränkt sich hier gemäß Gleichung (4.20) auf Änderungen der Prognosen für aufeinanderfolgende Zeitpunkte t , also aufeinanderfolgende Quartale, so dass sowohl unter Annahme A als auch unter Annahme B die Varianz-Kovarianz-Matrix des Störterms proportional zur Einheitsmatrix ist. Würde man die Änderungen der Prognosen für weiter entfernte Zeitpunkte betrachten, wären diese aufgrund der makro-ökonomischen Schocks korreliert, so dass die Kovarianzen nicht mehr Null wären (vgl. Clements et al. (2007), S. 127, Fußnote 4).

Doch Gleichung (4.23) soll nicht nur separat für jeden Prognosehorizont $h = 2, \dots, H$ geschätzt werden, sondern die Änderungen der Prognosen werden wiederum gemäß Clements et al. (2007) auch **gepoolt** und somit für alle Prognosehorizonte zusammen auf Effizienz untersucht. Dabei werden die Spaltenvektoren $r_{1,2} = (r_{2;1;2}, \dots, r_{T;1;2})'$, $r_{2,3} = (r_{2;2;3}, \dots, r_{T;2;3})'$ bis $r_{H-1,H} = (r_{2;H-1;H}, \dots, r_{T;H-1;H})'$ der aufeinanderfolgenden Änderungen zum $(H-1)(T-1)$ -dimensionalen Spaltenvektor

$$r = (r_{1,2}, r_{2,3}, \dots, r_{H-1,H})' \quad (4.26)$$

zusammengesetzt. Analog wird der Vektor der verzögerten Änderungen

$$r_{-1} = (r_{1;1;2}, \dots, r_{T-1;1;2}, r_{1;2;3}, \dots, r_{T-1;2;3}, \dots, r_{1;H-1;H}, \dots, r_{T-1;H-1;H})' \quad (4.27)$$

definiert, um dann anhand der Regressionsgleichung

$$r = \gamma r_{-1} + \delta_d + \omega \quad (4.28)$$

die Nullhypothese der Informationseffizienz, $H_0 : \gamma = 0$, zu testen (vgl. Clements et al. (2007), S. 127). In Gleichung (4.28) ist $\delta_d = (I_{H-1} \otimes i_{T-1})\delta^h$ und der Parametervektor $\delta^h = (\delta_1, \delta_2, \dots, \delta_{H-1})$. Somit stellt δ_d einen Vektor von Dummyvaria-

blen für die $(H - 1)$ Prognosehorizonte dar, wodurch horizont-spezifische Verzerrungen im Modell zugelassen werden (Variante II). Geht man von einer gleich großen Verzerrung für alle Prognosehorizonte aus (Variante I), wird in Gleichung (4.28) $\delta_d = i_{T-1, H-1} \delta$ gesetzt.

Die Varianz-Kovarianz-Matrix Υ des Störtermvektors ω hängt wiederum vom maximalen Prognosehorizont H ab. Für $H = 8$ hat sie unter der Nullhypothese die Form

$$\Upsilon = E[rr'] = \begin{bmatrix} A & B & C & D & E & F & G \\ B' & A & B & C & D & E & F \\ C' & B' & A & B & C & D & E \\ D' & C' & B' & A & B & C & D \\ E' & D' & C' & B' & A & B & C \\ F' & E' & D' & C' & B' & A & B \\ G' & F' & E' & D' & C' & B' & A \end{bmatrix}, \quad (4.29)$$

wobei $A = (2\sigma_\varepsilon^2 + \sigma_u^2)I_{T-1}$ gilt (vgl. auch Gleichung (4.25)), und die $(T - 1) \times (T - 1)$ -dimensionalen Komponentenmatrizen setzen sich folgendermaßen zusammen (vgl. Clements et al. (2007), S. 127):

$$B = \begin{bmatrix} -\sigma_\varepsilon^2 & \sigma_u^2 & 0 & \cdots & & 0 \\ 0 & -\sigma_\varepsilon^2 & \sigma_u^2 & 0 & \cdots & \vdots \\ \vdots & 0 & -\sigma_\varepsilon^2 & \sigma_u^2 & \ddots & \\ & \vdots & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \sigma_u^2 \\ 0 & \cdots & & & 0 & -\sigma_\varepsilon^2 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & \sigma_u^2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \sigma_u^2 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & & & & \ddots & \ddots & \sigma_u^2 \\ & & & & & & \ddots & 0 & 0 \\ 0 & \cdots & & & 0 & 0 & 0 \end{bmatrix},$$

D ist wie C, nur dass die Elemente (1,4), (2,5), (3,6)... σ_u^2 enthalten, während die Elemente (1,3), (2,4), (3,5)... gleich Null sind. Analog werden auch die Komponentenmatrizen E, F und G gebildet, das heißt in E sind die Elemente (1,5), (2,6),

(3,7)..., in F die Elemente (1,6), (2,7), (3,8)... und in G die Elemente (1,7), (2,8), (3,9)... gleich σ_u^2 .

Die bereits durch Gleichung (4.18) beziehungsweise Gleichung (4.19) geschätzten Varianzen $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ werden dann in die Komponentenmatrizen A bis G eingesetzt, um die geschätzte Varianz-Kovarianz-Matrix \hat{Y} zu bekommen. Diese wird anschließend zur Berechnung der Varianz-Kovarianz-Matrix des Parametervektors $(\hat{\gamma}, \hat{\delta}^h)$ (beziehungsweise $(\hat{\gamma}, \hat{\delta})$ für Variante I)

$$Cov[(\hat{\gamma}, \hat{\delta}^h)] = (X'X)^{-1}X'\hat{Y}X(X'X)^{-1} \quad (4.30)$$

mit der Regressormatrix $X = (r_{-1} \ \delta_d)$ verwendet, um konsistente Schätzer der Standardfehler des Parametervektors in der OLS-Regression der Gleichung (4.28) zu bekommen.

Als Nächstes soll die **Informationsmenge der vergangenen Änderungen der Prognosen desselben Prognoseziels** in Betracht gezogen werden (Fall 2). Die Nullhypothese der Informationseffizienz beinhaltet nun, dass die vergangenen Prognoseänderungen $\{r_{t,h-1+i,h+i}\}$ für $i = 1, 2, \dots$ keinen Erklärungsgehalt für die aktuelle Änderung desselben Prognoseziels $r_{t,h-1,h}$ haben. Aus Gleichung (4.22) lässt sich direkt ableiten, dass $Cov[r_{t,h-1,h}, r_{t,h,h+1}] = E[-\varepsilon_{th}^2] = -\sigma_\varepsilon^2$. Da die aktuelle Prognoseänderung also (zumindest unter Annahme A) mit der direkt vorhergehenden Änderung aufgrund des idiosynkratischen Störterms korreliert ist, wird für die Regressionsgleichung anstatt $r_{t,h,h+1}$ die ein weiteres Quartal früher zustande gekommene Änderung $r_{t,h+1,h+2}$ als Regressor verwendet (vgl. Clements et al. (2007), S. 128). Somit ergibt sich die Regressionsgleichung

$$r_{t,h-1,h} = \gamma r_{t,h+1,h+2} + \delta + \omega_t \quad (\forall \ t = 1, \dots, T), \quad (4.31)$$

um die Nullhypothese der Informationseffizienz, $H_0 : \gamma = 0$, **getrennt für jeden Prognosehorizont** $h = 2, \dots, T - 2$ zu testen. Die $(T \times T)$ -dimensionale Varianz-Kovarianz-Matrix des Störtermvektors ω ergibt sich analog zu Gleichung (4.25) als

$$\text{Cov}[\omega\omega'] = (2\sigma_\varepsilon^2 + \sigma_u^2)I_T. \quad (4.32)$$

Im Modell ohne $\varepsilon_{t,h}$ (Annahme B), ist $\sigma_\varepsilon^2 = 0$ und σ_u^2 bezeichnet wiederum vereinfachend $\sigma_{u^*}^2$ (vgl. Gleichung (4.11)).

Darüber hinaus werden die Änderungen auch wieder gemäß Clements et al. (2007) **gepoolt** und somit für alle Prognosehorizonte zusammen auf Informationseffizienz getestet. Dabei werden nun die Spaltenvektoren $r_{1,2} = (r_{1;1;2}, \dots, r_{T;1;2})'$, $r_{2,3} = (r_{1;2;3}, \dots, r_{T;2;3})'$ bis $r_{H-1,H} = (r_{1;H-1;H}, \dots, r_{T;H-1;H})'$ für alle Beobachtungen T definiert. Für $H = 8$ werden analog zu Gleichung (4.31) der zusammengesetzte Spaltenvektor der Prognoseänderungen

$$r = (r_{1,2}, r_{2,3}, \dots, r_{5,6})' \quad (4.33)$$

sowie der gepoolte Vektor der verzögerten Prognoseänderungen

$$r_{-1} = (r_{3,4}, r_{4,5}, \dots, r_{7,8})' \quad (4.34)$$

gebildet, um in der entsprechenden Regressionsgleichung

$$r = \gamma r_{-1} + \delta_d + \omega \quad (4.35)$$

die Nullhypothese der Informationseffizienz, $H_0 : \gamma = 0$, zu testen. In Gleichung (4.35) steht δ_d nun folglich für $\delta_d = (I_5 \otimes i_T)\delta^h$ mit dem Parametervektor $\delta^h = (\delta_1, \delta_2, \dots, \delta_5)$. Somit werden durch den Vektor δ_d der Dummyvariablen für die Prognosehorizonte wiederum horizont-spezifische Verzerrungen im Modell zugelassen (Variante II). Geht man von einer gleich großen Verzerrung für alle Prognosehorizonte aus (Variante I), wird in Gleichung (4.35) $\delta_d = i_{5T} \delta$ gesetzt.

Die Varianz-Kovarianz-Matrix Υ des Störtermvektors ω hängt ebenfalls stark vom maximalen Prognosehorizont H ab. Für $H = 8$ hat sie unter der Nullhypothese nun die Form

$$\Upsilon = E[rr'] = \begin{bmatrix} A & B & C & D & E \\ B' & A & B & C & D \\ C' & B' & A & B & C \\ D' & C' & B' & A & B \\ E' & D' & C' & B' & A \end{bmatrix}, \quad (4.36)$$

wobei $A = (2\sigma_\varepsilon^2 + \sigma_u^2)I_T$. Die $(T \times T)$ -Komponentenmatrizen B, C, D und E setzen sich analog zu den oben beschriebenen zusammen (vgl. auch Clements et al. (2007), S. 128).

Die bereits durch Gleichung (4.18) beziehungsweise Gleichung (4.19) geschätzten Varianzen $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ werden dann wiederum in die Komponentenmatrizen A bis E eingesetzt, um die geschätzte Varianz-Kovarianz-Matrix $\hat{\Upsilon}$ zu bekommen, die anschließend zur Berechnung der Varianz-Kovarianz-Matrix des Parametervektors $(\hat{\gamma}, \hat{\delta}^h)$ (beziehungsweise $(\hat{\gamma}, \hat{\delta})$ für Variante I)

$$Cov[(\hat{\gamma}, \hat{\delta}^h)] = (X'X)^{-1}X'\hat{\Upsilon}X(X'X)^{-1} \quad (4.37)$$

mit der Regressormatrix $X = (r_{-1} \delta_d)$ verwendet wird, um konsistente Schätzer der Standardfehler des Parametervektors in der OLS-Regression der Gleichung (4.35) zu bekommen.

In den beiden folgenden Kapiteln werden nun sowohl die verschiedenen Varianten des Tests auf Unverzerrtheit als auch die beiden vorgestellten Tests auf Informationseffizienz zuerst auf die eigenen SARIMA-Prognosen des Steueraufkommens und anschließend auf die offiziellen Prognosen des Arbeitskreises „Steuerschätzungen“ angewandt.

4.2. Untersuchung der Prognoserationalität der SARIMA-Schätzungen

Anhand der erläuterten Tests wird nun untersucht, ob die beiden in Kapitel 3 vorgeschlagenen Prognosemodelle geeignet sind, rationale Prognosen des Gesamtsteueraufkommens der BRD zu erstellen. Zunächst werden die Prognosen der ermittelten SARIMA-Modellspezifikationen für unterschiedliche maximale Prognosehorizonte ($H = 4, 8, 12, 16$) auf Unverzerrtheit getestet, wobei jeweils sowohl Variante I als auch Variante II unter beiden möglichen Annahmen bezüglich des idiosynkratischen Störterms berücksichtigt werden. Anschließend werden diejenigen Prognosen, die sich als unverzerrt herausstellen, außerdem auf Effizienz gemäß der beiden vorgestellten Informationsmengen getestet.

4.2.1. Test auf Unverzerrtheit der SARIMA-Prognosen

Für die kurzfristigen Prognosen des Gesamtsteueraufkommens mit einem maximalen Prognosehorizont von $H = 4$ Quartalen (Ein-Jahres-Prognosen) anhand des SARIMA(1,0,0)(1,0,1)₄-Modells, das sich im Box-Jenkins-Verfahren als das **beste Modell** herausstellte, ergeben sich für die **Variante I**, also mit einer gleich großen Verzerrung über alle Prognosehorizonte als Alternativhypothese, die Testergebnisse in Tabelle 4.1.

Sowohl unter Annahme A als auch unter Annahme B ergibt sich als Schätzwert für α eine insignifikante negative Verzerrung von ungefähr -1,4 Milliarden Euro. Das Gesamtsteueraufkommen der BRD wird also durch das SARIMA(1,0,0)(1,0,1)₄-Modell im betrachteten Zeitraum durchschnittlich um etwa 1,4 Milliarden Euro pro Quartal überschätzt, wenn der maximale Prognosehorizont auf $H = 4$ Quartale begrenzt ist. Gemäß des jeweiligen p -Wertes von circa 0,44 beziehungsweise 0,46 kann die Nullhypothese der Unverzerrtheit allerdings unter beiden möglichen Annahmen bezüglich des idiosynkratischen Störterms eindeutig nicht verworfen werden.

Tabelle 4.1.: **Ergebnisse des t -Tests für Variante I mit $H = 4$**

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -1,40 | -1,40 |
| $se(\hat{\alpha})$ | 1,82 | 1,88 |
| p -Wert | 0,44 | 0,46 |
| $\hat{\sigma}_\varepsilon^2$ | 5,71 | – |
| $\hat{\sigma}_u^2$ | 21,44 | 23,34 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, \dots, 4$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Für **Variante II** sind die Ergebnisse des F -Tests in Tabelle 4.2 zusammengestellt. Wenn für jeden Prognosehorizont h eine horizont-spezifische Verzerrung in der Alternativhypothese zugelassen wird (Variante II), nehmen die geschätzten Verzerrungen $\hat{\alpha}_h$ mit steigendem Prognosehorizont zu, wobei aber alle (einzeln anhand des t -Tests getestet) sowohl unter Annahme A als auch unter Annahme B insignifikant sind.³ Werden alle geschätzten Verzerrungen gemeinsam auf Signifikanz getestet, ergibt der entsprechende F -Test ebenfalls, dass die Nullhypothese der Unverzerrtheit der Prognosen, $\hat{\alpha}^H = 0$, unter beiden Annahmen bezüglich des idiosynkratischen Störterms nicht verworfen werden kann. Die Testentscheidung fällt allerdings unter Hinzunahme des Störterms ε_{th} (Annahme A) mit einem p -Wert von ungefähr 0,92 wesentlich deutlicher aus als unter Annahme B mit einem p -Wert von circa 0,15.

³Die detaillierten Regressionsergebnisse sind in Tabelle B.1 im Anhang dargestellt.

Tabelle 4.2.: **Ergebnisse des F -Tests für Variante II mit $H = 4$**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,24 | 1,72 |
| p -Wert | 0,92 | 0,15 |
| $\hat{\sigma}_\varepsilon^2$ | 5,52 | – |
| $\hat{\sigma}_u^2$ | 21,34 | 23,18 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, \dots, 4$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Wird der maximale Prognosehorizont also mit $H = 4$ auf ein Jahr begrenzt, bestätigen beide Testvarianten die **Unverzerrtheit** der eigenen Prognosen anhand des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells. Die Nullhypothese der Unverzerrtheit kann außerdem auch für die Prognosen des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells (mit $H = 4$), das sich im Box-Jenkins-Verfahren als die **zweitbeste Modellspezifikation** erwies, sowohl für Variante I als auch für Variante II nicht verworfen werden.⁴ Die Annahme bezüglich des idiosynkratischen Störterms spielt auch bei dieser Modellvariante keine entscheidende Rolle für die Testergebnisse.

Nun stellt sich jedoch die Frage, ob die Prognosen auch dann noch unverzerrt sind, wenn sich der maximale Prognosehorizont H erhöht. Dabei gilt es allerdings zu beachten, dass die Matrix Θ und damit auch Ψ sowie die Varianz-Kovarianz-Matrix Σ , wie in Kapitel 4.1.1 erläutert, maßgeblich von H abhängen. Um die SARIMA-Prognosen in Kapitel 4.3 mit denjenigen des Arbeitskreises „Steuerschätzungen“ vergleichen zu können, sind zusätzlich zu den bereits untersuchten Ein-

⁴Tabelle B.2 und Tabelle B.3 im Anhang stellen die entsprechenden Testergebnisse dar.

Jahres-Prognosen (mit $H = 4$) auch die Zwei-Jahres-Prognosen (mit $H = 8$), die Drei-Jahres-Prognosen (mit $H = 12$) sowie die Vier-Jahres-Prognosen (mit $H = 16$) von besonderem Interesse.

In Tabelle 4.3 sind daher zunächst für den maximalen Prognosehorizont von $H = 8$ die Ergebnisse des t -Tests (**Variante I**) auf Unverzerrtheit der Prognosen des Gesamtsteueraufkommens der BRD anhand des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells zusammengefasst.

Tabelle 4.3.: Ergebnisse des t -Tests für Variante I mit $H = 8$

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -2,82 | -2,82 |
| $se(\hat{\alpha})$ | 3,37 | 3,43 |
| p -Wert | 0,40 | 0,41 |
| $\hat{\sigma}_\varepsilon^2$ | 4,78 | – |
| $\hat{\sigma}_u^2$ | 23,72 | 24,56 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, \dots, 8$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Bei einem maximalen Prognosehorizont von $H = 8$ Quartalen erhöht sich der Schätzwert für α auf eine negative Verzerrung von ungefähr -2,82 Milliarden Euro. Das Gesamtsteueraufkommen der BRD wird nun also durch das SARIMA(1,0,0)(1,0,1)₄-Modell durchschnittlich um etwa 2,8 Milliarden Euro pro Quartal überschätzt. Gemäß der jeweiligen p -Werte kann die Nullhypothese der Unverzerrtheit allerdings auch noch für die Zwei-Jahres-Prognosen sowohl unter Annahme A als auch unter Annahme B eindeutig nicht verworfen werden.

Wenn für jeden Prognosehorizont h eine horizont-spezifische Verzerrung in der Alternativhypothese zugelassen wird, also in **Variante II**, ergeben sich die in Tabelle 4.4 dargestellten Ergebnisse des F -Tests.

Tabelle 4.4.: **Ergebnisse des F -Tests für Variante II mit $H = 8$**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,37 | 1,87 |
| p -Wert | 0,93 | 0,06 |
| $\hat{\sigma}_\varepsilon^2$ | 1,99 | – |
| $\hat{\sigma}_u^2$ | 23,82 | 24,17 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, \dots, 8$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Auch für die Zwei-Jahres-Prognosen nehmen die geschätzten Verzerrungen $\hat{\alpha}_h$ erwartungsgemäß mit steigendem Prognosehorizont zu. Allerdings sind auch für den maximalen Prognosehorizont $H = 8$ alle Verzerrungen (einzeln anhand des t -Tests getestet) sowohl unter Annahme A als auch unter Annahme B insignifikant.⁵ Wenn alle geschätzten Verzerrungen gemeinsam auf Signifikanz getestet werden, kann die Nullhypothese der Unverzerrtheit, $\hat{\alpha}^H = 0$, des entsprechenden F -Tests unter Annahme A eindeutig nicht verworfen werden. Wird unter Annahme B jedoch kein idiosynkratischer Störterm im Modell berücksichtigt, fällt das Testergebnis mit einem p -Wert von ungefähr 6% relativ knapp aus. Ausgehend von einem Signifikanzniveau von maximal 5% kann die Nullhypothese der Unverzerrtheit dennoch ebenfalls nicht verworfen werden.

⁵Die detaillierten Regressionsergebnisse sind in Tabelle B.4 im Anhang dargestellt.

Somit erweisen sich die Prognosen anhand des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells also auch für einen maximalen Prognosehorizont von $H = 8$ Quartalen in beiden Testvarianten als **unverzerrt**. Des Weiteren resultieren diese Testergebnisse ebenso für die Prognosen des **alternativen SARIMA(0,0,1)(1,0,1)₄-Modells**. Auch für diese Modellvariante kann ausgehend von einem Signifikanzniveau von maximal 5% sowohl für Variante I als auch für Variante II die Nullhypothese der Unverzerrtheit nicht verworfen werden.⁶

Im nächsten Schritt werden die Drei-Jahres-Prognosen untersucht, wofür die Matrix Θ und somit auch Ψ sowie die Varianz-Kovarianz-Matrix Σ erneut an den maximalen Prognosehorizont $H = 12$ angepasst werden. Tabelle 4.5 gibt die Ergebnisse des t -Tests auf Unverzerrtheit der Prognosen des Gesamtsteueraufkommens der BRD anhand des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells für **Variante I** und $H = 12$ wieder.

Tabelle 4.5.: **Ergebnisse des t -Tests für Variante I mit $H = 12$**

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -3,98 | -3,98 |
| $se(\hat{\alpha})$ | 5,71 | 5,50 |
| p -Wert | 0,49 | 0,47 |
| $\hat{\sigma}_\varepsilon^2$ | (-20,01) | – |
| $\hat{\sigma}_u^2$ | 33,66 | 31,26 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, \dots, 12$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

⁶Die Testergebnisse sind Tabelle B.5 und Tabelle B.6 im Anhang zu entnehmen.

Da sich unter Annahme A ein negativer Schätzwert für die Varianz σ_ε^2 ergibt, erscheint das Modell mit idiosynkratischem Störterm ε_{th} hier zwar fraglich. Allerdings unterscheiden sich die Schätzwerte für die Verzerrung α , ihres Standardfehlers sowie der Varianz σ_u^2 unter Annahme A kaum von denen unter Annahme B. Die Nullhypothese der Unverzerrtheit kann jedenfalls in beiden Fällen eindeutig nicht verworfen werden, wenn als Alternativhypothese dieselbe Verzerrung α für alle Prognosehorizonte h vorausgesetzt wird.

Wird in der Alternativhypothese jedoch für jeden Prognosehorizont h eine horizontspezifische Verzerrung zugelassen (**Variante II**), ergeben sich die in Tabelle 4.6 dargestellten Ergebnisse des F -Tests.

Tabelle 4.6.: **Ergebnisse des F -Tests für Variante II mit $H = 12$**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,18 | 1,44 |
| p -Wert | 0,999 | 0,14 |
| $\hat{\sigma}_\varepsilon^2$ | (-28,41) | – |
| $\hat{\sigma}_u^2$ | 34,22 | 30,81 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, \dots, 12$ Prognosehorizonte gegen die Alternativhypothese horizontspezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise drei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Die geschätzten Verzerrungen $\hat{\alpha}_h$ nehmen erneut mit steigendem Prognosehorizont zu, wobei alle (einzeln anhand des t -Tests getestet) sowohl unter Annahme A als auch unter Annahme B insignifikant sind.⁷ Werden alle geschätzten Verzerrungen

⁷Die detaillierten Regressionsergebnisse sind in Tabelle B.7 im Anhang dargestellt.

gemeinsam auf Signifikanz getestet, ergibt sich zwar unter Annahme A wiederum ein negativer Schätzwert für die Varianz σ_ε^2 . Die Nullhypothese der Unverzerrtheit kann jedoch unter beiden Annahmen auch für die Drei-Jahres-Prognosen eindeutig nicht verworfen werden. Unter Annahme B, also im Modell ohne ε_{th} , in dem $\sigma_\varepsilon^2 = 0$ gesetzt wird, fällt die Testentscheidung allerdings auch hier wieder deutlich knapper aus als im Modell mit idiosynkratischem Störterm.

Die Testergebnisse implizieren folglich, dass die Prognosen anhand des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells selbst dann noch als **unverzerrt** betrachtet werden können, wenn der maximale Prognosehorizont auf $H = 12$ erhöht wird. Diese Testentscheidungen ergeben sich ebenfalls für die Prognosen anhand des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells (für $h = 1, \dots, 12$ Prognosehorizonte).⁸

Abschließend werden die Vier-Jahres-Prognosen auf Unverzerrtheit getestet. Dabei werden die Matrizen Θ , Ψ sowie die Varianz-Kovarianz-Matrix Σ wiederum an den maximalen Prognosehorizont $H = 16$ angepasst. Für die Prognose des logarithmierten Gesamtsteueraufkommens anhand des SARIMA(1,0,0)(1,0,1)₄-Modells mit $H = 16$ ergeben sich für die Variante I die Testergebnisse in Tabelle 4.7.

Für die Vier-Jahres-Prognosen ergibt sich als Schätzwert für α eine insignifikante negative Verzerrung von ungefähr -4,8 Milliarden Euro. Das Gesamtsteueraufkommen der BRD wird also durch das SARIMA(1,0,0)(1,0,1)₄-Modell durchschnittlich um etwa 4,8 Milliarden Euro pro Quartal überschätzt, wenn der maximale Prognosehorizont auf $H = 16$ Quartale erhöht wird. Dabei ergibt sich unter Annahme A wieder ein negativer Schätzwert für die Varianz σ_ε^2 , so dass das Modell mit idiosynkratischem Störterm ε_{th} fraglich erscheint. Allerdings kann die Nullhypothese der Unverzerrtheit erneut sowohl unter Annahme A als auch unter Annahme B eindeutig nicht verworfen werden, wenn als Alternativhypothese dieselbe Verzerrung α für alle Prognosehorizonte h vorausgesetzt wird.

⁸Tabelle B.8 und Tabelle B.9 im Anhang stellen die entsprechenden Testergebnisse dar.

Tabelle 4.7.: **Ergebnisse des t -Tests für Variante I mit $H = 16$**

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -4,81 | -4,81 |
| $se(\hat{\alpha})$ | 8,40 | 7,86 |
| p -Wert | 0,57 | 0,54 |
| $\hat{\sigma}_\varepsilon^2$ | (-60,26) | – |
| $\hat{\sigma}_u^2$ | 43,86 | 38,38 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, \dots, 16$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Wenn für jeden Prognosehorizont h jedoch eine horizont-spezifische Verzerrung in der Alternativhypothese zugelassen wird, also in **Variante II**, ergeben sich die in Tabelle 4.8 dargestellten Ergebnisse des F -Tests.

Für den F -Test wirkt sich also die Annahme bezüglich des idiosynkratischen Störterms entscheidend auf das Testergebnis aus. Wird ε_{th} als prognose-spezifischer Fehler im Modell gelassen, ergibt sich zwar wiederum ein negativer Schätzwert für die Varianz σ_ε^2 , die Nullhypothese der Unverzerrtheit kann aber weiterhin nicht verworfen werden. Im Modell ohne ε_{th} , in dem $\sigma_\varepsilon^2 = 0$ gesetzt wird, ändert sich dagegen die Testentscheidung, so dass nicht davon ausgegangen werden kann, dass die Prognosen anhand des SARIMA(1,0,0)(1,0,1)₄-Modells uneingeschränkt als unverzerrt beurteilt werden können, wenn der maximale Prognosehorizont auf $H = 16$ erhöht wird. Auch die Prognosen anhand des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells (für $h = 1, \dots, 16$ Prognosehorizonte) liefern diese Testergebnisse.⁹

⁹Tabelle B.10 und Tabelle B.11 im Anhang stellen die entsprechenden Testergebnisse dar.

Tabelle 4.8.: **Ergebnisse des F -Tests für Variante II mit $H = 16$**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,14 | 2,88 |
| p -Wert | 0,99997 | 0,00014 |
| $\hat{\sigma}_\varepsilon^2$ | (-72,59) | – |
| $\hat{\sigma}_u^2$ | 44,60 | 38,00 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, \dots, 16$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise fünf Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Die bisherigen Testergebnisse implizieren folglich, dass die eigenen Ein-, Zwei- und Drei-Jahres-Prognosen anhand der beiden ausgewählten SARIMA-Modelle **unverzerrt** sind, wogegen die Testentscheidung für die Vier-Jahres-Prognosen von den Modellannahmen abhängt. Daher beschränkt sich die weitere Analyse der Informationseffizienz auf die **unverzerrten Prognosen mit einem maximalen Prognosehorizont von $H = 4$, $H = 8$ sowie $H = 12$ Quartalen.**

4.2.2. Test auf Informationseffizienz der SARIMA-Prognosen

Für die Ein-Jahres-Prognosen des Gesamtsteueraufkommens der BRD anhand des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells (mit $H = 4$) wird zunächst auf **traditionelle Art und Weise** die Regressiongleichung (4.23)

$$r_{t,h-1,h} = \gamma r_{t-1,h-1,h} + \delta + \omega_t \quad \forall \quad t = 2, \dots, T$$

separat für jeden Prognosehorizont ($h = 2, 3, 4$) geschätzt, um jeweils die Nullhypothese der Informationseffizienz, $H_0 : \gamma = 0$, wie in Kapitel 4.1.2 beschrieben zu testen. Zur Berechnung des Schätzers der Varianz-Kovarianz-Matrix des Störtermvektors ω gemäß Gleichung (4.25) werden beide Modellvarianten berücksichtigt, also einerseits das Modell mit ε_{th} (Annahme A) sowie andererseits das Modell ohne idiosynkratischen Störterm (Annahme B). Die Ergebnisse der entsprechenden t -Tests sind in Tabelle 4.9 dargestellt.

Tabelle 4.9.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 1 der Prognosen mit $H = 4$**

| h | $\hat{\gamma}$ | Annahme A | | Annahme B | |
|-----|----------------|--------------------|-----------|--------------------|-----------|
| | | $se(\hat{\gamma})$ | p -Wert | $se(\hat{\gamma})$ | p -Wert |
| 2 | -0,01 | 0,25 | 0,98 | 0,21 | 0,98 |
| 3 | -0,06 | 0,24 | 0,79 | 0,20 | 0,76 |
| 4 | -0,08 | 0,16 | 0,60 | 0,13 | 0,54 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2, 3, 4$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für jedes h sind unter beiden Annahmen die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells mit $H = 4$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Sowohl unter Annahme A als auch unter Annahme B kann die Nullhypothese der Informationseffizienz für alle Prognosehorizonte also eindeutig nicht verworfen werden, wenn der maximale Prognosehorizont auf $H = 4$ begrenzt wird. Im nächsten Schritt werden die Änderungen der SARIMA-Prognosen gemäß Gleichung (4.26) und Gleichung (4.27) **gepoolt** und die Regressionsgleichung (4.28)

$$r = \gamma r_{-1} + \delta_d + \omega$$

für alle Prognosehorizonte zusammen geschätzt, um die Nullhypothese der Informationseffizienz, $H_0 : \gamma = 0$, zu testen.

Tabelle 4.10 gibt die entsprechenden Testergebnisse für Variante I (gleich große Verzerrung über alle Prognosehorizonte) sowie Variante II (horizont-spezifische Verzerrungen) jeweils unter Annahme A (Modell mit ε_{th}) und unter Annahme B (Modell ohne ε_{th}) wieder.

Tabelle 4.10.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 1 der Prognosen mit $H = 4$**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | -0,06 | -0,06 | -0,06 | -0,06 |
| $se(\hat{\gamma})$ | 0,17 | 0,16 | 0,17 | 0,16 |
| p -Wert | 0,73 | 0,71 | 0,72 | 0,71 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 1). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells mit $H = 4$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Demnach kann die Nullhypothese der Informationseffizienz also ebenfalls eindeutig nicht verworfen werden, unabhängig davon, ob bei der Modellierung horizont-spezifische Verzerrungen zugelassen werden oder nicht. Auch die Annahme bezüglich des idiosynkratischen Störterms ε_{th} hat keinen Einfluss auf die Testentscheidung. Die Steuerschätzungen anhand des SARIMA(1,0,0)(1,0,1)₄-Modells mit einem maximalen Prognosehorizont von $H = 4$ Quartalen erweisen sich also als **effizient**, wenn als zugrunde liegende **Informationsmenge die Änderungen der Prognosen gleichbleibender aufeinanderfolgender Prognosehorizonte für ver-**

schiedene vergangene Prognoseziele $\{r_{t-1;h-1;h}, r_{t-2;h-1;h}, \dots\}$ betrachtet wird. Diese Testentscheidungen ergeben sich ebenfalls für die Prognosen anhand des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells (mit $H = 4$).¹⁰

Als Nächstes soll die **Informationsmenge der vergangenen Änderungen der Prognosen desselben Prognoseziels** in Betracht gezogen werden (Fall 2). Erneut werden zuerst die Prognosen des Gesamtsteueraufkommens der BRD anhand des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells (mit $H = 4$) auf **traditionelle Art und Weise** auf diese zweite Form der Informationseffizienz getestet. Dabei kann die Regressionsgleichung (4.31)

$$r_{t;h-1;h} = \gamma r_{t;h+1;h+2} + \delta + \omega_t \quad (\forall \quad t = 1, \dots, T),$$

aufgrund des geringen maximalen Prognosehorizonts jedoch lediglich für $h = 2$ geschätzt werden, um die Nullhypothese der Informationseffizienz, $H_0 : \gamma = 0$, zu testen. Zur Berechnung der geschätzten Varianz-Kovarianz-Matrix des Störtermvektors ω gemäß Gleichung (4.32) werden auch wieder beide Modellvarianten berücksichtigt, also einerseits das Modell mit ε_{th} (Annahme A) sowie andererseits das Modell ohne idiosynkratischen Störterm (Annahme B). Die Ergebnisse des entsprechenden t-Tests sind in Tabelle 4.11 dargestellt.

Auch für den zweiten Fall des Tests auf Informationseffizienz kann die Nullhypothese $H_0 : \gamma = 0$ eindeutig nicht verworfen werden. Dabei spielt es keine Rolle, ob der idiosynkratische Störterm ε_{th} im Modell gelassen wird (Annahme A) oder nicht (Annahme B). Die eigenen Steuerschätzungen anhand des SARIMA(1,0,0)(1,0,1)₄-Modells mit $H = 4$ können demnach also auch als **effizient** beurteilt werden, wenn als zugrunde liegende **Informationsmenge die vergangenen Änderungen der Prognosen desselben Prognoseziels** $\{r_{t;h-1+i;h+i}\}$ betrachtet wird. Dieses Testergebnis kommt ebenfalls für die Prognosen anhand des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells mit $H = 4$ zu Stande.¹¹ Für die Ein-Jahres-Pro-

¹⁰Tabelle B.12 und Tabelle B.13 im Anhang stellen die entsprechenden Testergebnisse dar.

¹¹Die Testergebnisse sind Tabelle B.14 im Anhang zu entnehmen.

Tabelle 4.11.: **Ergebnisse des separaten Tests auf Informationseffizienz gemäß Fall 2 der Prognosen mit $H = 4$**

| h | $\hat{\gamma}$ | Annahme A | | Annahme B | |
|-----|----------------|--------------------|-----------|--------------------|-----------|
| | | $se(\hat{\gamma})$ | p -Wert | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,03 | 0,16 | 0,84 | 0,13 | 0,81 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz für $h = 2$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t;h+1;h+2} + \delta + \omega_t$ (Fall 2). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Der Wert des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests sind für beide Annahmen auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells mit $H = 4$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

gnosen können die Änderungen der Prognosen nicht gemäß Gleichung (4.33) und Gleichung (4.34) gepoolt werden, da der maximale Prognosehorizont von $H = 4$ Quartalen dafür nicht ausreicht.

Nun stellt sich jedoch die Frage, ob sich die Prognosen auch dann noch als effizient erweisen, wenn sich der maximale Prognosehorizont H erhöht. Auch für die Zwei-Jahres-Prognosen des Gesamtsteueraufkommens der BRD anhand des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells mit $H = 8$ wird zunächst die Regressionsgleichung (4.23) **separat für jeden Prognosehorizont** ($h = 2, \dots, 8$) geschätzt, um jeweils die Nullhypothese der Informationseffizienz, $H_0 : \gamma = 0$, zu testen. Zur Berechnung des Schätzers der Varianz-Kovarianz-Matrix des Störtermvektors ω gemäß Gleichung (4.25) werden erneut beide Modellvarianten berücksichtigt, also einerseits das Modell mit ε_{th} (Annahme A) sowie andererseits das Modell ohne idiosynkratischen Störterm (Annahme B). Die Ergebnisse der entsprechenden t -Tests sind in Tabelle 4.12 dargestellt.

Tabelle 4.12.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 1 der Prognosen mit $H = 8$**

| h | $\hat{\gamma}$ | Annahme A | | Annahme B | |
|-----|----------------|--------------------|-----------|--------------------|-----------|
| | | $se(\hat{\gamma})$ | p -Wert | $se(\hat{\gamma})$ | p -Wert |
| 2 | -0,002 | 0,24 | 0,99 | 0,21 | 0,99 |
| 3 | -0,060 | 0,24 | 0,80 | 0,21 | 0,77 |
| 4 | -0,057 | 0,15 | 0,71 | 0,13 | 0,67 |
| 5 | -0,050 | 0,16 | 0,76 | 0,14 | 0,72 |
| 6 | -0,144 | 0,17 | 0,41 | 0,15 | 0,34 |
| 7 | -0,140 | 0,17 | 0,41 | 0,14 | 0,34 |
| 8 | -0,155 | 0,13 | 0,23 | 0,11 | 0,16 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2, \dots, 8$, anhand der Regressionsgleichung: $r_{t,h-1;h} = \gamma r_{t-1,h-1;h} + \delta + \omega_t$ (Fall 1). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für jedes h sind unter beiden Annahmen die Werte des geschätzten Koeffizienten $\hat{\gamma}$ auf drei, diejenigen seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells mit $H = 8$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Auch für die Prognosen mit maximalem Prognosehorizont $H = 8$ kann folglich die Nullhypothese der Informationseffizienz sowohl unter Annahme A als auch unter Annahme B für alle Prognosehorizonte h eindeutig nicht verworfen werden.

Im nächsten Schritt werden wiederum die Änderungen der SARIMA-Prognosen gemäß Gleichung (4.26) und Gleichung (4.27) **gepoolt** und die Regressionsgleichung (4.28) für alle Prognosehorizonte zusammen geschätzt, um die Nullhypothese der Informationseffizienz, $H_0 : \gamma = 0$, zu testen. Tabelle 4.13 gibt die entsprechenden Testergebnisse für beide Modellvarianten und beide Annahmen bezüglich des idiosynkratischen Störterms wieder.

Tabelle 4.13.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 1 der Prognosen mit $H = 8$**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | -0,10 | -0,10 | -0,10 | -0,10 |
| $se(\hat{\gamma})$ | 0,14 | 0,14 | 0,14 | 0,13 |
| p -Wert | 0,47 | 0,46 | 0,46 | 0,45 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 1). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells mit $H = 8$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Für alle betrachteten Modellvarianten sowie unter beiden Annahmen bezüglich des idiosynkratischen Störterms ε_{th} kann demnach die Nullhypothese der Informationseffizienz erneut eindeutig nicht verworfen werden. Wird also als zugrunde liegende Informationsmenge die Menge der Änderungen der Prognosen gleichbleibender aufeinanderfolgender Prognosehorizonte für verschiedene vergangene Prognoseziele $\{r_{t-1;h-1;h}, r_{t-2;h-1;h}, \dots\}$ (Fall 1) betrachtet, so können auch die Steuerschätzungen anhand des SARIMA(1,0,0)(1,0,1)₄-Modells mit maximalem Prognosehorizont $H = 8$ als **effizient** beurteilt werden. Abermals ergeben sich für die Prognosen anhand des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells die gleichen Testentscheidungen wie für das bevorzugte Modell.¹²

Als Nächstes wird für die Zwei-Jahres-Prognosen die Informationsmenge der vergangenen Änderungen der Prognosen desselben Prognoseziels in Betracht gezogen (Fall 2). Dabei werden zuerst auch wieder die Prognosen des Gesamtsteuerauf-

¹²Tabelle B.15 und Tabelle B.16 im Anhang stellen die entsprechenden Testergebnisse dar.

kommens der BRD anhand des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells (mit $H = 8$) auf traditionelle Art und Weise auf diese zweite Form der Informationseffizienz getestet.

Die Regressionsgleichung (4.31)

$$r_{t;h-1;h} = \gamma r_{t;h+1;h+2} + \delta + \omega_t \quad (\forall \quad t = 1, \dots, T),$$

wird also **separat** für die Prognosehorizonte $h = 2, \dots, 6$ geschätzt, um jeweils die Nullhypothese der Informationseffizienz, $H_0 : \gamma = 0$, zu testen. Erneut werden zur Berechnung der geschätzten Varianz-Kovarianz-Matrix des Störtermvektors ω gemäß Gleichung (4.32) beide Modellannahmen bezüglich des idiosynkratischen Störterms berücksichtigt, also sowohl Annahme A als auch Annahme B. Die Ergebnisse der entsprechenden t-Tests sind in Tabelle 4.14 dargestellt.

Tabelle 4.14.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 2 der Prognosen mit $H = 8$**

| h | $\hat{\gamma}$ | Annahme A | | Annahme B | |
|-----|----------------|--------------------|-----------|--------------------|-----------|
| | | $se(\hat{\gamma})$ | p -Wert | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,04 | 0,15 | 0,81 | 0,13 | 0,77 |
| 3 | -0,04 | 0,16 | 0,81 | 0,14 | 0,77 |
| 4 | -0,01 | 0,16 | 0,93 | 0,13 | 0,92 |
| 5 | 0,02 | 0,17 | 0,89 | 0,14 | 0,87 |
| 6 | 0,05 | 0,12 | 0,70 | 0,11 | 0,65 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2, \dots, 6$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t;h+1;h+2} + \delta + \omega_t$ (Fall 2). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für jedes h sind unter beiden Annahmen die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells mit $H = 8$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Somit kann für die Prognosen mit maximalem Prognosehorizont $H = 8$ also auch für den zweiten Fall des Tests auf Informationseffizienz die Nullhypothese $H_0 : \gamma = 0$ für alle Prognosehorizonte h eindeutig nicht verworfen werden. Dabei spielt es auch in diesem Fall keine Rolle, ob der idiosynkratische Störterm ε_{th} im Modell gelassen wird oder nicht.

Im zweiten Schritt können die Änderungen der Prognosen nun gemäß Gleichung (4.33) und Gleichung (4.34) **gepoolt** werden, um die Regressionsgleichung (4.35)

$$r = \gamma r_{-1} + \delta_d + \omega$$

für alle Prognosehorizonte zusammen zu schätzen. Tabelle 4.15 zeigt die entsprechenden Ergebnisse des Tests auf Informationseffizienz, $H_0 : \gamma = 0$, wiederum sowohl für Variante I (gleich große Verzerrung über alle Prognosehorizonte) als auch Variante II (horizont-spezifische Verzerrungen) jeweils unter Annahme A (Modell mit ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}).

Auch wenn der Test für alle Prognosehorizonte zusammen durchgeführt wird, kann die Nullhypothese der Informationseffizienz eindeutig nicht verworfen werden, unabhängig davon, ob bei der Modellierung horizont-spezifische Verzerrungen zugelassen werden (Variante I) oder nicht (Variante II). Auch die Annahme bezüglich des idiosynkratischen Störterms ε_{th} hat keinen Einfluss auf die Testentscheidung.

Die eigenen Steuerschätzungen anhand des SARIMA(1,0,0)(1,0,1)₄-Modells mit maximalem Prognosehorizont $H = 8$ können demnach also ebenfalls als **effizient** beurteilt werden, wenn als zugrunde liegende Informationsmenge die vergangenen Änderungen der Prognosen desselben Prognoseziels $\{r_{t;h-1+i;h+i}\}$ betrachtet wird. Diese Testergebnisse kommen ebenfalls für die Prognosen anhand des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells zu Stande.¹³

¹³Die Testergebnisse sind Tabelle B.17 und Tabelle B.18 im Anhang zu entnehmen.

Tabelle 4.15.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 2 der Prognosen mit $H = 8$**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | 0,01 | 0,01 | 0,02 | 0,02 |
| $se(\hat{\gamma})$ | 0,13 | 0,12 | 0,12 | 0,12 |
| p -Wert | 0,91 | 0,90 | 0,90 | 0,90 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 2). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells mit $H = 8$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Schließlich werden noch die unverzerrten Drei-Jahres-Prognosen auf Informationseffizienz getestet. Im ersten Schritt wird erneut für die Prognosen des Gesamtsteueraufkommens der BRD anhand des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells mit maximalem Prognosehorizont $H = 12$ zunächst die Regressionsgleichung (4.23) **separat für jeden Prognosehorizont** ($h = 2, \dots, 12$) geschätzt, um jeweils die Nullhypothese der Informationseffizienz, $H_0 : \gamma = 0$, zu testen. Zur Berechnung des Schätzers der Varianz-Kovarianz-Matrix des Störtermvektors ω werden zwar auch für die Drei-Jahres-Prognosen beide Modellannahmen bezüglich des idiosynkratischen Störterms berücksichtigt.

Allerdings ergeben sich unter Annahme A aufgrund des stark negativen Schätzwertes für σ_ε^2 (s. Tabelle 4.5) auch negative Werte für die geschätzten Varianzen des Parametervektors $(\hat{\gamma}, \hat{\delta})$. Dadurch können unter Annahme A keine Standardfehler und p -Werte für $\hat{\gamma}$ angegeben werden. Diese Modellannahme ist jedoch für die Prognosen mit einem maximalen Prognosehorizont von $H = 12$ Quartalen aufgrund des nega-

tiven Varianzschätzers ohnehin als fragwürdig einzustufen. Daher beschränken sich die Ergebnisse der t -Tests auf Informationseffizienz auf das Modell unter Annahme B, bei dem also $\sigma_\varepsilon^2 = 0$ gesetzt wird.

Die entsprechenden Testergebnisse (für Fall 1) sind in Tabelle 4.16 dargestellt. Demnach kann die Nullhypothese der Informationseffizienz unter Annahme B also auch für die Drei-Jahres-Prognosen eindeutig nicht verworfen werden, wenn für jeden Prognosehorizont h separat auf Informationseffizienz getestet wird.

Tabelle 4.16.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 1 der Prognosen mit $H = 12$**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,09 | 0,22 | 0,68 |
| 3 | 0,06 | 0,22 | 0,81 |
| 4 | 0,02 | 0,14 | 0,86 |
| 5 | 0,01 | 0,15 | 0,96 |
| 6 | -0,0005 | 0,15 | 0,997 |
| 7 | -0,10 | 0,15 | 0,53 |
| 8 | -0,15 | 0,11 | 0,20 |
| 9 | -0,14 | 0,12 | 0,24 |
| 10 | -0,11 | 0,11 | 0,34 |
| 11 | -0,08 | 0,11 | 0,47 |
| 12 | -0,09 | 0,09 | 0,33 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2, \dots, 12$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für jedes h sind unter beiden Annahmen die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei, drei bzw. vier Nachkommastellen gerundet. Die Stichprobe umfasst die Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells mit $H = 12$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Im nächsten Schritt werden die Änderungen der SARIMA-Prognosen gemäß Gleichung (4.26) und (4.27) wieder **gepoolt** und die Regressionsgleichung (4.28) für alle Prognosehorizonte zusammen geschätzt, um die Nullhypothese der Informationseffizienz, $H_0 : \gamma = 0$, zu testen. Tabelle 4.17 gibt die entsprechenden Testergebnisse für Variante I (gleich große Verzerrung über alle Prognosehorizonte) sowie Variante II (horizont-spezifische Verzerrungen) jeweils unter Annahme A (Modell mit ε_{th}) und unter Annahme B (Modell ohne ε_{th}) wieder.

Tabelle 4.17.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 1 der Prognosen mit $H = 12$**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | -0,07 | -0,07 | -0,07 | -0,07 |
| $se(\hat{\gamma})$ | 0,11 | 0,12 | 0,11 | 0,12 |
| p -Wert | 0,54 | 0,55 | 0,52 | 0,54 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 1). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells mit $H = 12$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Folglich kann die Nullhypothese der Informationseffizienz also abermals für alle betrachteten Modellvarianten sowie unter beiden Annahmen bezüglich des idiosynkratischen Störterms ε_{th} eindeutig nicht verworfen werden. Die Steuerschätzungen anhand des SARIMA(1,0,0)(1,0,1)₄-Modells erweisen sich also sogar für die Prognosen mit einem maximalen Prognosehorizont von $H = 12$ Quartalen als **effizient**, wenn als zugrunde liegende **Informationsmenge die Änderungen der Prognosen gleichbleibender aufeinanderfolgender Prognosehorizonte für verschiede-**

ne vergangene Prognoseziele $\{r_{t-1;h-1;h}, r_{t-2;h-1;h}, \dots\}$ betrachtet wird. Für die Prognosen anhand des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells (mit $H = 12$) ergeben sich abermals die gleichen Testentscheidungen wie für das bevorzugte Modell.¹⁴

Abschließend wird für die Drei-Jahres-Prognosen die **Informationsmenge der vergangenen Änderungen der Prognosen desselben Prognoseziels** in Betracht gezogen (Fall 2). Erneut werden zuerst die Prognosen des Gesamtsteueraufkommens der BRD anhand des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells (mit $H = 12$) auf traditionelle Art und Weise auf diese zweite Form der Informationseffizienz getestet. Dabei wird wieder die Regressionsgleichung (4.31) **separat für jeden Prognosehorizont** ($h = 2, \dots, 10$) geschätzt, um jeweils die Nullhypothese der Informationseffizienz, $H_0 : \gamma = 0$, zu testen. Zur Berechnung der geschätzten Varianz-Kovarianz-Matrix des Störtermvektors ω werden auch wieder beide Modellvarianten berücksichtigt, also einerseits das Modell mit ε_{th} (Annahme A) sowie andererseits das Modell ohne idiosynkratischen Störterm (Annahme B).

Allerdings ergeben sich unter Annahme A aufgrund des stark negativen Schätzwertes für σ_ε^2 auch in diesem Fall wieder negative Werte für die geschätzten Varianzen des Parametervektors $(\hat{\gamma}, \hat{\delta})$, so dass wiederum unter Annahme A keine Standardfehler und p -Werte für $\hat{\gamma}$ angegeben werden können. Da diese Modellannahme jedoch ohnehin aufgrund des negativen Varianzschätzers als fragwürdig einzustufen ist, beschränken sich die Ergebnisse der t -Tests auf Informationseffizienz erneut auf das Modell unter Annahme B, bei dem also $\sigma_\varepsilon^2 = 0$ gesetzt wird. Die entsprechenden Testergebnisse (für Fall 2) sind in Tabelle 4.18 dargestellt.

Im Gegensatz zum ersten Fall des Tests auf Informationseffizienz kann nun die Nullhypothese $H_0 : \gamma = 0$ lediglich bis zum Prognosehorizont $h = 5$ auf dem 5%-Signifikanzniveau nicht verworfen werden. Ab einem Prognosehorizont von $h = 6$ Quartalen implizieren die Testergebnisse dagegen, dass die vergangenen Änderungen der Prognosen desselben Prognoseziels Informationen enthalten, anhand derer

¹⁴Tabelle B.19 und Tabelle B.20 im Anhang stellen die entsprechenden Testergebnisse dar.

Tabelle 4.18.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 2 der Prognosen mit $H = 12$**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,24 | 0,14 | 0,088 |
| 3 | 0,20 | 0,15 | 0,193 |
| 4 | 0,16 | 0,13 | 0,232 |
| 5 | 0,21 | 0,14 | 0,154 |
| 6 | 0,33 | 0,11 | 0,005 |
| 7 | 0,26 | 0,12 | 0,035 |
| 8 | 0,26 | 0,11 | 0,022 |
| 9 | 0,25 | 0,11 | 0,025 |
| 10 | 0,31 | 0,09 | 0,001 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2, \dots, 10$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t;h+1;h+2} + \delta + \omega_t$ (Fall 2). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für jedes h sind unter beiden Annahmen die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei bzw. drei Nachkommastellen gerundet. Die Stichprobe umfasst die Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells mit $H = 12$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

die aktuelle Änderung der Prognosen vorhergesagt werden kann. Die Wahl der zu Grunde liegenden Informationsmenge bestimmt für die Prognosen mit einem maximalen Prognosehorizont von $H = 12$ Quartalen also maßgeblich, ob die Prognosen als effizient betrachtet werden können.

Im zweiten Schritt werden die Änderungen der Prognosen wieder gemäß Gleichung (4.33) und Gleichung (4.34) **gepoolt**, um die Regressionsgleichung (4.35) für alle Prognosehorizonte zusammen zu schätzen. Tabelle 4.19 zeigt die entsprechenden Ergebnisse des Tests auf Informationseffizienz, $H_0 : \gamma = 0$, wiederum sowohl für Variante I (gleich große Verzerrung über alle Prognosehorizonte) als auch Variante II

(horizont-spezifische Verzerrungen) jeweils unter Annahme A (Modell mit ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}).

Tabelle 4.19.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 2 der Prognosen mit $H = 12$**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | 0,26 | 0,26 | 0,26 | 0,26 |
| $se(\hat{\gamma})$ | 0,10 | 0,11 | 0,10 | 0,11 |
| p -Wert | 0,01 | 0,02 | 0,01 | 0,02 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 2). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells mit $H = 12$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Wird der Test für alle Prognosehorizonte zusammen durchgeführt, kann die Nullhypothese der Informationseffizienz auf dem 5%-Signifikanzniveau verworfen werden, unabhängig davon, ob bei der Modellierung horizont-spezifische Verzerrungen zugelassen werden (Variante I) oder nicht (Variante II). Auch die Annahme bezüglich des idiosynkratischen Störterms ε_{th} hat keinen Einfluss auf die Testentscheidung. Somit müssen die eigenen Steuerschätzungen anhand des SARIMA(1,0,0)(1,0,1)₄-Modells mit einem **maximalen Prognosehorizont von $H = 12$ als ineffizient** beurteilt werden, wenn als zugrunde liegende **Informationsmenge die vergangenen Änderungen der Prognosen desselben Prognoseziels** $\{r_{t,h-1+i,h+i}\}$ betrachtet wird. Diese Testergebnisse kommen ebenfalls für die Prognosen anhand des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells zu Stande.¹⁵

¹⁵Die Testergebnisse sind Tabelle B.21 und Tabelle B.22 im Anhang zu entnehmen.

4.2.3. Fazit zur Prognoserationalität der SARIMA-Schätzungen

Zusammenfassend lässt sich festhalten, dass sowohl die Prognosen des Gesamtsteuer-aufkommens der BRD anhand des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells als auch diejenigen anhand des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells bei einem maximalen Prognosehorizont von $H = 4$, $H = 8$ sowie $H = 12$ **Quartale für alle betrachteten Modellvarianten sowie unter allen Annahmen unverzerrt** sind.

Wird der maximale Prognosehorizont allerdings auf $H = 16$ Quartale erhöht, resultieren keine robusten Testergebnisse. Wird in der Alternativhypothese unterstellt, dass die Verzerrung für alle Prognosehorizonte gleich groß sein muss (Variante I), kann die Nullhypothese der Unverzerrtheit der Prognosen eindeutig nicht verworfen werden. Lässt man in der Alternativhypothese allerdings horizont-spezifische Verzerrungen zu (Variante II), so erweisen sich die untersuchten Prognosen lediglich im Modell mit idiosynkratischem Störterm (Annahme A) als unverzerrt. Dabei resultiert allerdings ein negativer Schätzwert für σ_ε^2 . Wird $\sigma_\varepsilon^2 = 0$ gesetzt (Annahme B) muss dagegen von verzerrten Vier-Jahres-Prognosen ausgegangen werden.

Des Weiteren erweisen sich die Prognosen mit einem maximalen Prognosehorizont von $H = 4$ **und** $H = 8$ für beide untersuchten Informationsmengen und unter allen betrachteten Modellannahmen als **effizient**.

Wird der maximale Prognosehorizont jedoch auf $H = 12$ erhöht, hat die Wahl der zu Grunde liegenden Informationsmenge einen maßgeblichen Einfluss darauf, ob die Prognosen als effizient betrachtet werden können. So implizieren die Testergebnisse, dass die Drei-Jahres-Prognosen auch dann noch **effizient** sind, wenn die **Änderungen der Prognosen gleichbleibender aufeinanderfolgender Prognosehorizonte für verschiedene vergangene Prognoseziele** $\{r_{t-1;h-1;h}, r_{t-2;h-1;h}, \dots\}$ (Fall 1) als zugrunde liegende Informationsmenge berücksichtigt wird.

Wird aber als Informationsmenge die vergangenen Änderungen der Prognosen desselben Prognoseziels $\{r_{t,h-1+i,h+i}\}$ (Fall 2) betrachtet und wird dabei für jeden Prognosehorizont einzeln auf Effizienz getestet, kann die Nullhypothese der Informationseffizienz lediglich bis zum Prognosehorizont $h = 5$ auf dem 5%-Signifikanzniveau nicht verworfen werden. Ab dem Prognosehorizont $h = 6$ sowie auch im gepoolten Vorgehen implizieren die Testergebnisse dagegen, dass die SARIMA-Prognosen mit einem **maximalen Prognosehorizont von $H = 12$ als ineffizient** beurteilt werden müssen, wenn als zugrunde liegende **Informationsmenge die vergangenen Änderungen der Prognosen desselben Prognoseziels $\{r_{t,h-1+i,h+i}\}$** berücksichtigt wird.

Somit erweisen sich die eigenen SARIMA-Prognosen des Gesamtsteueraufkommens der BRD bei einem maximalen Prognosehorizont von $H = 4$ **sowie $H = 8$ als unverzerrt und effizient**, so dass sie im Sinne der beiden betrachteten Informationsmengen auch als **rational** bezeichnet werden können. Wird der maximale Prognosehorizont jedoch auf $H = 12$ erhöht, können die unverzerrten SARIMA-Prognosen nur noch im Sinne der zuerst berücksichtigten Informationsmenge als effizient und damit auch als rational beurteilt werden.

Im folgenden Kapitel wird nun die Frage untersucht, inwieweit die offiziellen AKS-Prognosen ebenfalls unverzerrt und effizient sind, wenn sie anhand des vorgestellten gepoolten Ansatzes gemäß Clements et al. (2007) getestet werden.

4.3. Untersuchung der Prognoserationalität der AKS-Schätzungen

Die bisherigen empirischen Analysen der Prognosen des Arbeitskreises „Steuerschätzungen“ auf Unverzerrtheit von Becker und Buettner (2007), Büttner und Kauder (2008b) sowie Lehmann (2010) und diejenige auf Informationseffizienz von Becker und Buettner (2007) kommen zu dem Ergebnis, dass die AKS-Schätzungen unverzerrt und in den meisten der untersuchten Fälle auch effizient sind. Allerdings folgen diese Autoren dem traditionellen Ansatz, die Prognosen separat für verschiedene Prognosehorizonte zu testen. Im Folgenden werden die AKS-Prognosen nun erstmalig gemäß des gepoolten Ansatzes von Clements et al. (2007) auf ihre Prognoserationalität hin untersucht.

Wie bereits erwähnt überprüft der Arbeitskreis „Steuerschätzungen“ im November lediglich die Prognosen für das laufende Jahr und überarbeitet die Prognosen für das kommende Jahr, um eine möglichst aktuelle Grundlage für die Aufstellung des Bundeshaushaltes zu liefern (vgl. Bundesministerium der Finanzen (2005), S. 11). Zu diesem Zeitpunkt werden allerdings keine neuen bzw. überarbeiteten Steuerschätzungen für den mittelfristigen Zeitraum erstellt, so dass sich die weitere Analyse auf die **Mai-Prognosen des AKS** konzentriert. Denn im Mai erfolgt die Steuerschätzung sowohl für das laufende Jahr, was im Folgenden mit $h = 1$ bezeichnet wird, als auch für die vier Folgejahre ($h = 2, \dots, 5$). Somit stellt sich für die Mai-Schätzungen die Frage, ob sich die Prognosen auch dann als rational erweisen, wenn die Schätzungen für alle fünf Prognosehorizonte gemeinsam auf Unverzerrtheit und Informationseffizienz getestet werden.

Als Datengrundlage werden die auf der Internetseite des Bundesministeriums der Finanzen veröffentlichten Ergebnisse des Arbeitskreises „Steuerschätzungen“ von 1971 bis zum Jahr 2010 verwendet (vgl. Bundesministerium der Finanzen (2011a)). Dabei gilt es allerdings zu beachten, dass in den Jahren vor 1986 die Sitzung zur Erstellung der mittelfristigen Prognosen nicht immer im Mai stattfand. Während

beispielsweise 1971, 1972 und 1978 die entsprechenden Sitzungen bereits im Februar abgehalten wurden, erstellte der AKS die mittelfristigen Prognosen in den Jahren 1973, 1975 und 1977 erst im August. Als extreme Ausnahme muss darüber hinaus das Jahr 1976 betrachtet werden, als die entsprechende Sitzung erst im Dezember stattfand. Zumindest die 1976 erstellte Prognose für das laufende Jahr ($h = 1$) sowie diejenige für das darauf folgende Jahr ($h = 2$) lassen sich daher kaum mit den Prognosen der Prognosehorizonte $h = 1$ und $h = 2$ anderer Jahre vergleichen.

Möchte man also die Prognosen des Ausnahmejahres 1976 komplett vermeiden, können erst die Prognosefehler von 1981 an untersucht werden, da zum Prognosezeitpunkt Dezember 1976 mittelfristige Prognosen bis einschließlich 1980 ($h = 5$) vom AKS erstellt wurden. Damit umfasst die vorläufige **Prognosematrix** zur Untersuchung der Prognoserationalität gemäß Clements et al. (2007) die **AKS-Schätzungen für $T = 30$ Prognoseziele**, nämlich für die Jahre von 1981 bis 2010, für jeweils $h = 1$ bis $h = 5$ **Prognosehorizonte**. Alle Prognosewerte des vom AKS vorhergesagten Gesamtsteueraufkommens der BRD wurden dabei ebenfalls in Milliarden Euro umgerechnet.

4.3.1. Test auf Unverzerrtheit der AKS-Prognosen

Zunächst werden die Prognosen des Arbeitskreises „Steuerschätzungen“ für unterschiedliche maximale Prognosehorizonte auf Unverzerrtheit getestet, wobei jeweils sowohl Variante I als auch Variante II unter beiden möglichen Annahmen bezüglich des idiosynkratischen Störterms berücksichtigt werden. Für die AKS-Prognosen des Gesamtsteueraufkommens von 1981 bis 2010 mit einem **maximalen Prognosehorizont von vier Jahren**, also mit $H = 5$, ergeben sich für die **Variante I**, also mit einer gleich großen Verzerrung über alle Prognosehorizonte als Alternativhypothese, die Testergebnisse in Tabelle 4.20.

Tabelle 4.20.: **Ergebnisse des t -Tests für Variante I mit $H = 5$**

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -10,99 | -10,99 |
| $se(\hat{\alpha})$ | 13,10 | 11,64 |
| p -Wert | 0,40 | 0,35 |
| $\hat{\sigma}_\varepsilon^2$ | (-506,12) | – |
| $\hat{\sigma}_u^2$ | 611,17 | 473,13 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, \dots, 5$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

Unter beiden Annahmen ergibt sich als Schätzwert für α eine insignifikante negative Verzerrung der Prognose von ungefähr $-10,99$ Milliarden Euro. Durchschnittlich überschätzte der AKS das Gesamtsteueraufkommen zwischen 1981 und 2010 also um knapp 11 Milliarden Euro pro Jahr. Da sich unter Annahme A ein negativer Schätzwert für die Varianz σ_ε^2 ergibt, erscheint diese Modellannahme eher fraglich. Allerdings kann gemäß des jeweiligen p -Wertes die Nullhypothese der Unverzerrtheit sowohl unter Annahme A als auch unter Annahme B eindeutig nicht verworfen werden.

Für **Variante II**, bei der für jeden Prognosehorizont h eine horizont-spezifische Verzerrung in der Alternativhypothese zugelassen wird, sind die Regressionsergebnisse in Tabelle 4.21 zusammengestellt.

Tabelle 4.21.: **Regressionsergebnisse der horizont-spezifischen Verzerrungen der AKS-Prognosen mit $H = 5$**

| h | Annahme A | | | Annahme B | |
|-----|------------------|----------------------|-----------|----------------------|-----------|
| | $\hat{\alpha}_h$ | $se(\hat{\alpha}_h)$ | p -Wert | $se(\hat{\alpha}_h)$ | p -Wert |
| 1 | 0,97 | 1,87 | 0,60 | 3,97 | 0,81 |
| 2 | -4,00 | 7,95 | 0,62 | 7,88 | 0,61 |
| 3 | -11,34 | 12,69 | 0,37 | 11,74 | 0,34 |
| 4 | -16,90 | 17,19 | 0,33 | 15,55 | 0,28 |
| 5 | -23,67 | 21,57 | 0,27 | 19,32 | 0,22 |

Regressionsergebnisse der horizont-spezifischen Verzerrungen beim Test auf Unverzerrtheit der AKS-Prognosen mit $H = 5$, anhand der Regressionsgleichung $e = (i_T \otimes I_H) \alpha^H + v$. Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für jedes h sind unter beiden Annahmen die Werte der geschätzten horizont-spezifischen Verzerrungen $\hat{\alpha}_h$, ihrer Standardfehler $se(\hat{\alpha}_h)$ sowie die entsprechenden p -Werte der einzelnen t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

Die geschätzten Verzerrungen $\hat{\alpha}_h$ nehmen also auch für die AKS-Prognosen wie erwartet mit steigendem Prognosehorizont zu. Während der Arbeitskreis „Steuer-schätzungen“ das jährliche Gesamtsteueraufkommen des laufenden Jahres durchschnittlich noch um eine knappe Milliarde Euro unterschätzt, liegt für das Steueraufkommen des nächsten Jahres bereits eine durchschnittliche Überschätzung von ungefähr 4 Milliarden Euro vor. Die mittelfristigen Vier-Jahres-Prognosen weisen sogar eine durchschnittliche Verzerrung von circa 23,7 Milliarden Euro auf. Allerdings erweisen sich alle Verzerrungen (einzeln anhand des t -Tests getestet) sowohl unter Annahme A als auch unter Annahme B als insignifikant. Werden hingegen alle geschätzten Verzerrungen gemeinsam auf Signifikanz getestet, ergeben sich die in Tabelle 4.22 dargestellten Ergebnisse des F -Tests.

Tabelle 4.22.: **Ergebnisse des F -Tests für Variante II mit $H = 5$**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,24 | 8,77 |
| p -Wert | 0,94 | 0,0000 |
| $\hat{\sigma}_\varepsilon^2$ | (-577,15) | – |
| $\hat{\sigma}_u^2$ | 609,00 | 451,59 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, \dots, 5$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

Für den F -Test wirkt sich also die Annahme bezüglich des idiosynkratischen Störterms entscheidend auf das Testergebnis aus. Wird ε_{th} als prognose-spezifischer Fehler im Modell gelassen, ergibt sich zwar wiederum ein negativer Schätzwert für die Varianz σ_ε^2 , die Nullhypothese der Unverzerrtheit kann dann aber weiterhin nicht verworfen werden. Im plausibleren Modell ohne ε_{th} , in dem $\sigma_\varepsilon^2 = 0$ gesetzt wird, ändert sich dagegen das Testergebnis entscheidend, so dass für die mittelfristigen AKS-Prognosen mit maximalem Prognosehorizont $H = 5$ nicht uneingeschränkt davon ausgegangen werden kann, dass sie unverzerrt sind.

Die Prognosen der als Alternative vorgeschlagenen SARIMA-Modelle kommen jedoch bei einem maximalen Prognosehorizont von vier Jahren ($H = 16$) ebenfalls auf diese Testergebnisse, während sie sich bei einem maximalen Prognosehorizont von bis zu drei Jahren ($H = 4, 8, 12$) noch als unverzerrt erweisen. Daher werden im nächsten Schritt auch die AKS-Prognosen mit einem **maximalen Prognosehorizont von drei Jahren**, also mit $H = 4$, auf Unverzerrtheit getestet. Für die AKS-Prognosen des Gesamtsteueraufkommens von 1981 bis 2010 mit $H = 4$, erge-

ben sich für die Variante I die Testergebnisse in Tabelle 4.23. Für Variante II sind die Ergebnisse des F -Tests in Tabelle 4.24 zusammengestellt.¹⁶

Tabelle 4.23.: **Ergebnisse des t -Tests für Variante I mit $H = 4$**

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -7,82 | -7,82 |
| $se(\hat{\alpha})$ | 11,07 | 9,41 |
| p -Wert | 0,48 | 0,41 |
| $\hat{\sigma}_\varepsilon^2$ | (-582,21) | - |
| $\hat{\sigma}_u^2$ | 634,53 | 440,46 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, \dots, 4$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

Während die Ergebnisse des t -Tests in Variante I also auch für die Prognosen mit maximalem Prognosehorizont $H = 4$ unverzerrte Prognosen implizieren, kann die Nullhypothese der Unverzerrtheit des F -Tests in Variante II eindeutig verworfen werden, wenn aufgrund des negativen Schätzwertes für σ_ε^2 unter Annahme A kein idiosynkratischer Störterm im Modell berücksichtigt wird, also $\sigma_\varepsilon^2 = 0$ gesetzt wird (Annahme B). Somit hängt die Testentscheidung also auch für die AKS-Prognosen mit maximalem Prognosehorizont von drei Jahren wesentlich davon ab, ob horizontspezifische Verzerrungen als Alternative berücksichtigt werden und dabei der Störterm ε_{th} aus dem Modell genommen wird. Diese Testentscheidungen ergeben sich außerdem auch dann noch, wenn der maximale Prognosehorizont auf zwei Jahre ($H = 3$) beziehungsweise ein Jahr ($H = 2$) gesenkt wird. Die Tabellen C.2 bis C.5

¹⁶Die detaillierten Regressionsergebnisse sind in Tabelle C.1 im Anhang dargestellt

Tabelle 4.24.: **Ergebnisse des F -Tests für Variante II mit $H = 4$**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,23 | 11,35 |
| p -Wert | 0,92 | 0,0000 |
| $\hat{\sigma}_\varepsilon^2$ | (-625,47) | – |
| $\hat{\sigma}_u^2$ | 633,16 | 424,67 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, \dots, 4$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

im Anhang stellen die jeweiligen Testergebnisse dar. Die Prognosen der als Alternative vorgeschlagenen SARIMA-Modelle erweisen sich dagegen bei einem maximalen Prognosehorizont von bis zu drei Jahren ($H = 4, 8, 12$ Quartale) für alle Modellvarianten als unverzerrt.

Da bisher die AKS-Prognosen für die komplette Zeitspanne von 1981 bis 2010 auf Unverzerrtheit getestet wurden, fällt die deutsche Wiedervereinigung in den Untersuchungszeitraum. In der Literatur wird deshalb die Prognose für das Jahr 1990 oftmals weggelassen (vgl. z. B. Gebhardt (2001), S. 133 oder Lehmann (2010), S. 35). Daher wird nun außerdem die Prognosematrix auf die Prognoseziele 1980 bis 1989 sowie 1991 bis 2010 (weiterhin $T = 30$ Prognoseziele), für jeweils $h = 1$ bis $h = 5$ Prognosehorizonte, angepasst. Dadurch geht in diesem Fall zwar auch die 4-Jahres-Prognose des Ausnahmejahres 1976 mit in die Analyse ein. Allerdings dürfte es für diesen mittelfristigen Prognosehorizont kaum eine Rolle spielen, ob die Prognosen bereits im Mai oder erst im Dezember des Jahres 1976 erstellt wurden. Die Testergebnisse verändern sich jedoch auch nicht, wenn das **Wiedervereinigungsjahr ausgeschlossen** wird. Die Tabellen C.6 bis C.13 im Anhang geben die entsprechenden

Testergebnisse für die verschiedenen maximalen Prognosehorizonte ($H = 2, 3, 4, 5$) wieder. Somit spielt es also keine Rolle, ob das Jahr 1990 als Prognoseziel eliminiert wird oder nicht.

Der Arbeitskreis „Steuerschätzungen“ sollte erstmals im Mai 1991 die mittelfristige Prognose für das wiedervereinigte Deutschland erstellen, weist in der auf der Internetseite des BMF veröffentlichten Pressemitteilung vom 16. Mai 1991 zu den Ergebnissen der 92. Sitzung des AKS jedoch darauf hin, dass die Datenbasis noch unzureichend sei. Daher wäre es auch denkbar, das Prognosejahr 1991 aus der Analyse herauszunehmen. Möchte man jedoch alle Prognosen, die 1991 unter mangelhafter Datengrundlage erstellt wurden, eliminieren, müssen die **Prognoseziele 1991 bis 1995 weggelassen** werden. Damit sich der Stichprobenumfang nicht verkleinert, müssen also die Prognoseziele von 1976 an untersucht werden, als der Prognosezeitpunkt wie oben beschrieben noch sehr stark schwankte. Doch auch in diesem Fall werden die bisherigen Testergebnisse bestätigt. Die jeweiligen Testergebnisse sind für die maximalen Prognosehorizonte $H = 2, 3, 4, 5$ den Tabellen C.14 bis C.21 im Anhang zu entnehmen.

Als Kompromisslösung kann abschließend auch noch die Variante betrachtet werden, bei der lediglich **das Jahr 1991 ausgeschlossen** wird. Die Prognosematrix umfasst hierbei also die Prognoseziele 1980 bis 1990 sowie 1992 bis 2010, für ebenfalls jeweils $h = 1$ bis $h = 5$ Prognosehorizonte. Doch auch in diesem Fall gibt es keine veränderten Testentscheidungen. Die Tabellen C.22 bis C.29 im Anhang stellen die entsprechenden Testergebnisse für die verschiedenen maximalen Prognosehorizonte ($H = 2, 3, 4, 5$) dar.

Wird in der Alternativhypothese also unterstellt, dass die Verzerrung für alle Prognosehorizonte gleich groß sein muss (Variante I), kann die Nullhypothese der Unverzerrtheit der Prognosen in allen betrachteten Fällen eindeutig nicht verworfen werden. Lässt man in der Alternativhypothese hingegen horizont-spezifische Verzerrungen zu (Variante II), dann erweisen sich die untersuchten Prognosen lediglich

im Modell mit idiosynkratischem Störterm (Annahme A) als unverzerrt. Dabei resultiert allerdings ein negativer Schätzwert für σ_ε^2 . Wird $\sigma_\varepsilon^2 = 0$ gesetzt (Annahme B) muss jedoch von **verzerrten Prognosen** ausgegangen werden. Dies widerspricht den bisherigen empirischen Ergebnissen von Becker und Buettner (2007), Büttner und Kauder (2008b) und Lehmann (2010), die den traditionellen Test auf Unverzerrtheit aber lediglich einzeln für bestimmte Prognosehorizonte durchführen. Die Prognosen der beiden als Alternative vorgeschlagenen SARIMA-Modelle erweisen sich dagegen bei einem maximalen Prognosehorizont von bis zu drei Jahren ($H = 4, 8, 12$) unter Anwendung des gepoolten Ansatzes für alle Modellvarianten als unverzerrt.

Als Nächstes werden die offiziellen Prognosen des Arbeitskreises „Steuerschätzungen“ nun auf ihre Informationseffizienz gemäß Clements et al. (2007) getestet.

4.3.2. Test auf Informationseffizienz der AKS-Prognosen

Zunächst wird wieder die **ursprüngliche Prognosematrix** mit den AKS-Schätzungen des Gesamtsteueraufkommens für die Jahre 1981 bis 2010 mit einem maximalen Prognosehorizont von vier Jahren, also mit $H = 5$, zur Analyse der Informationseffizienz verwendet. Im ersten Schritt wird auch für die AKS-Prognosen auf **traditionelle Art und Weise** die Regressionsgleichung (4.23)

$$r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t \quad \forall \quad t = 2, \dots, T$$

separat für jeden Prognosehorizont ($h = 2, \dots, 5$) geschätzt, um jeweils die Nullhypothese der Informationseffizienz, $H_0 : \gamma = 0$, wie in Kapitel 4.1.2 beschrieben, zu testen. Zur Berechnung des Schätzers der Varianz-Kovarianz-Matrix des Störtermvektors ω gemäß Gleichung (4.25)

$$Cov[\omega\omega'] = (2\sigma_\varepsilon^2 + \sigma_u^2)I_{T-1}.$$

werden zwar wieder beide Modellvarianten berücksichtigt, also einerseits das Modell mit ε_{th} (Annahme A) sowie andererseits das Modell ohne idiosynkratischen Störterm (Annahme B).

Allerdings ergeben sich unter Annahme A aufgrund des stark negativen Schätzwertes für σ_ε^2 (s. Tabelle 4.20) auch negative Werte für die geschätzten Varianzen des Parametervektors $(\hat{\gamma}, \hat{\delta})$. Dadurch können unter Annahme A keine Standardfehler und p -Werte für $\hat{\gamma}$ angegeben werden. Diese Modellannahme ist jedoch aufgrund des negativen Varianzschätzers ohnehin als fragwürdig einzustufen. Daher beschränken sich die Ergebnisse der t -Tests auf Informationseffizienz auf das Modell unter Annahme B, bei dem also $\sigma_\varepsilon^2 = 0$ gesetzt wird. Die entsprechenden Testergebnisse (für Fall 1) sind in Tabelle 4.25 dargestellt.

Tabelle 4.25.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 5$**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,29 | 0,22 | 0,20 |
| 3 | 0,32 | 0,20 | 0,11 |
| 4 | 0,32 | 0,17 | 0,08 |
| 5 | 0,39 | 0,18 | 0,04 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2, \dots, 5$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für jedes h sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 5$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

Unter Annahme B kann die Nullhypothese der Informationseffizienz bis zu einem Prognosehorizont von drei Jahren ($h = 4$) auf dem 5%-Signifikanzniveau nicht verworfen werden. Für den mittelfristigen Prognosehorizont von vier Jahren ($h = 5$) implizieren die Testergebnisse dagegen, dass die Änderung der Prognosen mit aufeinanderfolgenden Prognosehorizonten (also $h = 4$ und $h = 5$) für ein bestimmtes Prognoseziel Informationen enthält, anhand derer die Änderung der Prognosen mit denselben Prognosehorizonten für das darauf folgende Prognoseziel vorhergesagt werden kann. Somit ist für die offiziellen Prognosen des AKS also lediglich für die verhältnismäßig lange Schätzdistanz von vier Jahren ein **schwacher Glättungseffekt** erkennbar.

Im nächsten Schritt werden die Änderungen der AKS-Prognosen gemäß der Gleichung (4.26) und der Gleichung (4.27) **gepoolt** und die Regressionsgleichung (4.28)

$$r = \gamma r_{-1} + \delta_d + \omega$$

für alle Prognosehorizonte zusammen geschätzt, um die Nullhypothese der Informationseffizienz, $H_0 : \gamma = 0$, zu testen. Tabelle 4.26 gibt die entsprechenden Testergebnisse für Variante I (gleich große Verzerrung über alle Prognosehorizonte) sowie Variante II (horizont-spezifische Verzerrungen) jeweils unter Annahme A (Modell mit ε_{th}) und unter Annahme B (Modell ohne ε_{th}) wieder.

Demnach kann die Nullhypothese der Informationseffizienz also unter Annahme A auf dem 5%-Signifikanzniveau noch verworfen werden, während sie unter Annahme B bei einem Signifikanzniveau von fünf Prozent gerade nicht mehr verworfen werden kann. Dieses Testergebnis ist unabhängig davon, ob bei der Modellierung horizont-spezifische Verzerrungen zugelassen werden oder nicht. Da sich unter Annahme A allerdings ein negativer Schätzwert für die Varianz des idiosynkratischen Störterms ε_{th} ergibt, scheint die Modellierung unter Annahme B, bei der $\sigma_\varepsilon^2 = 0$ gesetzt wird, plausibler.

Tabelle 4.26.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 5$**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | 0,33 | 0,33 | 0,33 | 0,33 |
| $se(\hat{\gamma})$ | 0,16 | 0,18 | 0,16 | 0,18 |
| p -Wert | 0,04 | 0,07 | 0,04 | 0,06 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 1). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 5$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

Somit können die Steuerschätzungen des AKS bei einem maximalen Prognosehorizont von $H = 5$ und einem Signifikanzniveau von fünf Prozent also als **effizient** beurteilt werden, wenn als zugrunde liegende **Informationsmenge die Änderungen der Prognosen gleichbleibender aufeinanderfolgender Prognosehorizonte für verschiedene vergangene Prognoseziele** $\{r_{t-1;h-1;h}, r_{t-2;h-1;h}, \dots\}$ betrachtet wird. Geht man allerdings lediglich von einem Signifikanzniveau von zehn Prozent aus, erweisen sich die AKS-Prognosen in diesem Fall als ineffizient.

Auch die AKS-Prognosen mit einem kleineren maximalen Prognosehorizont ($H = 4$, $H = 3$ sowie $H = 2$) implizieren diese Testentscheidungen. Die Tabellen C.30 bis C.34 im Anhang stellen die entsprechenden Testergebnisse dar. Außerdem spielt es auch keine Rolle, ob das Wiedervereinigungsjahr 1990 als Prognoseziel weggelassen wird oder nicht. Die jeweiligen Testergebnisse sind für die verschiedenen maximalen Prognosehorizonte ($H = 5, 4, 3, 2$) den Tabellen C.35 bis C.41 im Anhang zu entnehmen. Wird das Prognoseziel 1991 eliminiert, erhöhen sich die entsprechenden p -Werte zwar leicht. Dennoch lassen sich ebenfalls keine wesentlichen Änderungen

der Testentscheidungen beobachten. Die Tabellen C.42 bis C.48 im Anhang geben die diesbezüglichen Testergebnisse für die verschiedenen maximalen Prognosehorizonte ($H = 5, 4, 3, 2$) wieder.

Werden jedoch alle Prognosen, die 1991 unter mangelhafter Datengrundlage erstellt wurden, aus der Analyse ausgeschlossen, erweisen sich die AKS-Prognosen selbst für den mittelfristigen Prognosehorizont von vier Jahren ($h = 5$) sowohl unter Annahme A als auch unter Annahme B als effizient, wenn die Nullhypothese der Informationseffizienz gemäß Fall 1 separat für jeden Prognosehorizont getestet wird. Tabelle 4.27 stellt die entsprechenden Ergebnisse der separaten t -Tests für die AKS-Prognosen mit einem maximalen Prognosehorizont von $H = 5$ dar.

Tabelle 4.27.: Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 5$ ohne die Jahre 1991 bis 1995

| h | $\hat{\gamma}$ | Annahme A | | Annahme B | |
|-----|----------------|--------------------|-----------|--------------------|-----------|
| | | $se(\hat{\gamma})$ | p -Wert | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,17 | 0,13 | 0,20 | 0,20 | 0,41 |
| 3 | 0,21 | 0,12 | 0,10 | 0,19 | 0,28 |
| 4 | 0,17 | 0,11 | 0,14 | 0,17 | 0,33 |
| 5 | 0,18 | 0,12 | 0,14 | 0,18 | 0,33 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2, \dots, 5$, anhand der Regressionsgleichung: $r_{t,h-1;h} = \gamma r_{t-1,h-1;h} + \delta + \omega_t$ (Fall 1). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für jedes h sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 5$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1976 bis 1990 sowie 1996 bis 2010.

Ebenso kann die Nullhypothese der Informationseffizienz klar verworfen werden, wenn der Test für alle Prognosehorizonte zusammen durchgeführt wird. Das Testergebnis ist dabei davon unabhängig, ob bei der Modellierung horizont-spezifische

Verzerrungen zugelassen werden oder nicht. Auch die Annahme bezüglich des idiosynkratischen Störterms ε_{th} hat keinen Einfluss auf die Testentscheidung. Die diesbezüglichen Testergebnisse sind in Tabelle 4.28 dargestellt.

Tabelle 4.28.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 5$ ohne die Jahre 1991 bis 1995**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | 0,19 | 0,19 | 0,18 | 0,18 |
| $se(\hat{\gamma})$ | 0,16 | 0,17 | 0,14 | 0,16 |
| p -Wert | 0,25 | 0,27 | 0,21 | 0,26 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 1). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 5$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1976 bis 1990 sowie 1996 bis 2010.

Die Prognosen des Arbeitskreises „Steuerschätzungen“ **für die Jahre 1976 bis 1990 sowie 1996 bis 2010** können demnach also uneingeschränkt als **effizient** beurteilt werden, wenn als **zugrunde liegende Informationsmenge die Änderungen der Prognosen gleichbleibender aufeinanderfolgender Prognosehorizonte für verschiedene vergangene Prognoseziele** $\{r_{t-1;h-1;h}, r_{t-2;h-1;h}, \dots\}$ betrachtet wird. Auch bei einem kleineren maximalen Prognosehorizont ($H = 4$, $H = 3$ sowie $H = 2$) ergeben sich verbesserte Testergebnisse bezüglich dieser ersten Form der Informationseffizienz, wenn alle Prognosen des AKS, die 1991 unter mangelhafter Datengrundlage erstellt wurden, aus der Analyse ausgeschlossen werden. Die Tabellen C.49 bis C.53 im Anhang stellen die entsprechenden Testergebnisse dar.

Als Nächstes soll die **Informationsmenge der vergangenen Änderungen der Prognosen desselben Prognoseziels** in Betracht gezogen werden (Fall 2). Erneut wird zuerst wieder die **ursprüngliche Prognosematrix** mit den AKS-Schätzungen des Gesamtsteueraufkommens für die Jahre 1981 bis 2010 mit einem maximalen Prognosehorizont von vier Jahren, also mit $H = 5$, auf **traditionelle Art und Weise** auf diese zweite Form der Informationseffizienz getestet. Dabei wird die Regressionsgleichung (4.31)

$$r_{t;h-1;h} = \gamma r_{t;h+1;h+2} + \delta + \omega_t \quad (\forall \quad t = 1, \dots, T),$$

separat für jeden Prognosehorizont ($h = 2$ und $h = 3$) geschätzt, um jeweils die Nullhypothese der Informationseffizienz, $H_0 : \gamma = 0$, zu testen. Zur Berechnung der geschätzten Varianz-Kovarianz-Matrix des Störtermvektors ω gemäß Gleichung (4.32) werden auch wieder beide Modellvarianten berücksichtigt, also einerseits das Modell mit ε_{th} (Annahme A) sowie andererseits das Modell ohne idiosynkratischen Störterm (Annahme B).

Allerdings ergeben sich unter Annahme A aufgrund des stark negativen Schätzwertes für σ_ε^2 (s. Tabelle 4.20) auch in diesem Fall wieder negative Werte für die geschätzten Varianzen des Parametervektors $(\hat{\gamma}, \hat{\delta})$, so dass wiederum unter Annahme A keine Standardfehler und p -Werte für $\hat{\gamma}$ angegeben werden können. Da diese Modellannahme jedoch ohnehin aufgrund des negativen Varianzschätzers als fragwürdig einzustufen ist, beschränken sich die Ergebnisse der t -Tests auf Informationseffizienz erneut auf das Modell unter Annahme B, bei dem also $\sigma_\varepsilon^2 = 0$ gesetzt wird. Die entsprechenden Testergebnisse (für Fall 2) sind in Tabelle 4.29 dargestellt.

Unter Annahme B kann demnach die Nullhypothese $H_0 : \gamma = 0$ für beide Prognosehorizonte h auf dem 5%-Signifikanzniveau nicht verworfen werden. Dieses Testergebnis ergibt sich außerdem für die AKS-Prognosen mit maximalem Prognosehorizont $H = 4$.¹⁷

¹⁷Das entsprechende Testergebnis ist in Tabelle C.54 im Anhang dargestellt.

Tabelle 4.29.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 2 der AKS-Prognosen mit $H = 5$**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | -0,12 | 0,17 | 0,48 |
| 3 | -0,27 | 0,16 | 0,10 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2$ und $h = 3$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t;h+1;h+2} + \delta + \omega_t$ (Fall 2). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für beide Prognosehorizonte h sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 5$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

Im zweiten Schritt werden die Änderungen der Prognosen gemäß Gleichung (4.33) und Gleichung (4.34) **gepoolt**, um die Regressionsgleichung (4.35)

$$r = \gamma r_{-1} + \delta_d + \omega$$

für alle Prognosehorizonte zusammen zu schätzen. Tabelle 4.30 zeigt die entsprechenden Ergebnisse des Tests auf Informationseffizienz, $H_0 : \gamma = 0$, wiederum sowohl für Variante I (gleich große Verzerrung über alle Prognosehorizonte) als auch Variante II (horizont-spezifische Verzerrungen) jeweils unter Annahme A (Modell mit ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}).

Wird in der Alternativhypothese also unterstellt, dass die Verzerrung für alle Prognosehorizonte gleich groß sein muss (Variante I), kann die Nullhypothese der Informationseffizienz sowohl unter Annahme A als auch unter Annahme B auf dem 5%-Signifikanzniveau nicht verworfen werden. Lässt man in der Alternativhypothese hingegen horizont-spezifische Verzerrungen zu (Variante II), kann die Nullhypothese der Informationseffizienz im Modell mit idiosynkratischem Störterm (Annahme A)

Tabelle 4.30.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 2 der AKS-Prognosen mit $H = 5$**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | -0,20 | -0,20 | -0,20 | -0,20 |
| $se(\hat{\gamma})$ | 0,10 | 0,16 | 0,08 | 0,16 |
| p -Wert | 0,05 | 0,22 | 0,02 | 0,21 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 2). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 5$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

bei einem Signifikanzniveau von fünf Prozent verworfen werden. Aufgrund des negativen Schätzwertes für die Varianz des idiosynkratischen Störterms ε_{th} unter Annahme A, scheint die Modellierung unter Annahme B, bei der $\sigma_\varepsilon^2 = 0$ gesetzt wird, allerdings erneut plausibler. Im Modell ohne idiosynkratischen Störterm kann dann aber die Nullhypothese wiederum nicht verworfen werden, auch wenn in der Alternativhypothese horizont-spezifische Verzerrungen zugelassen werden. Somit können die AKS-Schätzungen im plausibleren Modell ebenfalls als **effizient** beurteilt werden, wenn als zugrunde liegende **Informationsmenge die vergangenen Änderungen der Prognosen desselben Prognoseziels** $\{r_{t,h-1+i;h+i}\}$ betrachtet wird.

Außerdem spielt es abermals keine Rolle, ob das Wiedervereinigungsjahr 1990 als Prognoseziel weggelassen wird oder nicht. Die Tabellen C.55 bis C.57 im Anhang geben die diesbezüglichen Testergebnisse für die beiden maximalen Prognosehorizonte $H = 5$ und $H = 4$ wieder. Wird das Prognoseziel 1991 eliminiert, gibt es ebenfalls keine veränderten Testentscheidungen, wie den Tabellen C.58 bis C.60 im Anhang zu entnehmen ist.

Werden jedoch alle Prognosen, die 1991 unter mangelhafter Datengrundlage erstellt wurden, aus der Analyse ausgeschlossen, verschlechtert sich in diesem Fall überraschenderweise die Informationseffizienz der AKS-Prognosen. Tabelle 4.31 stellt zunächst die Testergebnisse für die Prognosen mit maximalem Prognosehorizont $H = 5$ dar, wenn die Nullhypothese der Informationseffizienz gemäß Fall 2 separat für jeden Prognosehorizont getestet wird.

Tabelle 4.31.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 2 der AKS-Prognosen mit $H = 5$ ohne die Jahre 1991 bis 1995**

| h | $\hat{\gamma}$ | Annahme A | | Annahme B | |
|-----|----------------|--------------------|-----------|--------------------|-----------|
| | | $se(\hat{\gamma})$ | p -Wert | $se(\hat{\gamma})$ | p -Wert |
| 2 | -0,19 | 0,11 | 0,08 | 0,17 | 0,25 |
| 3 | -0,39 | 0,10 | 0,0005 | 0,15 | 0,02 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2$ und $h = 3$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t;h+1;h+2} + \delta + \omega_t$ (Fall 2). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für jedes h sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 5$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1976 bis 1990 sowie 1996 bis 2010.

Im Gegensatz zur ursprünglichen Stichprobe erweisen sich die AKS-Prognosen für die Jahre 1976 bis 1990 sowie 1996 bis 2010 nun für $h = 3$ sowohl unter Annahme A als auch unter Annahme B auf dem 5%-Signifikanzniveau **nicht mehr als effizient**.¹⁸

Wird der Test für alle Prognosehorizonte zusammen durchgeführt (siehe Tabelle 4.32), kann die Nullhypothese der Informationseffizienz nur noch in Variante I unter Annahme B auf dem 5%-Signifikanzniveau nicht verworfen werden. Werden

¹⁸Für die AKS-Prognosen mit maximalem Prognosehorizont $H = 4$ ergibt sich allerdings keine Änderung des Testergebnisses, wie Tabelle C.61 im Anhang zu entnehmen ist.

in der Alternativhypothese jedoch horizont-spezifische Verzerrungen zugelassen (Variante II), können die AKS-Prognosen für die Jahre 1976 bis 1990 sowie 1996 bis 2010 **nicht mehr als effizient** beurteilt werden, wenn als zugrunde liegende Informationsmenge die vergangenen Änderungen der Prognosen desselben Prognoseziels $\{r_{t;h-1+i;h+i}\}$ betrachtet wird.

Tabelle 4.32.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 2 der AKS-Prognosen mit $H = 5$ ohne die Jahre 1991 bis 1995**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | -0,30 | -0,30 | -0,30 | -0,30 |
| $se(\hat{\gamma})$ | 0,14 | 0,16 | 0,10 | 0,15 |
| p -Wert | 0,03 | 0,06 | 0,005 | 0,04 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 2). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei beziehungsweise drei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 5$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1976 bis 1990 sowie 1996 bis 2010.

Somit erweisen sich die Prognosen des Arbeitskreises „Steuerschätzungen“ **insgesamt** betrachtet in den meisten untersuchten Fällen zwar als **effizient**, unabhängig davon welche der beiden betrachteten Informationsmengen der Analyse zugrunde liegt. Allerdings erscheinen die Testergebnisse **nicht sonderlich robust** gegenüber der Auswahl der Stichprobe, den unterschiedlichen Modellannahmen sowie der Wahl des Signifikanzniveaus.

4.3.3. Fazit zur Prognoserationalität der AKS-Schätzungen

Zusammenfassend lässt sich festhalten, dass die bisherigen empirischen Ergebnisse zur Untersuchung der Prognosen des Arbeitskreises „Steuerschätzungen“ auf Unverzerrtheit von Becker und Buettner (2007), Büttner und Kauder (2008b) sowie Lehmann (2010) und diejenige auf Informationseffizienz von Becker und Buettner (2007) nur sehr eingeschränkt bestätigt werden können. Anstatt wie diese Autoren separat für verschiedene Prognosehorizonte auf Unverzerrtheit beziehungsweise Informationseffizienz zu testen, wurden die Prognosen des AKS in dieser Arbeit erstmalig gemäß des gepoolten Ansatzes von Clements et al. (2007) auf ihre Prognoserationalität hin untersucht.

Die Testentscheidung, ob die AKS-Prognosen verzerrt sind, hängt dabei von den getroffenen **Modellannahmen** ab. Wird in der Alternativhypothese unterstellt, dass die Verzerrung für alle Prognosehorizonte gleich groß sein muss (Variante I), kann die Nullhypothese der Unverzerrtheit der Prognosen in allen betrachteten Fällen eindeutig nicht verworfen werden. Lässt man in der Alternativhypothese hingegen horizont-spezifische Verzerrungen zu (Variante II), dann erweisen sich die untersuchten Prognosen lediglich im Modell mit idiosynkratischem Störterm (Annahme A) als unverzerrt. Dabei resultiert allerdings ein negativer Schätzwert für σ_ε^2 . Wird $\sigma_\varepsilon^2 = 0$ gesetzt (Annahme B) muss jedoch von **verzerrten Prognosen** ausgegangen werden.

Wird für jeden Prognosehorizont einzeln auf **Effizienz** getestet und werden dabei die Änderungen der Prognosen gleichbleibender aufeinanderfolgender Prognosehorizonte für verschiedene vergangene Prognoseziele $\{r_{t-1;h-1;h}, r_{t-2;h-1;h}, \dots\}$ (Fall 1) als zugrunde liegende Informationsmenge berücksichtigt, stellen sich die Prognosen des Arbeitskreises „Steuerschätzungen“ bis zu einem Prognosehorizont von drei Jahren als effizient dar. Für den mittelfristigen Prognosehorizont von vier Jahren ist dagegen ein **schwacher Glättungseffekt** erkennbar. Dieser verschwindet jedoch,

wenn alle Prognosen des AKS, die 1991 unter mangelhafter Datengrundlage erstellt wurden, aus der Analyse ausgeschlossen werden.

Wenn der Test für alle Prognosehorizonte zusammen durchgeführt wird, hängt die Testentscheidung für die AKS-Prognosen für 1981 bis 2010 abermals von den getroffenen Modellannahmen ab. Demnach kann die Nullhypothese der Informationseffizienz unter Annahme A auf dem 5%-Signifikanzniveau noch verworfen werden, während sie unter Annahme B bei einem Signifikanzniveau von fünf Prozent gerade nicht mehr verworfen werden kann. Da sich unter Annahme A allerdings ein negativer Schätzwert für die Varianz des idiosynkratischen Störterms ε_{th} ergibt, scheint die Modellierung unter Annahme B, bei der $\sigma_\varepsilon^2 = 0$ gesetzt wird, plausibler. Somit erweisen sich die AKS-Prognosen also unter der realistischeren Modellannahme als effizient.

Werden alle Prognosen, die 1991 unter mangelhafter Datengrundlage erstellt wurden, aus der Analyse ausgeschlossen, fällt das Testergebnis erneut eindeutiger aus. Denn dann kann die Nullhypothese der Informationseffizienz auch unter Annahme A nicht verworfen werden (außer für die Prognosen mit maximalem Prognosehorizont $H = 3$). Somit können die Steuerschätzungen des AKS also insgesamt in den meisten der untersuchten Fälle als **effizient** beurteilt werden, wenn als zugrunde liegende **Informationsmenge die Änderungen der Prognosen gleichbleibender aufeinanderfolgender Prognosehorizonte für verschiedene vergangene Prognoseziele** $\{r_{t-1;h-1;h}, r_{t-2;h-1;h}, \dots\}$ betrachtet wird.

Wird dagegen als zugrunde liegende Informationsmenge die vergangenen Änderungen der Prognosen desselben Prognoseziels $\{r_{t;h-1+i;h+i}\}$ (Fall 2) berücksichtigt, hängt die Testentscheidung vor allem von der untersuchten Stichprobe ab. Wenn der Test für jeden Prognosehorizont separat durchgeführt wird, erweisen sich die AKS-Prognosen für die Jahre 1981 bis 2010 als effizient, während die Nullhypothese der Informationseffizienz für die AKS-Prognosen für die Jahre 1976 bis 1990

sowie 1996 bis 2010 bei einem Prognosehorizont von zwei Jahren ($h = 3$) auf dem 5%-Signifikanzniveau verworfen werden kann.

Ebenso verändert sich die Testentscheidung in **Abhängigkeit von der Stichprobe**, wenn der Test für alle Prognosehorizonte zusammen durchgeführt wird. Auch dann implizieren die Testergebnisse, dass die AKS-Prognosen für die Jahre 1981 bis 2010 unter der plausibleren Modellannahme B (mit $\sigma_\varepsilon^2 = 0$) effizient sind, wenn als zugrunde liegende **Informationsmenge die vergangenen Änderungen der Prognosen desselben Prognoseziels** $\{r_{t;h-1+i;h+i}\}$ betrachtet wird. Werden jedoch alle Prognoseziele, die 1991 unter mangelhafter Datengrundlage erstellt wurden, aus der Stichprobe genommen und durch die Prognosen für die Jahre 1976 bis 1980 ersetzt, erweisen sich die AKS-Prognosen für diesen Zeitraum unter derselben Modellannahme in Variante II (mit horizont-spezifischen Verzerrungen) bei einem Signifikanzniveau von fünf Prozent nicht mehr als effizient.

Somit können die offiziellen Prognosen des Arbeitskreises „Steuerschätzungen“ im untersuchten Zeitraum in den meisten der untersuchten Fälle zwar als **effizient** beurteilt werden. Dennoch erweisen sich die Testergebnisse nicht robust gegenüber der genauen Auswahl der Stichprobe, den unterschiedlichen Modellannahmen sowie der Wahl des Signifikanzniveaus. Aufgrund der signifikanten Verzerrung im Modell mit den plausibelsten Annahmen sind die offiziellen Prognosen des Arbeitskreises „Steuerschätzungen“ allerdings **nicht als rational** einzustufen. Die als Alternative vorgeschlagenen SARIMA-Prognosen erwiesen sich hingegen bei einem maximalen Prognosehorizont von bis zu drei Jahren ($H = 4, 8, 12$) als unverzerrt sowie für einen maximalen Prognosehorizont von bis zu zwei Jahren ($H = 4$ und $H = 8$) als effizient und können dementsprechend auch als rational bezeichnet werden.

Im folgenden Kapitel wird nun die Treffsicherheit der offiziellen AKS-Prognosen mit derjenigen der als Alternative vorgeschlagenen SARIMA-Prognosen anhand ausgewählter deskriptiver Gütemaße verglichen.

4.4. Vergleich der Prognosegenauigkeit anhand deskriptiver Gütemaße

Präzise Prognosen sollten einerseits unverzerrt und effizient sein. Andererseits sollten sie eine hohe Treffsicherheit aufweisen. Abschließend wird nun untersucht, ob die als Alternative vorgeschlagenen SARIMA-Prognosen für die Jahre 2001 bis 2010 eine mindestens vergleichbare Prognosegenauigkeit aufweisen wie die offiziellen AKS-Schätzungen dieses Zeitraums. Anhand deskriptiver Gütemaße wird dementsprechend sowohl die horizontale als auch vertikale Treffsicherheit der Prognosen verglichen.¹⁹ Wie in Kapitel 2.3 beschrieben, wird bei der horizontalen Treffsicherheitsanalyse die Entwicklung der Prognosegenauigkeit für bestimmte Prognoseziele untersucht (**fixed event**), während die vertikale Treffsicherheitsanalyse die Prognosegüte der Prognosen mit einem festgelegten Prognosehorizont betrachtet (**fixed horizon**) (vgl. z. B. Osterloh (2008), S. 28). Außerdem werden die ausgewählten Gütemaße anschließend ebenfalls für alle Prognoseziele und Prognosehorizonte zusammen (**gepooled**) berechnet und verglichen.²⁰

Ausgangspunkt der verwendeten deskriptiven Gütemaße ist der Prozentfehler beziehungsweise **Percentage Error** PE_{th} , der sich als

$$PE_{th} = \frac{e_{th}}{A_t} 100 \quad (4.38)$$

berechnet (vgl. beispielsweise Bowen und Starr (1982), S. 23). Der Prognosefehler e_{th} (gemäß Gleichung (4.2)) wird also zum tatsächlichen Wert A_t in Relation gesetzt, so

¹⁹Auf statistische Tests wie den üblicherweise angewandten Diebold-Mariano-Test wird bei der horizontalen sowie vertikalen Treffsicherheitsanalyse verzichtet, da die jeweilige Anzahl an Beobachtungen hierfür zu gering ist (vgl. Diebold und Mariano (1995), S. 254 und Harvey et al. (1997), S. 282).

²⁰Es wird wiederum auf den Diebold-Mariano-Test verzichtet, da dieser für Prognosen mit einem bestimmten Prognosehorizont vorgesehen ist, und daher die durch das Poolen mehrerer Prognosehorizonte entstehende serielle Korrelation vermutlich nicht adäquat berücksichtigt. Außerdem ist selbst im gepoolten Fall die jeweilige Anzahl an Beobachtungen mit 30 bzw. 40 Beobachtungswerten immer noch als kritisch einzustufen.

dass der Prozentfehler als relatives Maß nicht von der Skala der verwendeten Daten abhängt.

Damit sich positive und negative Abweichungen nicht kompensieren können, werden die **Absolutwerte** des Prozentfehlers gemittelt. Dabei wird also wie in der Literatur üblich von einer symmetrischen Verlustfunktion ausgegangen, obwohl Überschätzungen des Steueraufkommens für gewöhnlich von der Öffentlichkeit als schlechter bewertet werden als Unterschätzungen des gleichen Ausmaßes, da erstere Steuerausfälle und somit Haushaltslöcher nach sich ziehen (vgl. beispielsweise auch Körner (1995), S. 14).

Somit erhält man für die horizontale Treffsicherheitsanalyse (**fixed event**) den mittleren absoluten Prozentfehler für das Prognoseziel t bzw. **Mean Absolute Percentage Error MAPE_t**

$$\text{MAPE}_t = \frac{1}{H} \sum_{h=1}^H |\text{PE}_{th}|, \quad (4.39)$$

der als Maß für die durchschnittliche Prognosegenauigkeit für ein bestimmtes Prognoseziel t verwendet werden kann.

Analog ergibt sich für die vertikale Treffsicherheitsanalyse (**fixed horizon**) der mittlere absolute Prozentfehler für den Prognosehorizont h bzw. **Mean Absolute Percentage Error MAPE_h**

$$\text{MAPE}_h = \frac{1}{T} \sum_{t=1}^T |\text{PE}_{th}|, \quad (4.40)$$

der entsprechend ein Maß für die durchschnittliche Prognosegenauigkeit für einen bestimmten Prognosehorizont h darstellt (vgl. z. B. Diebold (1998), S. 344).

Im **gepoolten** Ansatz wird infolgedessen für alle Prognoseziele t und Prognosehorizonte h der mittlere absolute Prozentfehler bzw. Mean Absolute Percentage Error MAPE als

$$\text{MAPE} = \frac{1}{T} \sum_{t=1}^T \text{MAPE}_t = \frac{1}{H} \sum_{h=1}^H \text{MAPE}_h \quad (4.41)$$

berechnet.

Sollen große Abweichungen stärker gewichtet werden als kleine, so empfiehlt sich als Gütemaß der mittlere **quadratische** Prozentfehler beziehungsweise **Mean Squared Percentage Error** (MSPE) sowie dessen Wurzel, der **Root Mean Squared Percentage Error** (RMSPE).

Wiederum wird zwischen der horizontalen Treffsicherheitsanalyse mit dem mittleren quadratischen Prozentfehler für das Prognoseziel t bzw. Mean Squared Percentage Error MSPE_t

$$\text{MSPE}_t = \frac{1}{H} \sum_{h=1}^H \text{PE}_{th}^2 \quad (4.42)$$

sowie dem dazugehörigen Root Mean Squared Percentage Error RMSPE_t

$$\text{RMSPE}_t = \sqrt{\text{MSPE}_t} \quad (4.43)$$

und der vertikalen Treffsicherheitsanalyse mit dem mittleren quadratischen Prozentfehler für den Prognosehorizont h bzw. Mean Squared Percentage Error MSPE_h

$$\text{MSPE}_h = \frac{1}{T} \sum_{t=1}^T \text{PE}_{th}^2 \quad (4.44)$$

sowie dem dazugehörigen Root Mean Squared Percentage Error RMSPE_h

$$\text{RMSPE}_h = \sqrt{\text{MSPE}_h} \quad (4.45)$$

unterschieden (vgl. z. B. Diebold (1998), S. 344).

Demzufolge wird beim gepoolten Ansatz für alle Prognoseziele t und Prognosehorizonte h der mittlere quadratische Prozentfehler bzw. Mean Squared Percentage Error MSPE als

$$\text{MSPE} = \frac{1}{T} \sum_{t=1}^T \text{MSPE}_t = \frac{1}{H} \sum_{h=1}^H \text{MSPE}_h \quad (4.46)$$

berechnet.

Den Root Mean Squared Percentage Error RMSPE erhält man entsprechend als

$$\text{RMSPE} = \sqrt{\text{MSPE}} \quad (4.47)$$

für alle Prognoseziele t und Prognosehorizonte h zusammen.

Der RMSPE wird häufig für den Vergleich verschiedener Prognosemodelle verwendet. So können die offiziellen AKS-Prognosen ebenfalls mit den als Alternative vorgeschlagenen SARIMA-Prognosen verglichen werden, indem die **modifizierten Theilschen Ungleichheitskoeffizienten**

$$U'_t = \frac{\text{RMSPE}_t(\text{AKS})}{\text{RMSPE}_t(\text{SARIMA})} \quad (4.48)$$

für das Prognoseziel t ,

$$U'_h = \frac{\text{RMSPE}_h(\text{AKS})}{\text{RMSPE}_h(\text{SARIMA})} \quad (4.49)$$

für den Prognosehorizont h sowie

$$U' = \frac{\text{RMSPE}(\text{AKS})}{\text{RMSPE}(\text{SARIMA})} \quad (4.50)$$

für alle Prognoseziele t und Prognosehorizonte h zusammen berechnet werden (vgl. Theil (1966), S. 28 sowie Davis und Fagan (1997), S. 707).

Sofern $U' < 1$ gilt, weisen die AKS-Schätzungen tendenziell eine höhere Prognosegenauigkeit auf als die SARIMA-Prognosen. Für $U' = 1$ werden beide Prognosemodelle gleich bewertet, während $U' > 1$ darauf hindeutet, dass sich die alternativen SARIMA-Prognosen besser zur Prognose des Gesamtsteueraufkommens eignen als die offiziellen Prognosen des Arbeitskreises "Steuerschätzungen".

Um die beschriebenen deskriptiven Gütemaße sowohl für die SARIMA-Schätzungen als auch die AKS-Prognosen berechnen und vergleichen zu können, müssen die Quartalsprognosen, die anhand der beiden SARIMA-Modelle erstellt wurden, die gleiche Struktur wie die Jahresprognosen des Arbeitskreises „Steuerschätzungen“ aufweisen.

Die SARIMA-Schätzungen, deren 55 Prognosezeitpunkte vom ersten Quartal 1997 bis zum dritten Quartal 2010 reichen (vgl. Kapitel 3.3.3), wurden allerdings für 40 Prognoseziele, nämlich für alle Quartale zwischen dem ersten Quartal 2001 bis zum vierten Quartal 2010, erstellt. Die AKS-Prognosen, deren 14 Prognosezeitpunkte jeweils die Mai-Sitzungen der Jahre 1997 bis 2010 umfassen, beinhalten dagegen für die Jahre 2001 bis 2010 lediglich zehn Prognoseziele.

Somit müssen die SARIMA-Prognosen zunächst ebenfalls **zu Jahresprognosen kumuliert** werden, wobei sich allerdings zwei grundlegend verschiedene Vorgehensweisen ergeben. Entweder wird für die SARIMA-Schätzungen der gleiche Informationsstand wie für die Mai-Sitzungen des Arbeitskreises "Steuerschätzungen" vorausgesetzt (Methode 1), oder aber wird für die als Alternative vorgeschlagene SARIMA-Schätzmethode die Verwendung aller verfügbarer Informationen erlaubt (Methode 2). Im Folgenden werden beide Ansätze separat betrachtet, und jeweils eine horizontale, vertikale und gepoolte Treffsicherheitsanalyse durchgeführt.

4.4.1. Vergleich bei gleichem Informationsstand zum möglichst ähnlichen Prognosezeitpunkt

Die AKS-Prognosen (mit $h = 1, \dots, 5$) wurden in den Jahren 1997 bis 2010 immer im Mai unter Verwendung aller zum Sitzungstermin verfügbarer Informationen erstellt. Um für die SARIMA-Schätzungen den **gleichen Informationsstand zum möglichst ähnlichen Prognosezeitpunkt** zu simulieren, ergibt sich daher als erste Vorgehensweise der Kumulation der Quartalsprognosen, nur diejenigen SARIMA-Schätzungen zu verwenden, die im zweiten Quartal prognostiziert wurden (**Methode 1**).

Eine Prognose für das laufende Jahr wie die des AKS mit $h = 1$ ist für die SARIMA-Schätzungen nicht möglich. Die Quartalsprognosen mit $h = 3$, $h = 4$, $h = 5$ und $h = 6$ werden zur Jahresprognose für das kommende Jahr ($h = 2$ im Sinne des AKS) kumuliert, die Addition der Prognosen mit $h = 7$, $h = 8$, $h = 9$ und $h = 10$

ergibt die Zwei-Jahres-Prognose für das darauffolgende Jahr ($h = 3$ im Sinne des AKS). Analog werden die Quartalsprognosen mit $h = 11$, $h = 12$, $h = 13$ und $h = 14$ zur Drei-Jahres-Prognose ($h = 4$ im Sinne des AKS) kumuliert. Bei dieser Vorgehensweise ist die Konstruktion von Vier-Jahres-Prognosen ($h = 5$ im Sinne des AKS) allerdings nicht möglich.

Somit umfassen die zur weiteren Analyse zugrunde liegenden Prognosematrizen der beiden SARIMA-Schätzungen jeweils die Prognosen für zehn Prognoseziele t , nämlich für die Jahre 2001 bis 2010, sowie für drei Prognosehorizonte, nämlich $h = 2, 3, 4$ im Sinne des AKS.

Zur Analyse der Prognosegüte werden als Erstes sowohl für die AKS-Prognosen als auch für die beiden SARIMA-Schätzungen die entsprechenden Werte der **Prozentfehler** PE_{th} gemäß Gleichung (4.38) berechnet. Tabelle 4.33 stellt beispielhaft die Werte des Prozentfehlers der AKS-Prognosen der Jahre 2001 bis 2010 für alle Prognosehorizonte $h = 1, \dots, 5$ dar. Für die beiden SARIMA-Schätzungen sind die Werte der Prozentfehler gemäß Methode 1 in den Tabellen D.1 und D.2 im Anhang dargestellt.

Es fällt auf, dass für die AKS-Prognosen im untersuchten Zeitraum häufiger negative Werte des Prozentfehlers PE_{th} als positive zu beobachten sind. So **überschätzte** der Arbeitskreis "Steuerschätzungen" das Gesamtsteueraufkommen der BRD in 32 der untersuchten 50 Fälle, vor allem in den Jahren **2001 bis 2005**.

Außerdem zeigt sich, dass in dieser ersten Hälfte des Beobachtungszeitraums die Überschätzungen des Steueraufkommens der Zwei-, Drei- und Vier-Jahres-Prognosen meistens sogar über zehn Prozent betragen. Bemerkenswerterweise sind dagegen die Absolutwerte des Prozentfehlers während und direkt nach der Finanzmarktkrise für alle Prognosehorizonte vergleichsweise gering.

Im Folgenden wird nun die horizontale, vertikale sowie gepoolte Prognosegüte der AKS-Schätzungen mit derjenigen der SARIMA-Prognosen verglichen.

Tabelle 4.33.: **Prozentfehler PE_{th} der AKS-Prognosen**

| Prognoseziel t | Prognosehorizont h: | | | | |
|----------------|---------------------|-------|--------|--------|--------|
| | 1 | 2 | 3 | 4 | 5 |
| 2001 | -2,11 | -7,84 | -7,27 | -7,91 | -11,67 |
| 2002 | -2,97 | -7,54 | -12,04 | -10,71 | -13,52 |
| 2003 | -1,72 | -7,30 | -11,27 | -17,74 | -15,31 |
| 2004 | -0,23 | -4,86 | -12,60 | -16,60 | -23,63 |
| 2005 | 1,57 | -0,31 | -3,67 | -12,43 | -16,58 |
| 2006 | 4,69 | 6,51 | 3,01 | -0,76 | -9,71 |
| 2007 | 0,72 | 8,21 | 12,36 | 8,40 | 5,09 |
| 2008 | 1,21 | 1,05 | 9,59 | 13,08 | 8,93 |
| 2009 | -0,57 | -8,97 | -9,73 | 0,52 | 4,37 |
| 2010 | 3,82 | 3,80 | -12,18 | -12,12 | -1,30 |

Werte des auf zwei Nachkommastellen gerundeten Prozentfehlers PE_{th} der AKS-Prognosen des Gesamtsteueraufkommens der BRD für die Prognoseziele $t = 2001$ bis 2010 und Prognosehorizonte $h = 1, \dots, 5$.

Horizontale Treffsicherheitsanalyse

Auf Grundlage der Prozentfehler PE_{th} wird nun zunächst die Prognosegüte anhand der dargestellten deskriptiven Gütemaße **für die Prognoseziele 2001 bis 2010** untersucht (**fixed event**). Tabelle 4.34 zeigt die berechneten Werte des $MAPE_t$ und des $RMSPE_t$ der AKS-Prognosen sowie beider SARIMA-Modelle für diese Jahre. Außerdem ist für diese Prognoseziele der Theilsche Ungleichheitskoeffizient U'_t aufgeführt, der gemäß Gleichung (4.48) für beide SARIMA-Modelle jeweils für jedes Jahr t den Wert des $RMSPE_t$ der AKS-Prognosen zum Wert des $RMSPE_t$ der entsprechenden SARIMA-Prognosen ins Verhältnis setzt. Da die Quartalsprognosen gemäß Methode 1 kumuliert wurden, werden bei der Berechnung aller Maße jeweils die Prognosen der Prognosehorizonte $h = 2, 3, 4$ berücksichtigt. Die Prognosen des laufenden Jahres sowie die Vier-Jahres-Prognosen des AKS gehen demzufolge also nicht in die Berechnung der Gütemaße ein.

Tabelle 4.34.: **Horizontale Treffsicherheitsanalyse gemäß Methode 1**

| Gütemaß | Prognoseziel t | AKS | SARIMA M1 | SARIMA M2 |
|--------------------|----------------|-------|-----------|-----------|
| MAPE _t | 2001 | 7,67 | 9,10 | 9,07 |
| | 2002 | 10,10 | 14,62 | 14,65 |
| | 2003 | 12,10 | 12,19 | 12,23 |
| | 2004 | 11,35 | 8,41 | 8,40 |
| | 2005 | 5,47 | 8,44 | 8,35 |
| | 2006 | 3,43 | 8,95 | 8,90 |
| | 2007 | 9,66 | 4,97 | 4,93 |
| | 2008 | 7,91 | 4,80 | 4,77 |
| | 2009 | 6,41 | 18,74 | 18,67 |
| | 2010 | 9,37 | 16,88 | 16,88 |
| RMSPE _t | 2001 | 7,68 | 9,82 | 9,81 |
| | 2002 | 10,27 | 15,91 | 15,93 |
| | 2003 | 12,85 | 13,59 | 13,62 |
| | 2004 | 12,35 | 8,95 | 8,95 |
| | 2005 | 7,49 | 10,58 | 10,49 |
| | 2006 | 4,16 | 10,17 | 10,10 |
| | 2007 | 9,84 | 5,53 | 5,49 |
| | 2008 | 9,38 | 4,82 | 4,78 |
| | 2009 | 7,65 | 18,77 | 18,69 |
| | 2010 | 10,16 | 19,74 | 19,71 |
| U' _t | 2001 | | 0,78 | 0,78 |
| | 2002 | | 0,65 | 0,64 |
| | 2003 | | 0,95 | 0,94 |
| | 2004 | | 1,38 | 1,38 |
| | 2005 | | 0,71 | 0,71 |
| | 2006 | | 0,41 | 0,41 |
| | 2007 | | 1,78 | 1,79 |
| | 2008 | | 1,95 | 1,96 |
| | 2009 | | 0,41 | 0,41 |
| | 2010 | | 0,51 | 0,52 |

Werte des auf zwei Nachkommastellen gerundeten mittleren absoluten Prozentfehlers MAPE_t sowie der Wurzel aus dem mittleren quadratischen Prozentfehler RMSPE_t der AKS-Prognosen (AKS) sowie der Prognosen des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells (SARIMA M1) und des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells (SARIMA M2) für die Prognoseziele $t = 2001$ bis 2010. Der Theilsche Ungleichheitskoeffizient U'_t wird gemäß Formel (4.48) für beide SARIMA-Modelle berechnet. Die Quartalsprognosen wurden gemäß Methode 1 kumuliert (vgl. Kapitel 4.4.1), so dass bei der Berechnung des MAPE_t sowie RMSPE_t jeweils die Prognosen der Prognosehorizonte $h = 2, 3, 4$ berücksichtigt werden.

Den in Tabelle 4.34 dargestellten Ergebnissen ist zu entnehmen, dass sowohl das im Box-Jenkins-Verfahren bevorzugte SARIMA(1,0,0)(1,0,1)₄-Modell als auch das alternative SARIMA(0,0,1)(1,0,1)₄-Modell in den meisten Jahren eine schlechtere Prognosegüte als die Prognosen des Arbeitskreises "Steuerschätzungen" erzielen. Lediglich für die Prognoseziele 2004, 2007 und 2008 weisen die berechneten Gütemaße auf eine höhere Treffsicherheit der SARIMA-Prognosen hin. Dies ist insbesondere am Theilschen Ungleichheitskoeffizienten U'_t erkennbar, der für diese Prognoseziele Werte deutlich größer eins annimmt.

Des Weiteren ist festzustellen, dass der mittlere absolute Prozentfehler der AKS-Prognosen mit $h = 2, 3, 4$ lediglich für das Prognoseziel 2006 unter fünf Prozent liegt. Außerdem fällt auf, dass die Ein-, Zwei- und Drei-Jahres-Prognosen des AKS für die Jahre 2002 bis 2004 mit Werten des $MAPE_t$ von jeweils über zehn Prozent eine höhere durchschnittliche Prognoseungenauigkeit aufweisen, als diejenigen der übrigen Jahre, insbesondere diejenigen für die Jahre 2008 und 2009, als die Finanzmarktkrise Deutschland erreichte bzw. 2009 auf die Realwirtschaft übergriff.

Die SARIMA-Modelle prognostizierten dagegen neben den Jahren 2002 und 2003 gerade für die Jahre 2009 und 2010 mit einem mittleren absoluten Prozentfehler von fast 19 bzw. 17 Prozent am deutlich schlechtesten im untersuchten Zeitraum. Dabei spielt es keine Rolle, ob das SARIMA(1,0,0)(1,0,1)₄-Modell oder das SARIMA(0,0,1)(1,0,1)₄-Modell der Prognose zugrunde liegt. So zeigt sich deutlich, dass die **Prognosequalität der direkten Zeitreihenmodelle stark nachlässt**, wenn am Ende des Schätzzeitraums ein **exogener Schock** – wie hier die Finanzmarktkrise mit ihren realwirtschaftlichen Auswirkungen – auftritt.

Abbildung 4.2 veranschaulicht beispielhaft die beschriebene Entwicklung des $MAPE_t$ für die AKS-Prognosen sowie für die Prognosen des SARIMA(1,0,0)(1,0,1)₄-Modells.

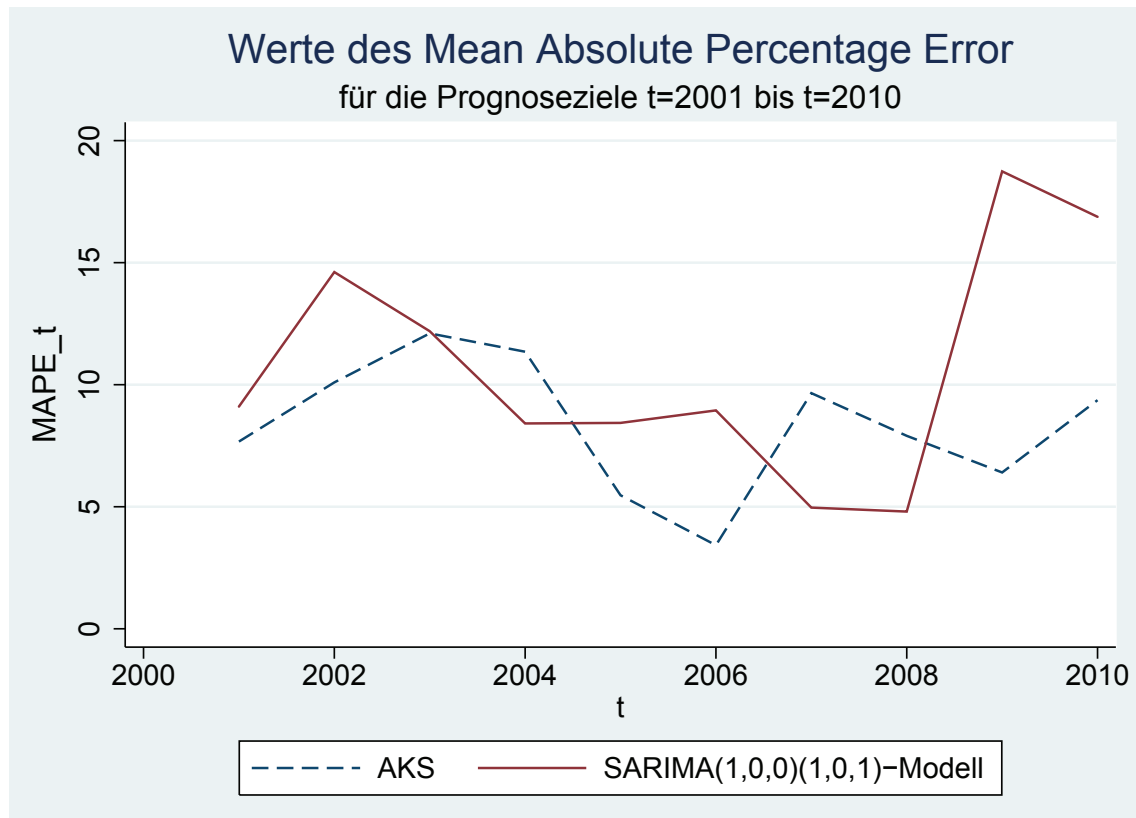


Abbildung 4.2.: Abbildung der Werte des $MAPE_t$ der AKS-Prognosen sowie der Prognosen des $SARIMA(1,0,0)(1,0,1)_4$ -Modells gemäß Methode 1

Vertikale Treffsicherheitsanalyse

Als Nächstes wird die Prognosegüte in Abhängigkeit von der Schätzdistanz untersucht (**fixed horizon**). Tabelle 4.35 zeigt die Werte des $MAPE_h$ und des $RMSPE_h$ der AKS-Prognosen für die Prognosehorizonte $h = 1, \dots, 5$ sowie des bevorzugten und alternativen SARIMA-Modells für $h = 2, 3, 4$. Außerdem ist für diese Prognosehorizonte der Theilsche Ungleichheitskoeffizient U'_h gemäß Formel (4.49) aufgeführt. Zur Berechnung aller Maße werden jeweils die Prognosen der Jahre 2001 bis 2010 berücksichtigt, wobei die Quartalsprognosen zuvor gemäß Methode 1 zu Jahresprognosen kumuliert wurden.

Tabelle 4.35.: **Vertikale Treffsicherheitsanalyse gemäß Methode 1**

| Gütemaß | Prognosehorizont h | AKS | SARIMA M1 | SARIMA M2 |
|-----------|----------------------|-------|-----------|-----------|
| $MAPE_h$ | 1 | 1,96 | – | – |
| | 2 | 5,64 | 6,09 | 6,09 |
| | 3 | 9,37 | 11,50 | 11,48 |
| | 4 | 10,03 | 14,53 | 14,49 |
| | 5 | 11,01 | – | – |
| $RMSPE_h$ | 1 | 2,40 | – | – |
| | 2 | 6,33 | 7,64 | 7,63 |
| | 3 | 9,97 | 13,09 | 13,06 |
| | 4 | 11,45 | 16,06 | 16,02 |
| | 5 | 12,68 | – | – |
| U'_h | 2 | | 0,83 | 0,83 |
| | 3 | | 0,76 | 0,76 |
| | 4 | | 0,71 | 0,71 |

Werte des auf zwei Nachkommastellen gerundeten mittleren absoluten Prozentfehlers $MAPE_h$ sowie der Wurzel aus dem mittleren quadratischen Prozentfehler $RMSPE_h$ der AKS-Prognosen (AKS) sowie der Prognosen des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells (SARIMA M1) und des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells (SARIMA M2) für die Prognosehorizonte h . Der Theilsche Ungleichheitskoeffizient U'_h wird gemäß Formel (4.49) für beide SARIMA-Modelle berechnet. Die Quartalsprognosen wurden gemäß Methode 1 kumuliert (vgl. Kapitel 4.4.1). Bei der Berechnung des $MAPE_h$ sowie $RMSPE_h$ werden jeweils die Prognosen der Jahre 2001 bis 2010 berücksichtigt.

Anhand der Werte des $MAPE_h$ und des $RMSPE_h$ wird deutlich, dass für alle Prognosehorizonte $h = 2, 3, 4$ die AKS-Prognosen eine höhere Prognosegüte als die Prognosen der beiden Zeitreihenmodelle aufweisen. Dementsprechend nimmt der Theilsche Ungleichheitskoeffizient U'_h für diese Prognosehorizonte auch jeweils Werte kleiner eins an. Des Weiteren ist für das alternative SARIMA(0,0,1)(1,0,1)₄-Modell eine tendenziell leicht höhere Treffsicherheit als für das im Box-Jenkins-Verfahren bevorzugte SARIMA(1,0,0)(1,0,1)₄-Modell festzustellen. Der Unterschied in der Prognosegüte beider Zeitreihenmodelle ist jedoch verschwindend klein.

Außerdem nimmt die Prognoseungenauigkeit wie erwartet für alle Modellvarianten bei wachsendem Prognosehorizont zu. Während der Arbeitskreis „Steuerschätzungen“ beispielsweise das Steueraufkommen des laufenden Jahres im untersuchten Zeitraum noch mit einem mittleren absoluten Prozentfehler von knapp unter zwei Prozent prognostizierte, liegt der Wert des MAPE_h für die Prognosen des kommenden Jahres ($h = 2$) bereits über fünf Prozent. Mit weiter zunehmendem Prognosehorizont steigt die Prognoseungenauigkeit der AKS-Schätzungen auf bis zu elf Prozent für die Vier-Jahres-Prognosen an.

Für die SARIMA-Schätzungen ist diese Abnahme der Treffsicherheit bei wachsendem Prognosehorizont allerdings noch deutlicher ausgeprägt. So weisen bereits die Zwei-Jahres-Prognosen mit einem Wert des mittleren absoluten Prozentfehlers von circa 11,5 Prozent eine größere Prognoseungenauigkeit auf als die Vier-Jahres-Prognosen des AKS.

Abbildung 4.3 stellt diese Entwicklung der vertikalen Treffsicherheit beispielhaft anhand der Werte des MAPE_h für die AKS-Prognosen sowie für die Prognosen des $\text{SARIMA}(0,0,1)(1,0,1)_4$ -Modells graphisch dar.

Gepoolte Treffsicherheitsanalyse

Abschließend wird die gepoolte Treffsicherheitsanalyse durchgeführt. Tabelle 4.36 gibt die Werte des MAPE und des RMSPE der AKS-Prognosen sowie der Prognosen beider SARIMA-Modelle wieder. Außerdem ist für beide SARIMA-Modelle jeweils der Theilsche Ungleichheitskoeffizient U' gemäß Formel (4.50) aufgeführt. Die Quartalsprognosen wurden gemäß Methode 1 zu Jahresprognosen kumuliert. Zur Berechnung aller Maße wurden jeweils die Prognosen der Jahre 2001 bis 2010 und die Prognosehorizonte $h = 2, 3, 4$ berücksichtigt.

Auch bei der gepoolten Analyse bestätigen die berechneten deskriptiven Gütemaße den **Prognosen des Arbeitskreises „Steuerschätzungen“ eine höhere Treffsicherheit** als den Prognosen der als Alternative vorgeschlagenen Zeitreihenmodelle.

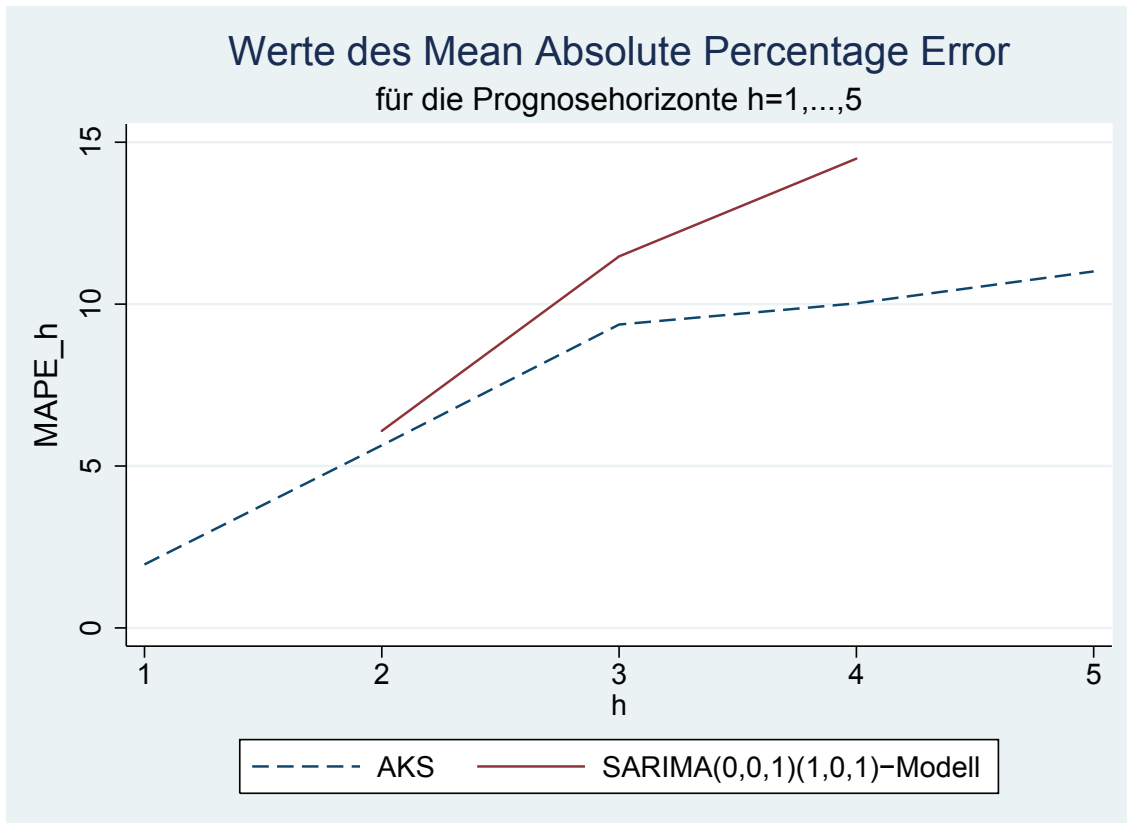


Abbildung 4.3.: Abbildung der Werte des $MAPE_h$ der AKS-Prognosen sowie der Prognosen des SARIMA(0,0,1)(1,0,1)₄-Modells gemäß Methode 1

Außerdem ist für die beiden SARIMA-Modelle kein großer Unterschied in der Prognosegüte festzustellen. Die Prognosen anhand des SARIMA(0,0,1)(1,0,1)₄-Modells weisen lediglich wieder eine tendenziell leicht höhere Treffsicherheit als diejenigen anhand des SARIMA(1,0,0)(1,0,1)₄-Modells auf.

So liegt der mittlere absolute Prozentfehler beider SARIMA-Prognosen mit einem Wert von über 10,5 Prozent deutlich höher als derjenige der AKS-Prognosen. Der Arbeitskreis „Steuerschätzungen“ erzielte mit seinen Ein-, Zwei- und Drei-Jahres-Prognosen für die Jahre 2001 bis 2010 nämlich einen mittleren absoluten Prozentfehler von unter 8,5 Prozent, was allerdings dennoch als hoch einzustufen ist. Denn dieser relative Wert entsprach einem mittleren absoluten Prognosefehler von immerhin über 40 Milliarden Euro pro Jahr.

Tabelle 4.36.: **Gepoolte Treffsicherheitsanalyse gemäß Methode 1**

| Gütemaß | AKS | SARIMA M1 | SARIMA M2 |
|---------|------|-----------|-----------|
| MAPE | 8,35 | 10,71 | 10,68 |
| RMSPE | 9,50 | 12,75 | 12,72 |
| U' | | 0,74 | 0,75 |

Werte des auf zwei Nachkommastellen gerundeten mittleren absoluten Prozentfehlers MAPE sowie der Wurzel aus dem mittleren quadratischen Prozentfehler RMSPE der AKS-Prognosen (AKS) sowie der Prognosen des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells (SARIMA M1) und des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells (SARIMA M2). Der Theilsche Ungleichheitskoeffizient U' wird gemäß Formel (4.50) für beide SARIMA-Modelle berechnet. Die Quartalsprognosen wurden gemäß Methode 1 kumuliert (vgl. Kapitel 4.4.1), so dass bei der Berechnung des MAPE sowie RMSPE jeweils die Prognosen der Prognosehorizonte $h = 2, 3, 4$ sowie der Jahre 2001 bis 2010 berücksichtigt werden.

Grundlage der bisherigen Ergebnisse ist jedoch die Kumulation der Quartalsprognosen gemäß Methode 1, das heißt es wurden nur diejenigen SARIMA-Schätzungen verwendet, die im zweiten Quartal prognostiziert wurden, um für die SARIMA-Prognosen den **gleichen Informationsstand wie für die Mai-Sitzungen des AKS** zu simulieren. Als Nächstes wird nun gemäß Methode 2 für die als Alternative vorgeschlagene SARIMA-Schätzmethode die Verwendung aller verfügbarer Informationen erlaubt.

4.4.2. Vergleich unter Verwendung aller verfügbarer Informationen zu allen Prognosezeitpunkten

Als alternative Vorgehensweise der Bildung der Jahresprognosen werden nun alle Prognosen, also auch diejenigen, die im ersten, dritten und vierten Quartal erstellt wurden, bei der Kumulation berücksichtigt (**Methode 2**). Daher steigt einerseits die Anzahl der Prognosezeitpunkte, andererseits werden zu jedem Prognosezeitpunkt alle verfügbaren Informationen verwendet, die sich aus den Quartalsdaten des Steueraufkommens ergeben. Diese Vorgehensweise wird im Folgenden an einem Beispiel erläutert.

Um die Jahresprognose für das Prognoseziel 2008 zu bilden, sind zwei Schritte notwendig. Als Erstes wird eine Durchschnittsprognose für das erste Quartal 2008 aus der Prognose zum Prognosezeitpunkt viertes Quartal 2007 ($h = 1$), derjenigen des dritten Quartals 2007 ($h = 2$), des zweiten Quartals 2007 ($h = 3$) sowie des ersten Quartals 2007 ($h = 4$) gebildet. Dieser Durchschnitt ergibt dann die Prognose für das erste Quartal 2008 mit $h = 2$ im Sinne des AKS (Prognose für das kommende Jahr).

Analog wird aus den Prognosen zum Prognosezeitpunkt viertes Quartal 2006 ($h = 5$), derjenigen des dritten Quartals 2006 ($h = 6$), des zweiten Quartals 2006 ($h = 7$) sowie des ersten Quartals 2006 ($h = 8$) das arithmetische Mittel gebildet, um die Prognose für das erste Quartal 2008 mit $h = 3$ im Sinne des AKS (Zwei-Jahres-Prognose) zu berechnen. Gleichmaßen werden die Durchschnittsprognosen für das erste Quartal 2008 mit $h = 4$ und $h = 5$ im Sinne des AKS (Drei- und Vier-Jahres-Prognose) ermittelt.

Dabei gehen also alle Prognosehorizonte $h = 1, \dots, 16$ der SARIMA-Prognosen in die Berechnung ein. Im zweiten Schritt werden analog die Prognosen für das zweite, dritte und vierte Quartal 2008, jeweils für $h = 2, \dots, 5$ im Sinne des AKS, konstruiert. Anschließend werden die entsprechenden Durchschnittsprognosen der Quartale zu Jahresprognosen (für $h = 2, \dots, 5$) kumuliert.

Für die folgende horizontale, vertikale sowie gepoolte Treffsicherheitsanalyse umfassen somit die Prognosematrizen der beiden SARIMA-Schätzungen nun jeweils die Prognosen für zehn Prognoseziele t , nämlich für die Jahre 2001 bis 2010, sowie für vier Prognosehorizonte, nämlich $h = 2, \dots, 5$ im Sinne des AKS. Daraus werden wieder die entsprechenden Werte der Prozentfehler PE_{th} berechnet.²¹

Horizontale Treffsicherheitsanalyse

Auf Grundlage dieser Prozentfehler PE_{th} wird nun erneut zunächst die Prognosegüte **für die Prognoseziele 2001 bis 2010** untersucht (**fixed event**). Tabelle 4.37 zeigt die berechneten Werte des $MAPE_t$ und des $RMSPE_t$ der AKS-Prognosen sowie des SARIMA(1,0,0)(1,0,1)₄- und des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells für diese Jahre. Außerdem ist für diese Prognoseziele wiederum der Theilsche Ungleichheitskoeffizient U'_t für beide SARIMA-Modelle aufgeführt. Da die Quartalsprognosen gemäß Methode 2 kumuliert wurden, können bei der Berechnung aller Maße nun jeweils die Prognosen der Prognosehorizonte $h = 2, \dots, 5$ berücksichtigt werden. Somit gehen die **Vier-Jahres-Prognosen** in die Berechnung der Gütemaße ein, während die Prognosen des laufenden Jahres des AKS weiterhin nicht berücksichtigt werden können.

Vergleicht man nun die Werte des $MAPE_t$ im untersuchten Zeitraum, so zeigt sich, dass die SARIMA-Prognosen in sechs der zehn betrachteten Jahre eine höhere durchschnittliche Prognosegenauigkeit als die Prognosen des Arbeitskreises „Steuerschätzungen“ aufweisen. Wird der $RMSPE_t$ zugrunde gelegt, werden also größere Abweichungen stärker gewichtet als kleine, so besitzen nur noch die SARIMA-Prognosen für die vier Prognoseziele 2004, 2005, 2007 und 2008 eine höhere Prognosegüte als die AKS-Prognosen. Dementsprechend sind für diese Jahre auch die Werte des Theilschen Ungleichheitskoeffizienten U'_t größer als eins.

²¹Die Tabellen D.3 und D.4 im Anhang geben die Werte der Prozentfehler für die beiden SARIMA-Schätzungen gemäß Methode 2 wieder.

Tabelle 4.37.: **Horizontale Treffsicherheitsanalyse gemäß Methode 2 mit Vier-Jahres-Prognosen**

| Gütemaß | Prognoseziel t | AKS | SARIMA M1 | SARIMA M2 |
|--------------------|----------------|-------|-----------|-----------|
| MAPE _t | 2001 | 8,67 | 8,49 | 8,48 |
| | 2002 | 10,95 | 13,20 | 13,20 |
| | 2003 | 12,91 | 12,53 | 12,52 |
| | 2004 | 14,42 | 7,17 | 7,15 |
| | 2005 | 8,25 | 4,26 | 4,20 |
| | 2006 | 5,00 | 6,65 | 6,58 |
| | 2007 | 8,51 | 5,67 | 5,61 |
| | 2008 | 8,16 | 6,05 | 6,01 |
| | 2009 | 5,90 | 12,06 | 12,07 |
| | 2010 | 7,35 | 14,70 | 14,69 |
| RMSPE _t | 2001 | 8,84 | 9,43 | 9,43 |
| | 2002 | 11,17 | 15,20 | 15,22 |
| | 2003 | 13,50 | 16,65 | 16,67 |
| | 2004 | 15,94 | 9,96 | 9,99 |
| | 2005 | 10,52 | 6,04 | 5,93 |
| | 2006 | 6,05 | 6,99 | 6,92 |
| | 2007 | 8,90 | 6,09 | 6,03 |
| | 2008 | 9,27 | 7,43 | 7,37 |
| | 2009 | 6,97 | 13,55 | 13,55 |
| | 2010 | 8,82 | 17,13 | 17,12 |
| U' _t | 2001 | | 0,94 | 0,94 |
| | 2002 | | 0,73 | 0,73 |
| | 2003 | | 0,81 | 0,81 |
| | 2004 | | 1,60 | 1,60 |
| | 2005 | | 1,74 | 1,77 |
| | 2006 | | 0,87 | 0,87 |
| | 2007 | | 1,46 | 1,47 |
| | 2008 | | 1,25 | 1,26 |
| | 2009 | | 0,51 | 0,51 |
| | 2010 | | 0,51 | 0,52 |

Werte des auf zwei Nachkommastellen gerundeten mittleren absoluten Prozentfehlers MAPE_t sowie der Wurzel aus dem mittleren quadratischen Prozentfehler RMSPE_t der AKS-Prognosen (AKS) sowie der Prognosen des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells (SARIMA M1) und des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells (SARIMA M2) für die Prognoseziele $t = 2001$ bis 2010. Der Theilsche Ungleichheitskoeffizient U'_t wird gemäß Formel (4.48) für beide SARIMA-Modelle berechnet. Die Quartalsprognosen wurden gemäß Methode 2 kumuliert (vgl. Kapitel 4.4.2). Bei der Berechnung des MAPE_t sowie RMSPE_t werden jeweils die Prognosen der Prognosehorizonte $h = 2, 3, 4, 5$ berücksichtigt.

Insgesamt kann nun eher von einer ähnlich guten Treffsicherheit der als Alternative vorgeschlagenen SARIMA-Modelle im Vergleich zu den AKS-Prognosen ausgegangen werden. Dabei spielt es wiederum keine große Rolle, welches der beiden SARIMA-Modelle der Prognose zugrunde liegt.

Die Werte des MAPE_t der AKS-Prognosen zeigen außerdem, dass unter Berücksichtigung der Prognosen mit $h = 5$ der mittlere absolute Prozentfehler in den meisten Jahren jeweils größer ist als derjenige ohne die Vier-Jahres-Prognosen (vgl. Tabelle 4.34). Als Ausnahmen davon sind die Prognoseziele 2007, 2009 und 2010 zu nennen. Folglich hat sich die durchschnittliche Prognosegüte gemäß des MAPE_t für diese Jahre durch die Berücksichtigung der Prognosen mit der größten Schätzdistanz sogar verbessert.

Vergleicht man ebenso die Werte des MAPE_t der SARIMA-Prognosen, ist in Tabelle 4.37 für die meisten Jahre eine höhere durchschnittliche Prognosegenauigkeit festzustellen. Allerdings bleibt offen, ob sich diese Verbesserung der Prognosegüte auf die Kumulation der Quartalsprognosen gemäß Methode 2 zurückführen lässt, oder aber die Hinzunahme der Vier-Jahres-Prognosen einen entscheidenden Einfluss hat.

Des Weiteren zeigt sich, dass die SARIMA-Modelle erneut für die Jahre 2002, 2003 sowie 2009 und 2010 mit einem mittleren absoluten Prozentfehler von über zwölf Prozent am deutlich schlechtesten im untersuchten Zeitraum prognostizierten. Allerdings fallen die Werte des MAPE_t für die Prognoseziele 2009 und 2010 nun für beide SARIMA-Modelle nicht ganz so hoch aus wie bei der bisherigen Analyse gemäß Methode 1 und ohne Berücksichtigung der Vier-Jahres-Prognosen.

Abbildung 4.4 veranschaulicht beispielhaft die Entwicklung des MAPE_t für die AKS-Prognosen sowie für die Prognosen des $\text{SARIMA}(1,0,0)(1,0,1)_4$ -Modells, die gemäß Methode 2 zu Jahresprognosen kumuliert wurden. Bei der Berechnung des mittleren absoluten Prozentfehlers für die Prognoseziele 2001 bis 2010 wurden dabei die Vier-Jahres-Prognosen berücksichtigt.

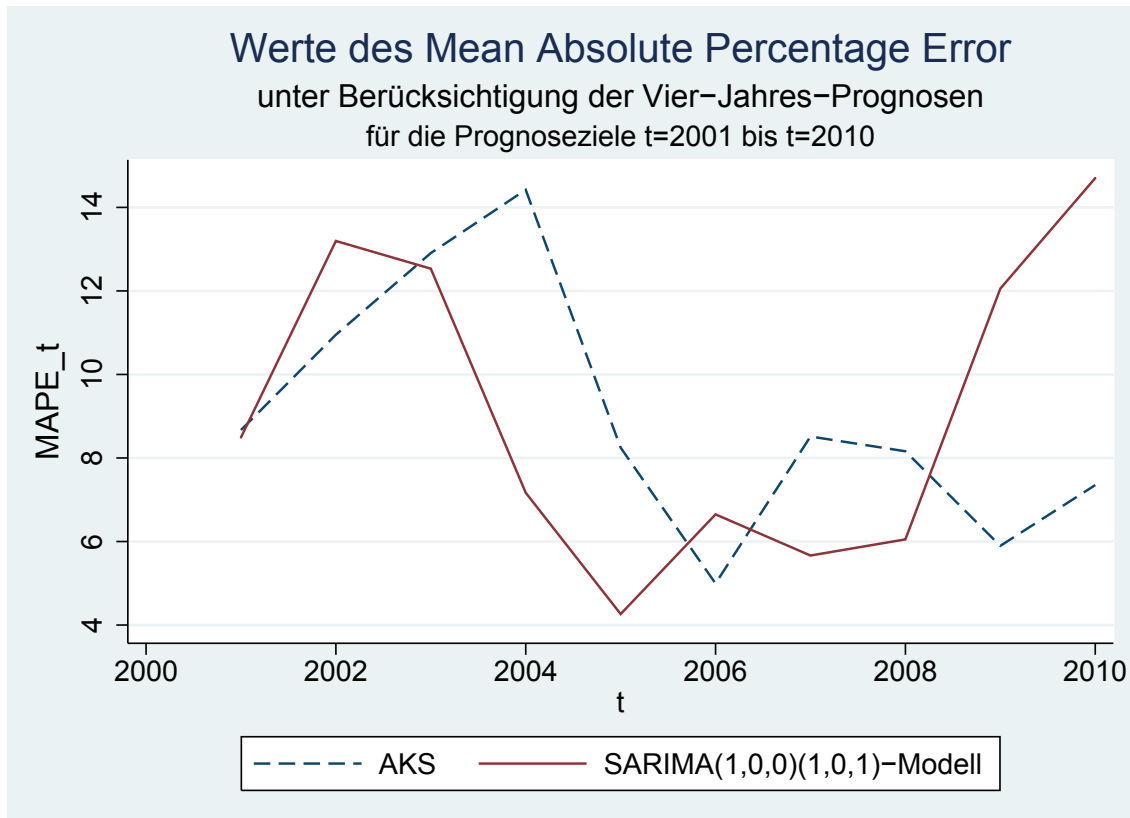


Abbildung 4.4.: Abbildung der Werte des $MAPE_t$ der AKS-Prognosen sowie der Prognosen des $SARIMA(1,0,0)(1,0,1)_4$ -Modells gemäß Methode 2 unter Berücksichtigung der Vier-Jahres-Prognosen

Um den Einfluss der beiden unterschiedlichen Methoden der Kumulation der Quartalsprognosen zu Jahresprognosen zu isolieren, wird im nächsten Schritt die horizontale Treffsicherheit der Prognosen gemäß Methode 2 **ohne Berücksichtigung der Vier-Jahres-Prognosen** untersucht. Die entsprechenden Werte des $MAPE_t$ und des $RMSPE_t$ der AKS-Prognosen sowie des $SARIMA(1,0,0)(1,0,1)_4$ - und des alternativen $SARIMA(0,0,1)(1,0,1)_4$ -Modells sind für die Jahre 2001 bis 2010 in Tabelle 4.38 dargestellt. Außerdem ist für diese Prognoseziele wiederum der Theilsche Ungleichheitskoeffizient U'_t für beide SARIMA-Modelle aufgeführt. Die Quartalsprognosen wurden gemäß Methode 2 zu Jahresprognosen kumuliert. Bei der Berechnung aller Maße werden nun allerdings jeweils wieder nur die Prognosen der Prognosehorizonte $h = 2, 3, 4$ berücksichtigt. Somit gehen die Vier-Jahres-Prognosen also wie bei Methode 1 nicht in die Berechnung der Gütemaße ein.

Tabelle 4.38.: **Horizontale Treffsicherheitsanalyse gemäß Methode 2 ohne Vier-Jahres-Prognosen**

| Gütemaß | Prognoseziel t | AKS | SARIMA M1 | SARIMA M2 |
|--------------------|----------------|-------|-----------|-----------|
| MAPE _t | 2001 | 7,67 | 10,45 | 10,47 |
| | 2002 | 10,10 | 11,66 | 11,66 |
| | 2003 | 12,10 | 6,88 | 6,86 |
| | 2004 | 11,35 | 3,43 | 3,36 |
| | 2005 | 5,47 | 1,83 | 1,81 |
| | 2006 | 3,43 | 5,69 | 5,64 |
| | 2007 | 9,66 | 5,70 | 5,66 |
| | 2008 | 7,91 | 3,94 | 3,91 |
| | 2009 | 6,41 | 14,10 | 14,11 |
| | 2010 | 9,37 | 9,87 | 9,87 |
| RMSPE _t | 2001 | 7,68 | 10,78 | 10,80 |
| | 2002 | 10,27 | 14,22 | 14,24 |
| | 2003 | 12,85 | 8,92 | 8,96 |
| | 2004 | 12,35 | 4,42 | 4,33 |
| | 2005 | 7,49 | 2,03 | 1,99 |
| | 2006 | 4,16 | 5,90 | 5,85 |
| | 2007 | 9,84 | 6,26 | 6,21 |
| | 2008 | 9,38 | 4,72 | 4,70 |
| | 2009 | 7,65 | 15,27 | 15,26 |
| | 2010 | 10,16 | 10,35 | 10,37 |
| U' _t | 2001 | | 0,71 | 0,71 |
| | 2002 | | 0,72 | 0,72 |
| | 2003 | | 1,44 | 1,43 |
| | 2004 | | 2,79 | 2,86 |
| | 2005 | | 3,68 | 3,76 |
| | 2006 | | 0,71 | 0,71 |
| | 2007 | | 1,57 | 1,58 |
| | 2008 | | 1,99 | 2,00 |
| | 2009 | | 0,50 | 0,50 |
| | 2010 | | 0,98 | 0,98 |

Werte des auf zwei Nachkommastellen gerundeten mittleren absoluten Prozentfehlers MAPE_t sowie der Wurzel aus dem mittleren quadratischen Prozentfehler RMSPE_t der AKS-Prognosen (AKS) sowie der Prognosen des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells (SARIMA M1) und des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells (SARIMA M2) für die Prognoseziele $t = 2001$ bis 2010. Der Theilsche Ungleichheitskoeffizient U'_t wird gemäß Formel (4.48) für beide SARIMA-Modelle berechnet. Die Quartalsprognosen wurden gemäß Methode 2 kumuliert (vgl. Kapitel 4.4.1). Bei der Berechnung des MAPE_t sowie RMSPE_t werden jeweils nur die Prognosen der Prognosehorizonte $h = 2, 3, 4$ berücksichtigt.

Anhand der Werte des MAPE_t sowie des RMSPE_t beider SARIMA-Prognosen zeigt sich, dass bei Kumulation der Quartalsprognosen gemäß Methode 2 für fast alle Prognoseziele eine höhere Treffsicherheit erzielt wird, als unter Kumulation der Quartalsprognosen gemäß Methode 1 (vgl. Tabelle 4.34). Lediglich für die Jahre 2001 sowie 2007 hat sich die Prognosegüte unter Verwendung aller verfügbarer Informationen leicht verschlechtert. Für diese beiden Prognoseziele sowie für das Jahr 2009 verbessert sich dagegen die durchschnittliche Prognosegenauigkeit, wenn die Vier-Jahres-Prognosen bei der Berechnung des MAPE_t berücksichtigt werden (vgl. Tabelle 4.37).

Für die Prognoseziele 2003 und 2004 sank bei Kumulation der Quartalsprognosen gemäß Methode 2 der mittlere absolute Prozentfehler um ungefähr fünf Prozentpunkte. Die Verringerung des MAPE_t beträgt für die Jahre 2005 und 2010 sogar über sechs Prozentpunkte. Dennoch prognostizieren die SARIMA-Modelle für das Jahr 2010 – neben den Jahren 2001, 2002 sowie 2009 – immer noch deutlich schlechter als für die übrigen Prognoseziele.

Während mit Methode 1 beide SARIMA-Modelle in den meisten Jahren eine schlechtere Prognosegüte als die Prognosen des Arbeitskreises „Steuerschätzungen“ erzielen, weisen die berechneten Gütemaße gemäß Methode 2 für die Hälfte der Prognoseziele auf eine höhere Treffsicherheit der SARIMA-Prognosen hin. Dies ist insbesondere am Theilschen Ungleichheitskoeffizienten U'_t erkennbar, der für die Jahre 2004 bzw. 2005 sogar außergewöhnlich hohe Werte größer zwei beziehungsweise drei annimmt.

Betrachtet man den untersuchten Zeitraum insgesamt, erreichen die Zeitreihenschätzungen, die gemäß Methode 2 alle verfügbaren Informationen nutzen, also eine vergleichbare Prognosegüte wie diejenigen des AKS. Dabei spielt es erneut keine große Rolle, ob das im Box-Jenkins-Verfahren bevorzugte SARIMA(1,0,0)(1,0,1)₄-Modell oder das alternative SARIMA(0,0,1)(1,0,1)₄-Modell der Prognose zugrunde liegt.

Abbildung 4.5 veranschaulicht die Entwicklung des $MAPE_t$ für die AKS-Prognosen sowie beispielhaft für die Prognosen des $SARIMA(1,0,0)(1,0,1)_4$ -Modells, die gemäß Methode 2 zu Jahresprognosen kumuliert wurden. Dabei wurden die Vier-Jahres-Prognosen bei der Berechnung des mittleren absoluten Prozentfehlers für die Prognoseziele 2001 bis 2010 nicht berücksichtigt.

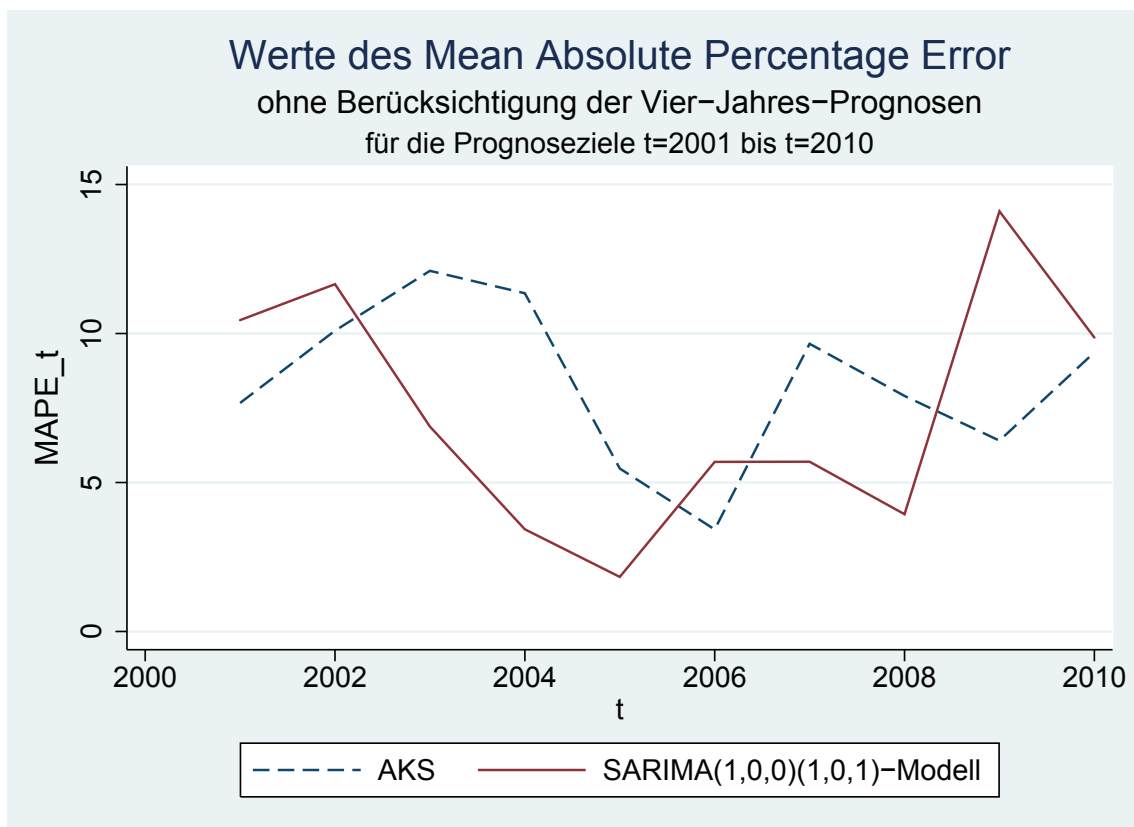


Abbildung 4.5.: Abbildung der Werte des $MAPE_t$ der AKS-Prognosen sowie der Prognosen des $SARIMA(1,0,0)(1,0,1)_4$ -Modells gemäß Methode 2 ohne Berücksichtigung der Vier-Jahres-Prognosen

Vertikale Treffsicherheitsanalyse

Als Nächstes wird die Prognosegüte wieder in Abhängigkeit vom Prognosehorizont untersucht (**fixed horizon**). Tabelle 4.39 zeigt dementsprechend die Werte des $MAPE_h$ und des $RMSPE_h$ der AKS-Prognosen für die Prognosehorizonte $h = 1, \dots, 5$ sowie beider SARIMA-Modelle für $h = 2, \dots, 5$. Außerdem ist für diese Prognoseho-

izonte wieder der Theilsche Ungleichheitskoeffizient U'_h aufgeführt. Zur Berechnung aller Maße werden die Prognosen der Jahre 2001 bis 2010 berücksichtigt, wobei die Quartalsprognosen zuvor gemäß Methode 2 zu Jahresprognosen kumuliert wurden.

Tabelle 4.39.: Vertikale Treffsicherheitsanalyse gemäß Methode 2

| Gütemaß | Prognosehorizont h | AKS | SARIMA M1 | SARIMA M2 |
|-----------|--------------------|-------|-----------|-----------|
| $MAPE_h$ | 1 | 1,96 | – | – |
| | 2 | 5,64 | 3,78 | 3,77 |
| | 3 | 9,37 | 7,17 | 7,17 |
| | 4 | 10,03 | 11,10 | 11,07 |
| | 5 | 11,01 | 14,25 | 14,20 |
| $RMSPE_h$ | 1 | 2,40 | – | – |
| | 2 | 6,33 | 4,74 | 4,73 |
| | 3 | 9,97 | 8,33 | 8,33 |
| | 4 | 11,45 | 12,87 | 12,86 |
| | 5 | 12,68 | 16,82 | 16,80 |
| U'_h | 2 | | 1,34 | 1,34 |
| | 3 | | 1,20 | 1,20 |
| | 4 | | 0,89 | 0,89 |
| | 5 | | 0,75 | 0,75 |

Werte des auf zwei Nachkommastellen gerundeten mittleren absoluten Prozentfehlers $MAPE_h$ sowie der Wurzel aus dem mittleren quadratischen Prozentfehler $RMSPE_h$ der AKS-Prognosen (AKS) sowie der Prognosen des bevorzugten $SARIMA(1,0,0)(1,0,1)_4$ -Modells (SARIMA M1) und des alternativen $SARIMA(0,0,1)(1,0,1)_4$ -Modells (SARIMA M2) für die Prognosehorizonte h . Der Theilsche Ungleichheitskoeffizient U'_h wird gemäß Formel (4.49) für beide SARIMA-Modelle berechnet. Die Quartalsprognosen wurden gemäß Methode 2 kumuliert (vgl. Kapitel 4.4.1). Bei der Berechnung des $MAPE_h$ sowie $RMSPE_h$ werden jeweils die Prognosen der Jahre 2001 bis 2010 berücksichtigt.

Unter Verwendung aller verfügbarer Informationen (Methode 2) erzielen die SARIMA-Schätzungen für alle Prognosehorizonte $h = 2, \dots, 5$ eine höhere Treffsicherheit gemäß der Werte des $MAPE_h$ und des $RMSPE_h$ als bei Kumulation der Quartalsprognosen gemäß Methode 1 (vgl. Tabelle 4.35).

Während bei gleichem Informationsstand zum möglichst ähnlichen Prognosezeitpunkt (Methode 1) die Prognosen des Arbeitskreises „Steuerschätzungen“ für alle Prognosehorizonte $h = 2, 3, 4$ eine höhere Prognosegüte als die Schätzungen der beiden Zeitreihenmodelle aufweisen, ist nun für die **SARIMA-Prognosen für das kommende Jahr ($h = 2$) sowie für die Zwei-Jahres-Prognosen ($h = 3$) eine höhere Treffsicherheit als für die AKS-Prognosen** festzustellen. Dementsprechend nimmt der Theilsche Ungleichheitskoeffizient U'_h für diese Prognosehorizonte auch jeweils Werte größer eins an. Für die mittelfristige Schätzdistanz von drei bzw. vier Jahren ($h = 4$ und $h = 5$) prognostizieren dagegen die offiziellen Schätzungen des AKS wieder besser als die SARIMA-Modelle.

Dieses Ergebnis bestätigt die in der Literatur verbreitete Auffassung, dass sich die direkten stochastischen Zeitreihenmodelle hauptsächlich für die kurzfristige Prognose hervorragend eignen. Ein ähnliches Ergebnis erzielte beispielsweise auch schon Flascha (1985) für den Untersuchungszeitraum 1969 bis 1984 (vgl. Kapitel 2.3 bzw. Flascha (1985), S. 169).

Für das alternative SARIMA(0,0,1)(1,0,1)₄-Modell ist dabei wieder eine tendenziell leicht höhere Treffsicherheit als für das bevorzugte SARIMA(1,0,0)(1,0,1)₄-Modell festzustellen. Der Unterschied in der Prognosegüte beider Zeitreihenmodelle ist jedoch nach wie vor nicht besonders groß.

Außerdem nimmt die Prognoseungenauigkeit, wie erwartet, erneut für alle Modellvarianten bei wachsendem Prognosehorizont zu. Abbildung 4.6 stellt die beschriebenen Ergebnisse der vertikalen Treffsicherheitsanalyse gemäß Methode 2 beispielhaft anhand der Werte des MAPE_h für die AKS-Prognosen sowie für die Prognosen des SARIMA(0,0,1)(1,0,1)₄-Modells graphisch dar.

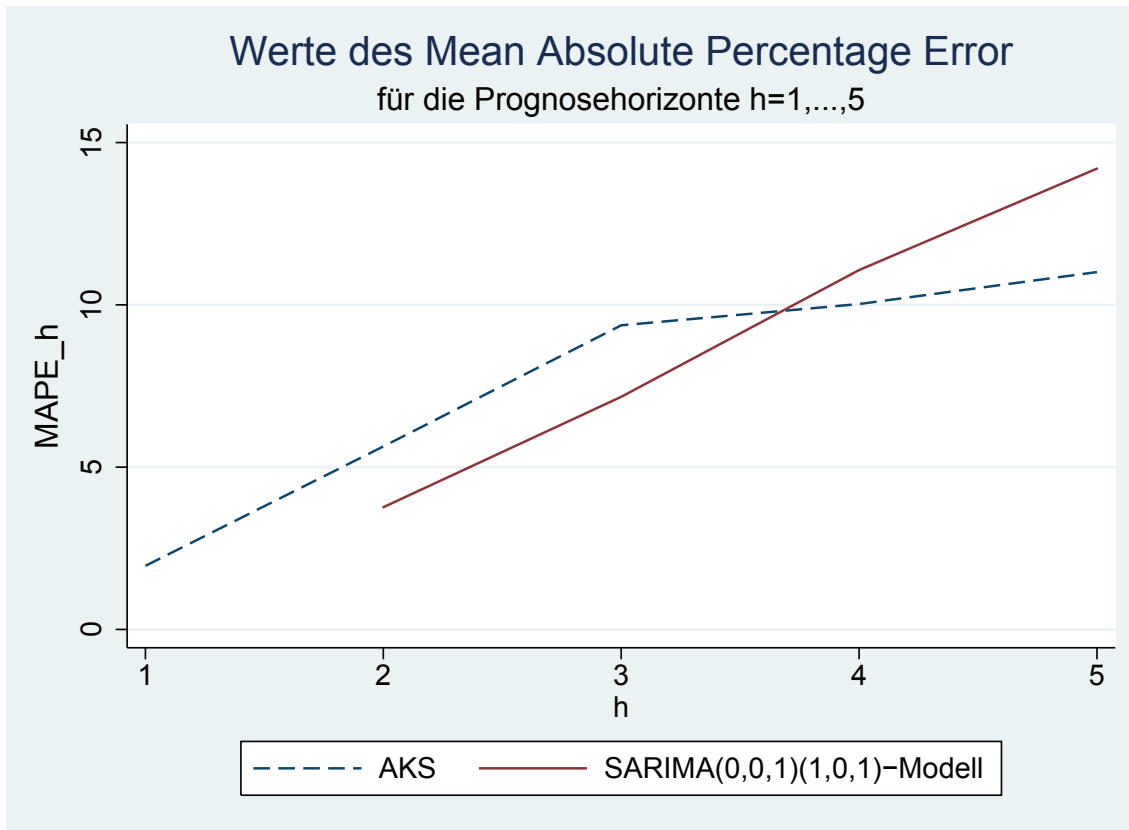


Abbildung 4.6.: Abbildung der Werte des $MAPE_h$ der AKS-Prognosen sowie der Prognosen des $SARIMA(0,0,1)(1,0,1)_4$ -Modells gemäß Methode 2

Gepoolte Treffsicherheitsanalyse

Wird die Treffsicherheitsanalyse für alle Prognoseziele und Prognosehorizonte zusammen durchgeführt, können die Vier-Jahres-Prognosen wiederum zunächst bei der Berechnung der deskriptiven Gütemaße berücksichtigt werden. Um den Einfluss der beiden unterschiedlichen Methoden der Kumulation der Quartalsprognosen zu Jahresprognosen zu isolieren, wird die Analyse anschließend wieder ohne Berücksichtigung der Prognosen mit $h = 5$ durchgeführt.

Werden also die **Vier-Jahres-Prognosen** wieder zur Berechnung verwendet, ergeben sich die in Tabelle 4.40 dargestellten Werte des MAPE und des RMSPE der AKS-Prognosen sowie der Prognosen beider SARIMA-Modelle. Außerdem ist für beide SARIMA-Modelle erneut jeweils der Theilsche Ungleichheitskoeffizient U' aufgeführt. Die Quartalsprognosen wurden gemäß Methode 2 zu Jahresprognosen kumuliert. Zur Berechnung aller Maße wurden jeweils wieder die Prognosen der Jahre 2001 bis 2010 und die Prognosehorizonte $h = 2, \dots, 5$ berücksichtigt.

Tabelle 4.40.: **Gepoolte Treffsicherheitsanalyse gemäß Methode 2 mit Vier-Jahres-Prognosen**

| Gütemaß | AKS | SARIMA M1 | SARIMA M2 |
|---------|-------|-----------|-----------|
| MAPE | 9,01 | 9,08 | 9,05 |
| RMSPE | 10,38 | 11,62 | 11,61 |
| U' | | 0,89 | 0,89 |

Werte des auf zwei Nachkommastellen gerundeten mittleren absoluten Prozentfehlers MAPE sowie der Wurzel aus dem mittleren quadratischen Prozentfehler RMSPE der AKS-Prognosen (AKS) sowie der Prognosen des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells (SARIMA M1) und des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells (SARIMA M2). Der Theilsche Ungleichheitskoeffizient U' wird gemäß Formel (4.50) für beide SARIMA-Modelle berechnet. Die Quartalsprognosen wurden gemäß Methode 2 kumuliert (vgl. Kapitel 4.4.1). Bei der Berechnung des MAPE sowie RMSPE werden jeweils die Prognosen der Prognosehorizonte $h = 2, \dots, 5$ sowie der Jahre 2001 bis 2010 berücksichtigt.

Auch unter Verwendung aller verfügbarer Informationen erzielen die als Alternative vorgeschlagenen SARIMA-Schätzungen insgesamt immer noch eine leicht niedrigere Treffsicherheit als die AKS-Prognosen, wenn die Vier-Jahres-Prognosen bei der Berechnung der deskriptiven Gütemaße berücksichtigt werden. Die Werte des MAPE unterscheiden sich allerdings nur noch auf der zweiten Nachkommastelle. So beträgt sowohl der mittlere absolute Prozentfehler der AKS-Prognosen als auch derjenige der Zeitreihenmodelle ungefähr neun Prozent. Werden große Abweichungen stärker

gewichtet als kleine, so resultiert auf Basis des RMSPE eine **leicht höhere Treffsicherheit** der Prognosen des Arbeitskreises „Steuerschätzungen“, so dass auch der Theilsche Ungleichheitskoeffizient U' jeweils einen Wert kleiner eins annimmt.

Außerdem ist erneut für die beiden alternativen SARIMA-Modelle kein großer Unterschied in der Prognosegüte festzustellen. Die Prognosen anhand des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells scheinen lediglich eine tendenziell leicht höhere Treffsicherheit als diejenigen anhand des im Box-Jenkins-Verfahren bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells aufzuweisen.

Um den Einfluss der beiden unterschiedlichen Methoden der Kumulation der Quartalsprognosen zu Jahresprognosen zu isolieren, wird abschließend noch die gepoolte Treffsicherheit der Prognosen gemäß Methode 2 **ohne Berücksichtigung der Vier-Jahres-Prognosen** untersucht. Dabei ergeben sich die in Tabelle 4.41 dargestellten Werte des MAPE und des RMSPE der AKS-Prognosen sowie der Prognosen beider SARIMA-Modelle. Außerdem ist für beide SARIMA-Modelle erneut jeweils der Theilsche Ungleichheitskoeffizient U' aufgeführt. Die Quartalsprognosen wurden gemäß Methode 2 zu Jahresprognosen kumuliert. Zur Berechnung aller Maße wurden jeweils die Prognosen der Jahre 2001 bis 2010 und nur die Prognosehorizonte $h = 2, 3, 4$ berücksichtigt.

Während bei Kumulation der Quartalsprognosen gemäß Methode 1 der mittlere absolute Prozentfehler der SARIMA-Prognosen über 10,5 Prozent beträgt (vgl. Tabelle 4.36), erzielen die SARIMA-Schätzungen unter Verwendung aller verfügbarer Informationen (Methode 2) für die Ein-, Zwei- und Drei-Jahres-Prognosen einen Wert des MAPE unter 7,5 Prozent, was allerdings dennoch große absolute Über- bzw. Unterschätzungen des Steueraufkommens der BRD widerspiegelt.

Damit liegt aber nicht nur der Wert des jeweiligen mittleren absoluten Prozentfehlers beider SARIMA-Schätzungen um einen Prozentpunkt unter demjenigen der AKS-Prognosen. Auch die Werte des RMSPE weisen auf eine **leicht höhere Treff-**

Tabelle 4.41.: **Gepoolte Treffsicherheitsanalyse gemäß Methode 2 ohne Vier-Jahres-Prognosen**

| Gütemaß | AKS | SARIMA M1 | SARIMA M2 |
|---------|------|-----------|-----------|
| MAPE | 8,35 | 7,35 | 7,34 |
| RMSPE | 9,50 | 9,26 | 9,26 |
| U' | | 1,03 | 1,03 |

Werte des auf zwei Nachkommastellen gerundeten mittleren absoluten Prozentfehlers MAPE sowie der Wurzel aus dem mittleren quadratischen Prozentfehler RMSPE der AKS-Prognosen (AKS) sowie der Prognosen des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells (SARIMA M1) und des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells (SARIMA M2). Der Theilsche Ungleichheitskoeffizient U' wird gemäß Formel (4.50) für beide SARIMA-Modelle berechnet. Die Quartalsprognosen wurden gemäß Methode 2 kumuliert (vgl. Kapitel 4.4.1). Bei der Berechnung des MAPE sowie RMSPE werden jeweils nur die Prognosen der Prognosehorizonte $h = 2, 3, 4$ sowie der Jahre 2001 bis 2010 berücksichtigt.

sicherheit der SARIMA-Prognosen hin, so dass auch der Theilsche Ungleichheitskoeffizient U' jeweils einen Wert knapp größer eins annimmt.

Dabei spielt es erneut keine große Rolle, ob das im Box-Jenkins-Verfahren bevorzugte SARIMA(1,0,0)(1,0,1)₄-Modell oder das alternative SARIMA(0,0,1)(1,0,1)₄-Modell der Prognose zugrunde liegt. Insgesamt kann bei Nutzung aller verfügbarer Informationen der SARIMA-Prognosen von einer mindestens vergleichbaren Prognosegüte der Zeitreihenmodelle sowie der offiziellen AKS-Prognosen ausgegangen werden, wenn die Vier-Jahres-Prognosen nicht berücksichtigt werden.

4.4.3. Fazit zum Vergleich der Prognosegenauigkeit anhand deskriptiver Gütemaße

Zusammenfassend lässt sich festhalten, dass die Ergebnisse der deskriptiven Analyse stark von der Vorgehensweise der Kumulation der Quartalsprognosen zu Jahresprognosen abhängen. Wird gemäß Methode 1 der **gleiche Informationsstand zum möglichst ähnlichen Prognosezeitpunkt** vorausgesetzt, erzielen die **Prognosen des Arbeitskreises „Steuerschätzungen“ fast durchweg eine höhere Prognosegenauigkeit** als die Schätzungen der als Alternative vorgeschlagenen SARIMA-Modelle. Lediglich für die Prognoseziele 2004, 2007 und 2008 weisen die Werte des MAPE_t beziehungsweise des RMSPE_t auf eine höhere Treffsicherheit der SARIMA-Prognosen hin.

Werden die Quartalsprognosen jedoch gemäß Methode 2 zu Jahresprognosen kumuliert, so dass also **alle verfügbaren Informationen zu allen Prognosezeitpunkten** für die Steuerschätzung genutzt werden, verbessert sich die Prognosegenauigkeit der SARIMA-Schätzungen erheblich. Allein für die Jahre 2001 und 2007 verschlechtert sich die Treffsicherheit der Zeitreihenmodelle, wenn die Jahresprognosen gemäß Methode 2 gebildet werden.

Außerdem ist sowohl gemäß Methode 1 als auch gemäß Methode 2 für das alternative $\text{SARIMA}(0,0,1)(1,0,1)_4$ -Modell eine tendenziell leicht höhere Treffsicherheit als für das im Box-Jenkins-Verfahren bevorzugte $\text{SARIMA}(1,0,0)(1,0,1)_4$ -Modell festzustellen. Der Unterschied in der Prognosegüte beider Zeitreihenmodelle ist jedoch vernachlässigbar klein.

Insgesamt kann von einer **vergleichbaren Prognosegüte der AKS-Prognosen sowie der SARIMA-Schätzungen** ausgegangen werden, wenn für letztere alle verfügbaren Informationen genutzt werden. Dabei zeigt sich deutlich, dass sich die direkten stochastischen Zeitreihenmodelle vor allem für die kurzfristige Prognose eignen, während der AKS für längere Schätzdistanzen besser prognostiziert. So

weisen die Werte des MAPE_h beziehungsweise des RMSPE_h auf eine **höhere Treffsicherheit der SARIMA-Prognosen für die Prognosehorizonte $h = 2$ und $h = 3$** hin. Für die Drei- und Vier-Jahres-Prognosen ($h = 4$ und $h = 5$) bescheinigen die berechneten deskriptiven Gütemaße jedoch den AKS-Prognosen eine höhere Prognosegenauigkeit.

Somit sind die vorgeschlagenen SARIMA-Modelle zumindest als Kontrollschätzung, wenn nicht sogar als bessere Alternative für die AKS-Prognosen des kommenden Jahres sowie für die Zwei-Jahres-Prognosen zu empfehlen. Allerdings bleibt einzuwenden, dass sich die Prognosegüte der direkten stochastischen Zeitreihenmodelle erheblich verschlechtert, wenn am Ende des Schätzzeitraums ein exogener Schock wie beispielsweise die vergangene Finanzmarktkrise auftritt.

5. Fazit und Ausblick

Die Steuereinnahmen stellen für Bund, Länder und Kommunen die Haupteinnahmequelle dar, so dass ihrer Vorausschätzung eine dementsprechend große finanzpolitische Bedeutung gebührt. In der Bundesrepublik Deutschland werden seit 1955 die Prognosen des Steueraufkommens vom Arbeitskreis „Steuerschätzungen“ erstellt. Im Mai erfolgen die Steuerschätzungen für den mittelfristigen Zeitraum, das heißt für das laufende Jahr sowie die vier Folgejahre. Kurz vor Abschluss der Haushaltsberatungen im November finden sowohl eine Überprüfung der Schätzungen für das laufende Jahr als auch eine Überarbeitung der Prognose für das kommende Jahr statt. Diese Steuerschätzungen bilden dann die Basis für die Ansätze des Haushaltsgesetzes des Bundes (vgl. Bundesministerium der Finanzen (2005), S. 11).

Durch die Mitarbeit vieler unabhängiger Experten sollen die Steuerschätzungen allein auf fachlichen Erwägungen basieren und frei von politischer Einflussnahme sein. So erarbeiten zunächst acht der Mitglieder, nämlich die fünf Wirtschaftsforschungsinstitute, die Bundesbank, der Sachverständigenrat und das BMF jeweils unabhängig voneinander mit eigenen Methoden und Modellen Schätzvorschläge für jede Einzelsteuer. Diese werden vor der Öffentlichkeit geheim gehalten und bei der Zusammenkunft des AKS so lange diskutiert und angepasst, bis ein Konsens aller Mitglieder erreicht wird. Allerdings sind die Schätzungen des AKS an die von der Bundesregierung vorgegebenen Grundannahmen der gesamtwirtschaftlichen Entwicklung gebunden (vgl. Bundesministerium der Finanzen (2005), S. 10). Diese so genannten Eckdaten für die kurzfristige Wirtschaftsentwicklung stellen jedoch Zielprojektionen der Bundesregierung dar. Somit ist die mit der Gründung des AKS eigentlich

angestrebte Objektivierung der Steuerschätzung und Unabhängigkeit von tagespolitischen Ziel- und Wunschvorstellungen fraglich. Jedenfalls besteht das Risiko, dass allein schon durch den Zielcharakter der gesamtwirtschaftlichen Projektion mehr von der angestrebten als der zu erwartenden Entwicklung bei den Prognosen ausgegangen wird, was zu einer Tendenz der Überschätzung des Steueraufkommens führen könnte (vgl. z. B. Gebhardt (2001), S. 129). Die vorliegende Arbeit untersuchte daher zum einen die Fragestellung, ob die offiziellen Schätzungen des AKS die für eine Prognose wünschenswerten Eigenschaften der Unverzerrtheit und Informations-effizienz erfüllen. Zum anderen wurden saisonale ARIMA-Modelle als alternative Schätzmethode vorgeschlagen und untersucht, wie treffsicher diese im Vergleich zu den AKS-Prognosen sind.

Dazu wurde im **Grundlagenkapitel** zunächst der institutionelle Rahmen des AKS, seine Zielsetzung sowie seine Vorgehensweise bei der Erarbeitung der Steuerprognosen vorgestellt. Daraufhin wurde die Vielzahl der Bestimmungsfaktoren des Steueraufkommens gemäß Leibrecht (2003) in die Ist- und Soll-Determinanten der Präzision der Steuerprognose systematisiert und im Detail erläutert. Anschließend wurden die zur Wahl stehenden qualitativen sowie direkten und indirekten quantitativen Prognosemethoden vorgestellt. Des Weiteren wurden die bei der Steuerprognose auftretenden Probleme in Abhängigkeit von der Wahl der Prognosemethode und der zu prognostizierenden Steuerart aufgeführt. Der Arbeitskreis „Steuerschätzungen“ verwendet vorwiegend indirekte steuerschuldorientierte Verfahren, für die zahlreiche Input-Daten erforderlich sind, die überwiegend selbst Prognosen mit entsprechenden Prognosefehlern darstellen (vgl. von der Lippe (1986)).

Bevor die saisonalen ARIMA-Modelle als alternatives direktes Verfahren vorgeschlagen und geschätzt wurden, erfolgte als Abschluss des zweiten Kapitels ein Überblick über die bisherigen empirischen Untersuchungen zur Qualität der AKS-Prognosen. Zusammenfassend kann festgehalten werden, dass die offiziellen Steuerschätzungen des Arbeitskreises „Steuerschätzungen“ gemäß bisheriger Studien als unverzerrt und weitgehend effizient eingestuft werden können (vgl. Becker und Buettner (2007),

Büttner und Kauder (2008b) und Lehmann (2010)). Dabei gilt allerdings zu beachten, dass alle Autoren nach traditionellem Ansatz separate Tests für die jeweils untersuchten Prognosehorizonte durchführten. Außerdem bescheinigen die empirischen Analysen der letzten 15 Jahre den AKS-Prognosen eine relativ hohe Treffsicherheit (vgl. Gebhardt (2001) und Büttner und Kauder (2008b)). Dennoch beobachteten die Autoren häufiger Überschätzungen des Steueraufkommens als Unterschätzungen (vgl. Gebhardt (2001), Bundesministerium der Finanzen (2005) und Lehmann (2010)). Außerdem kam es immer wieder zu Jahren mit so extremen Überschätzungen, dass sich diese enorm auf die Finanzhaushalte aller Ebenen auswirkten und politische Maßnahmen bis hin zu Haushaltssperren im laufenden Fiskaljahr notwendig machten. Insgesamt fordern fast alle Autoren eine Steigerung der Prognosequalität durch weitere Verbesserungen der Schätzmethoden (vgl. Flascha (1985), von der Lippe (1986) und Gebhardt (2001)). Des Weiteren erzielten Flaschas kurz- und mittelfristige ARIMA-Prognosen für das Gesamtsteueraufkommen im Zeitraum 1969 bis 1984 eine größere Treffsicherheit als die offiziellen AKS-Schätzungen.

Im **dritten Kapitel** wurde zur Prognose des Gesamtsteueraufkommens der Bundesrepublik Deutschland ebenfalls diese direkte stochastische Methode als Alternative vorgeschlagen. Die univariate Zeitreihenanalyse versucht eine Variable allein durch die Gesetzmäßigkeiten ihrer in der Vergangenheit beobachteten Realisationen zu erklären und daraus zukünftige Werte dieser Variablen zu prognostizieren. Dadurch werden weder vergangene noch für die Zukunft prognostizierte Werte erklärender Variablen benötigt. Außerdem besteht keine Notwendigkeit, die komplexe Steuerstruktur, sich ändernde steuerrechtliche Gegebenheiten bzw. die vielfältigen Interdependenzen mit den konjunkturellen Größen in einem ökonometrischen Modell abzubilden. Zunächst wurden die der univariaten Zeitreihenanalyse zugrunde liegenden saisonalen ARIMA-Modelle schrittweise entwickelt. Sodann folgte die Beschreibung der Datengrundlage sowie -aufbereitung. Als Datenquelle dienten die vom Statistischen Bundesamt bereitgestellten Tabellen der Quartalsdaten der kassemäßigen Steuereinnahmen der Bundesrepublik Deutschland für die Jahre 1961 bis 2010. Um einen möglichen Strukturbruch durch die deutsche Wiedervereinigung

zu modellieren, wurde außerdem eine Dummyvariable generiert, die für alle Quartale ab dem Jahr 1991 den Wert eins annimmt. Des Weiteren wurden aufgrund der Währungsumstellung alle Werte bis 2002 in Euro umgerechnet. Somit lag der weiteren Analyse die Zeitreihe des Gesamtsteueraufkommens der BRD für alle Quartale von 1961 bis 2010 in Millionen Euro zugrunde.

Daraufhin wurden anhand der Box-Jenkins-Methode eigene Steuerprognosen erstellt. Um eine stationäre Zeitreihe zu erzeugen, waren als Datentransformationen einerseits die Bildung des natürlichen Logarithmus der Zeitreihe sowie andererseits die Bildung der ersten Differenz des logarithmierten Gesamtsteueraufkommens nötig. Für die dadurch näherungsweise generierten Wachstumsraten des Gesamtsteueraufkommens wurden anhand des Box-Jenkins-Verfahrens mögliche Modellspezifikationen abgeleitet und zahlreiche Modellvarianten geschätzt. Gemäß der Informationskriterien stellte sich das SARIMA(1,0,0)(1,0,1)₄-Modell als das beste Modell heraus, gefolgt vom SARIMA(0,0,1)(1,0,1)₄-Modell. Als Vergleichsmodell zu den beiden sparsam parametrisierten Modellen wurde außerdem das SARIMA(2,0,1)(1,0,1)₄-Modell ausgewählt. Nach Überprüfung der Residuen wurden sodann für jede Modellvariante unter Verwendung eines Holdout-Sets dynamische Prognosen erstellt. Dabei wurde der Prognosezeitpunkt schrittweise um ein Quartal erhöht, so dass die dadurch entstandene Prognosematrix die Prognosen für $T = 40$ Prognoseziele für jeweils $h = 1$ bis $h = 16$ Prognosehorizonte umfasste. Somit standen der weiteren Analyse Prognosewerten des Gesamtsteueraufkommens (in Milliarden Euro) der BRD für 40 Quartale bzw. zehn Jahre, nämlich vom ersten Quartal 2001 bis zum vierten Quartal 2010 bei einer wachsenden Schätzdistanz von bis zu 16 Quartalen beziehungsweise vier Jahren zur Verfügung. Außerdem wurden auf die gleiche Art und Weise Prognosematrizen für einen maximalen Prognosehorizont von $H = 12$ (Drei-Jahres-Prognosen) sowie für $H = 8$ (Zwei-Jahres-Prognosen) und für einen maximalen Prognosehorizont von $H = 4$ (Ein-Jahres-Prognosen) erstellt. Aufgrund von Schätzproblemen musste das höher parametrisierte SARIMA(2,0,1)(1,0,1)₄-Modell im Weiteren als Vergleichsmodell verworfen werden.

Um die Prognosegüte der beiden anderen ausgewählten SARIMA-Modelle beurteilen zu können, wurde im **vierten Kapitel** zunächst der gepoolte Modellansatz gemäß Clements et al. (2007) zur Analyse der Prognoserationalität, das heißt der Unverzerrtheit und Informationseffizienz, vorgestellt. Während bei der traditionellen Analyse der Prognoserationalität nach Mincer und Zarnowitz (1969) sowie Holden und Peel (1990) typischerweise nur Prognosen mit derselben Schätzdistanz für unterschiedliche Prognoseziele untersucht werden, stellen Clements et al. (2007) hingegen einen Ansatz vor, der die Prognosen für unterschiedliche Prognoseziele über alle Prognosehorizonte zusammenführt und gemeinsam untersucht. Ausgangspunkt ist dabei die Zerlegung des Prognosefehlers in die zu testende Verzerrung, den idiosynkratischen Störterm, sowie die aggregierten makro-ökonomischen Schocks. Darauf aufbauend wird die Nullhypothese getestet, dass die Prognosen für alle Prognosehorizonte unverzerrt sind. Als Alternativhypothese kommen zwei unterschiedliche Varianten in Frage. In Variante I wird angenommen, dass die Verzerrung für alle Prognosehorizonte gleich groß sein muss, während in Variante II horizont-spezifische Verzerrungen zugelassen werden.

Des Weiteren werden nochmals zwei Modellalternativen unterschieden je nachdem, welche Annahme bezüglich des idiosynkratischen Störterms getroffen wird. Entweder wird dieser unter Annahme A als prognose-spezifischer idiosynkratischer Fehler interpretiert, der beispielsweise durch ineffiziente Nutzung von Informationen, Messfehler oder Fehlspezifikationen im Prognosemodell hervorgerufen wird (vgl. Clements et al. (2007), S. 124), oder aber unter Annahme B komplett weggelassen. Die Nullhypothese der Unverzerrtheit wurde bei allen durchgeführten Tests sowohl für Variante I als auch Variante II unter jeweils beiden Annahmen getestet, um die Robustheit des Testergebnisses zu überprüfen. Außerdem wurde die Korrelationsstruktur der Prognosen bzw. Prognosefehler für unterschiedliche maximale Prognosehorizonte ($H = 4, 8, 12, 16$) entwickelt. Ferner wurde beim Test auf Informationseffizienz als Fall 1 einerseits die Informationsmenge der Änderungen gleichbleibender aufeinanderfolgender Prognosehorizonte für verschiedene vergangene Prognoseziele

betrachtet. Andererseits wurde als Fall 2 die Informationsmenge der vergangenen Änderungen der Prognosen desselben Prognoseziels berücksichtigt.

Sowohl die Prognosen des Gesamtsteueraufkommens der BRD anhand des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells als auch diejenigen anhand des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells erwiesen sich bei einem maximalen Prognosehorizont von bis zu drei Jahren ($H = 4, 8, 12$) für alle betrachteten Modellvarianten sowie unter allen Annahmen als unverzerrt. Die Testentscheidungen für die Vier-Jahres-Prognosen hingen dagegen von den Modellannahmen ab, so dass für den maximalen Prognosehorizont $H = 16$ keine robusten Testergebnisse resultierten. Außerdem erwiesen sich die Prognosen mit einem maximalen Prognosehorizont von bis zu zwei Jahren ($H = 4, 8$) für beide untersuchten Informationsmengen und unter allen betrachteten Modellannahmen als effizient. Wird der maximale Prognosehorizont jedoch auf $H = 12$ erhöht, hat die Wahl der zu Grunde liegenden Informationsmenge einen maßgeblichen Einfluss auf die Testentscheidung. So implizierten die Ergebnisse, dass die Drei-Jahres-Prognosen auch dann noch effizient sind, wenn die Änderungen der Prognosen gleichbleibender aufeinanderfolgender Prognosehorizonte für verschiedene vergangene Prognoseziele (Fall 1) als zugrunde liegende Informationsmenge berücksichtigt wird. Wird aber als Informationsmenge die vergangenen Änderungen der Prognosen desselben Prognoseziels (Fall 2) betrachtet, erwiesen sich die SARIMA-Prognosen mit einem maximalen Prognosehorizont von $H = 12$ als ineffizient.

Somit konnten die eigenen SARIMA-Prognosen des Gesamtsteueraufkommens der BRD bei einem maximalen Prognosehorizont von bis zu zwei Jahren als unverzerrt und effizient beurteilt werden, so dass sie im Sinne der beiden betrachteten Informationsmengen auch als rational bezeichnet werden können. Wird der maximale Prognosehorizont jedoch auf drei Jahre erhöht, erwiesen sich die unverzerrten SARIMA-Prognosen nur noch im Sinne der zuerst berücksichtigten Informationsmenge als effizient und damit auch als rational.

In der vorliegenden Arbeit wurden die Mai-Prognosen des Arbeitskreises „Steuerschätzungen“ erstmalig ebenfalls gemäß des gepoolten Ansatzes von Clements et al. (2007) auf ihre Prognoserationalität hin untersucht. Die Testentscheidungen, ob die AKS-Prognosen verzerrt sind, hingen dabei allerdings durchweg von den getroffenen Modellannahmen ab. Wurde in der Alternativhypothese nämlich unterstellt, dass die Verzerrung für alle Prognosehorizonte gleich groß sein muss (Variante I), konnte die Nullhypothese der Unverzerrtheit der Prognosen in allen betrachteten Fällen eindeutig nicht verworfen werden. Wenn in der Alternativhypothese hingegen horizont-spezifische Verzerrungen zugelassen wurden (Variante II), dann erwiesen sich die untersuchten Prognosen lediglich im Modell mit idiosynkratischem Störterm (Annahme A) als unverzerrt. Dabei resultierte allerdings ein negativer Schätzwert für dessen Varianz. Wurde diese deshalb gleich Null gesetzt (Annahme B), erwiesen sich die AKS-Prognosen dagegen als (negativ) verzerrt. Dieses Ergebnis widerspricht zwar den bisherigen empirischen Ergebnissen von Becker und Büttner (2007), Büttner und Kauder (2008b) und Lehmann (2010). Diese Autoren führten den traditionellen Test auf Unverzerrtheit aber lediglich einzeln für bestimmte Prognosehorizonte durch, anstatt für alle Prognosehorizonte gemeinsam zu testen.

Auch beim Test auf Informationseffizienz wurden für die Prognosen des Arbeitskreises „Steuerschätzungen“ keine robusten Testergebnisse erzielt. In den meisten Fällen erwiesen sich die AKS-Prognosen zwar als effizient, die Testentscheidungen hingen dabei allerdings von den getroffenen Modellannahmen, der Wahl des Signifikanzniveaus und in besonderem Maße von der genauen Auswahl der Stichprobe ab. Wurden nämlich alle Prognosen, die 1991 nach der Wiedervereinigung unter mangelhafter Datengrundlage erstellt worden sind, aus der Analyse ausgeschlossen, so verbesserten sich einerseits die Testergebnisse eindeutig, wenn als zugrunde liegende Informationsmenge die Änderungen der Prognosen gleichbleibender aufeinanderfolgender Prognosehorizonte für verschiedene vergangene Prognoseziele (Fall 1) betrachtet wurden. Wurden dagegen als zugrunde liegende Informationsmenge die vergangenen Änderungen der Prognosen desselben Prognoseziels (Fall 2) berücksichtigt, erwiesen sich die AKS-Prognosen teilweise nicht mehr als effizient, wenn alle

Prognosen, die 1991 erstellt worden sind, aus der Analyse ausgeschlossen wurden. Schon aufgrund der signifikanten Verzerrung im Modell mit den plausibelsten Annahmen sind die offiziellen Prognosen des Arbeitskreises „Steuerschätzungen“ somit gemäß des gepoolten Ansatzes nicht als rational einzustufen.

Abschließend wurde in Kapitel 4 anhand deskriptiver Gütemaße eine vergleichende horizontale, vertikale sowie gepoolte Treffsicherheitanalyse der SARIMA- sowie AKS-Prognosen für die Jahre 2001 bis 2010 durchgeführt. Dabei zeigte sich, dass die Ergebnisse stark von der Vorgehensweise der Kumulation der Quartalsprognosen zu Jahresprognosen abhängen. Wurde nämlich für die SARIMA-Prognosen der gleiche Informationsstand zum möglichst ähnlichen Prognosezeitpunkt wie für die Mai-Sitzungen des Arbeitskreises „Steuerschätzungen“ vorausgesetzt, erzielten die AKS-Prognosen fast durchweg eine höhere Prognosegenauigkeit als die Schätzungen der SARIMA-Modelle.

Wurde dagegen für die als Alternative vorgeschlagene SARIMA-Schätzmethode die Verwendung aller verfügbarer Informationen zu allen Prognosezeitpunkten erlaubt, verbesserte sich die Prognosegenauigkeit der SARIMA-Schätzungen in nahezu allen Fällen erheblich. Insgesamt kann dann von einer vergleichbaren Prognosegüte der AKS-Prognosen sowie der SARIMA-Schätzungen ausgegangen werden. Dabei zeigte sich außerdem deutlich, dass sich die direkten stochastischen Zeitreihenmodelle vor allem für die kurzfristige Prognose eignen, während der AKS für längere Schätzdistanzen besser prognostizierte. So wiesen die berechneten Gütemaße auf eine höhere Treffsicherheit der SARIMA-Prognosen für die Prognosehorizonte $h = 2$ und $h = 3$ hin, während sie für die Drei- und Vier-Jahres-Prognosen ($h = 4$ und $h = 5$) den AKS-Prognosen eine höhere Prognosegenauigkeit bescheinigten. Diesbezüglich wurde also das Ergebnis von Flascha (1985) für den Untersuchungszeitraum 2001 bis 2010 bestätigt. Des Weiteren zeigte sich, dass sich die Prognosegüte der direkten stochastischen Zeitreihenmodelle erheblich verschlechtert, wenn am Ende des Schätzzeitraums ein exogener Schock wie beispielsweise die vergangene Finanzmarktkrise auftritt.

Zusammenfassend kann also festgehalten werden, dass die signifikante (negative) Verzerrung der AKS-Prognosen im plausibelsten Modell die in der Literatur befürchtete Tendenz zur Überschätzung des Steueraufkommens durch den AKS (vgl. z. B. Gebhardt (2001), S. 129) bestätigt. Die Ursache dieser Verzerrung bleibt allerdings ungeklärt. Zukünftige Forschung könnte beispielsweise untersuchen, ob die Prognosefehler der vorgegebenen gesamtwirtschaftlichen Eckdaten die durchschnittliche Überschätzung des Steueraufkommens im gepoolten Modell nach sich ziehen. Außerdem wäre von Interesse, ob es bei der Steuerschätzung zu einer direkten politischen Einflussnahme kommt. Becker und Buettner (2007) stellten in ihrer Analyse der Mai-Prognosen für das laufende Jahr ($h = 1$) bereits fest, dass der AKS in Wahljahren das Gesamtsteueraufkommen tendenziell unterschätzt. Geprüft werden sollte darüber hinaus, ob sich dieses Ergebnis auch bestätigt, wenn gemäß des gepoolten Ansatzes alle Prognosehorizonte zusammen untersucht werden. Grundsätzlich wäre im Hinblick auf zukünftige Forschung eine Erhöhung der Transparenz in Bezug auf die vom AKS verwendeten Methoden der Steuerschätzung wünschenswert. Nur so könnten die veröffentlichten Endergebnisse von Außenstehenden nachvollzogen werden, und es könnten dementsprechend gezielte Anmerkungen und Empfehlungen für eine Verbesserung der Schätzmethode eingebracht werden.

Im Gegensatz zu den AKS-Prognosen erwiesen sich die mit der als Alternative vorgeschlagenen SARIMA-Schätzmethode erstellten Ein- und Zwei-Jahres-Prognosen unter allen getroffenen Annahmen als unverzerrt und effizient. Außerdem erzielten sie eine höhere Treffsicherheit als die offiziellen Steuerschätzungen, so dass die Anwendung der direkten stochastischen Zeitreihenmodelle für die kurze Frist von bis zu zwei Jahren generell durchaus zu empfehlen ist. Für die Drei- und Vier-Jahres-Prognosen bescheinigten die durchgeführten Analysen dagegen den Prognosen des Arbeitskreises „Steuerschätzungen“ eine höhere Prognosegüte, so dass die SARIMA-Modelle für die mittelfristige Steueraufkommensprognose demzufolge nicht als Alternative geeignet zu sein scheinen.

Im Hinblick auf zukünftige Forschung steht jedoch noch einerseits die Frage offen, inwiefern die Prognosegüte der SARIMA-Schätzungen von der Steuerart abhängt. So könnte beispielsweise geprüft werden, ob die direkten stochastischen Zeitreihenmodelle zur Prognose der Verbrauchssteuern besser geeignet sind als diejenigen des AKS. Des Weiteren könnte versucht werden, eine geeignete Variable zur Abbildung der wirtschaftlichen Entwicklung, wie z. B. ein Konjunkturindikator, in das Zeitreihenmodell mit aufzunehmen. Im einfachsten Fall könnte die Modellierung durch die so genannten (saisonalen) ARIMAX-Modelle erfolgen. Alternativ könnte untersucht werden, ob zwischen der Zeitreihe des Steueraufkommens und derjenigen der Konjunkturvariablen eine Kointegrationsbeziehung besteht. In diesem Fall sollten die Prognosen des Steueraufkommens anhand eines (Vektor-)Fehlerkorrekturmodells (VECM) ermittelt werden, um anschließend ihre Treffsicherheit mit derjenigen der AKS-Schätzungen zu vergleichen.

Bei der Analyse der Prognosegenauigkeit wäre außerdem noch von Interesse, wie sich die Annahme einer asymmetrischen Verlustfunktion auf die Ergebnisse auswirken würde. Schließlich stehen vor allem starke Überschätzungen des Steueraufkommens, die entweder zu drastischen Sparmaßnahmen oder zu einer höheren Staatsverschuldung als vorgesehen führen, bei der Öffentlichkeit und den Wählern besonders in der Kritik.

A. Ergebnistabellen und Schaubilder zu den SARIMA-Schätzungen

Tabelle A.1.: Schätzergebnisse des SARIMA(1,0,0)(1,0,1)₄-Modells

| Y_t | Koeffizient | Standardfehler | p -Wert |
|---------------------|-------------|----------------|-----------|
| c | 0,012 | 0,042 | 0,772 |
| Y_{t-1} | -0,233 | 0,055 | 0,000 |
| Y_{t-4} | 0,985 | 0,012 | 0,000 |
| ε_{t-4} | -0,563 | 0,061 | 0,000 |

Schätzergebnisse des SARIMA(1,0,0)(1,0,1)₄-Modells: $(1 - \phi_1 L)(1 - \phi_4 L^4)Y_t = c + (1 - \theta_4 L^4)\varepsilon_t$. Zur Bestimmung des Maximums sind 17 Iterationen notwendig. Die Stichprobe umfasst die 199 Wachstumsraten des Gesamtsteueraufkommens der BRD vom zweiten Quartal 1961 bis zum vierten Quartal 2010. Die Schätzwerte für die Koeffizienten der Konstanten, \hat{c} , des nicht-saisonalen AR(1)-Terms, $\hat{\phi}_1$, des saisonalen AR(1)-Terms, $\hat{\phi}_4$, des saisonalen MA(1)-Terms, $\hat{\theta}_4$, der dazugehörigen Standardfehler sowie p -Werte sind auf drei Nachkommastellen gerundet.

Tabelle A.2.: Schätzergebnisse des SARIMA(0,0,1)(1,0,1)₄-Modells

| Y_t | Koeffizient | Standardfehler | p -Wert |
|---------------------|-------------|----------------|-----------|
| c | 0,012 | 0,042 | 0,769 |
| ε_{t-1} | -0,225 | 0,049 | 0,000 |
| Y_{t-4} | 0,985 | 0,012 | 0,000 |
| ε_{t-4} | -0,571 | 0,062 | 0,000 |

Schätzergebnisse des SARIMA(0,0,1)(1,0,1)₄-Modells: $(1 - \phi_4 L^4)Y_t = c + (1 - \theta_1 L)(1 - \theta_4 L^4)\varepsilon_t$. Zur Bestimmung des Maximums sind 14 Iterationen notwendig. Die Stichprobe umfasst die 199 Wachstumsraten des Gesamtsteueraufkommens der BRD vom zweiten Quartal 1961 bis zum vierten Quartal 2010. Die Schätzwerte für die Koeffizienten der Konstanten, \hat{c} , des nicht-saisonalen MA(1)-Terms, $\hat{\theta}_1$, des saisonalen AR(1)-Terms, $\hat{\phi}_4$, des saisonalen MA(1)-Terms, $\hat{\theta}_4$, der dazugehörigen Standardfehler sowie p -Werte sind auf drei Nachkommastellen gerundet.

Tabelle A.3.: Schätzergebnisse des SARIMA(2,0,1)(1,0,1)₄-Modells

| Y_t | Koeffizient | Standardfehler | p -Wert |
|---------------------|-------------|----------------|-----------|
| c | 0,013 | 0,009 | 0,156 |
| Y_{t-1} | 0,711 | 0,064 | 0,000 |
| Y_{t-2} | 0,167 | 0,082 | 0,043 |
| ε_{t-1} | -0,977 | 0,032 | 0,000 |
| Y_{t-4} | 0,983 | 0,013 | 0,000 |
| ε_{t-4} | -0,513 | 0,068 | 0,000 |

Schätzergebnisse des SARIMA(2,0,1)(1,0,1)₄-Modells:
 $(1 - \phi_1 L)(1 - \phi_2 L)(1 - \phi_4 L^4)Y_t = c + (1 - \theta_1 L)(1 - \theta_4 L^4)\varepsilon_t$.
 Zur Bestimmung des Maximums sind 27 Iterationen notwendig. Die Stichprobe umfasst die 199 Wachstumsraten des Gesamtsteueraufkommens der BRD vom zweiten Quartal 1961 bis zum vierten Quartal 2010. Die Schätzwerte für die Koeffizienten der Konstanten, \hat{c} , des nicht-saisonalen AR(1)-Terms, $\hat{\phi}_1$, des nicht-saisonalen AR(2)-Terms, $\hat{\phi}_2$, des nicht-saisonalen MA(1)-Terms, $\hat{\theta}_1$, des saisonalen AR(1)-Terms, $\hat{\phi}_4$, des saisonalen MA(1)-Terms, $\hat{\theta}_4$, der dazugehörigen Standardfehler sowie p -Werte sind auf drei Nachkommastellen gerundet.

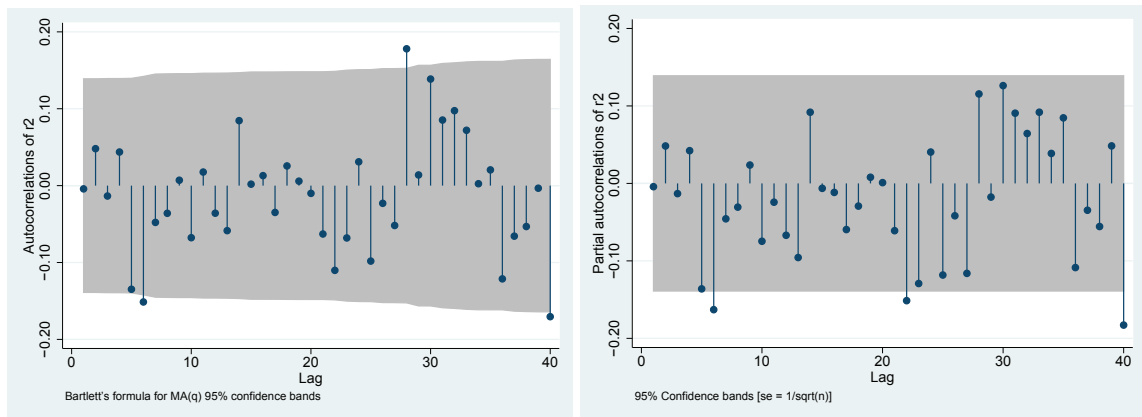


Abbildung A.1.: Abbildung der empirischen ACF sowie der empirischen PACF des SARIMA(0,0,1)(1,0,1)₄-Modells

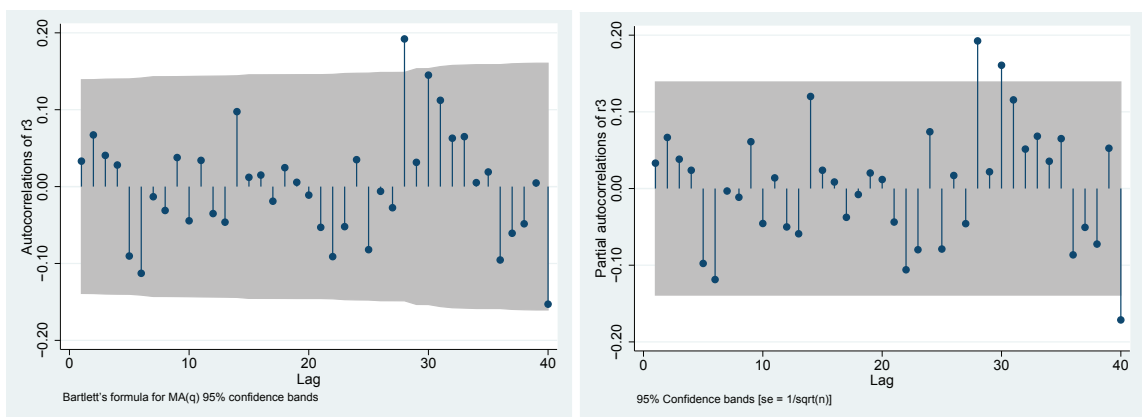


Abbildung A.2.: Abbildung der empirischen ACF sowie der empirischen PACF des SARIMA(2,0,1)(1,0,1)₄-Modells

B. Ergebnistabellen zur Analyse der SARIMA-Prognosen

Tabelle B.1.: **Regressionsergebnisse der horizont-spezifischen Verzerrungen der Prognosen anhand des bevorzugten Modells mit $H = 4$**

| h | Annahme A | | | Annahme B | |
|-----|------------------|----------------------|-----------|----------------------|-----------|
| | $\hat{\alpha}_h$ | $se(\hat{\alpha}_h)$ | p -Wert | $se(\hat{\alpha}_h)$ | p -Wert |
| 1 | -0,60 | 0,82 | 0,47 | 0,76 | 0,43 |
| 2 | -1,07 | 1,50 | 0,48 | 1,52 | 0,48 |
| 3 | -1,54 | 2,20 | 0,49 | 2,27 | 0,50 |
| 4 | -2,40 | 2,91 | 0,41 | 3,01 | 0,43 |

Regressionsergebnisse der horizont-spezifischen Verzerrungen beim Test auf Unverzerrtheit der Prognosen anhand des ersten Modells mit $H = 4$, anhand der Regressionsgleichung $e = (i_T \otimes I_H) \alpha^H + v$. Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für jedes h sind unter beiden Annahmen die Werte der geschätzten horizont-spezifischen Verzerrungen $\hat{\alpha}_h$, ihrer Standardfehler $se(\hat{\alpha}_h)$ sowie die entsprechenden p -Werte der einzelnen t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.2.: **Ergebnisse des t -Tests für das Alternativmodell und Variante I mit $H = 4$**

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -1,46 | -1,46 |
| $se(\hat{\alpha})$ | 1,80 | 1,88 |
| p -Wert | 0,42 | 0,44 |
| $\hat{\sigma}_\varepsilon^2$ | 6,70 | – |
| $\hat{\sigma}_u^2$ | 21,07 | 23,31 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, \dots, 4$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro anhand des SARIMA(0,0,1)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010.

Tabelle B.3.: **Ergebnisse des F -Tests für das Alternativmodell und Variante II mit $H = 4$**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,24 | 1,84 |
| p -Wert | 0,92 | 0,12 |
| $\hat{\sigma}_\varepsilon^2$ | 6,50 | – |
| $\hat{\sigma}_u^2$ | 20,98 | 23,14 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, \dots, 4$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro anhand des SARIMA(0,0,1)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010.

Tabelle B.4.: **Regressionsergebnisse der horizont-spezifischen Verzerrungen der Prognosen anhand des bevorzugten Modells mit $H = 8$**

| h | Annahme A | | | Annahme B | |
|-----|------------------|----------------------|-----------|----------------------|-----------|
| | $\hat{\alpha}_h$ | $se(\hat{\alpha}_h)$ | p -Wert | $se(\hat{\alpha}_h)$ | p -Wert |
| 1 | -0,56 | 0,84 | 0,51 | 0,78 | 0,47 |
| 2 | -1,06 | 1,57 | 0,50 | 1,56 | 0,50 |
| 3 | -1,57 | 2,31 | 0,50 | 2,32 | 0,50 |
| 4 | -2,48 | 3,05 | 0,42 | 3,08 | 0,42 |
| 5 | -3,42 | 3,79 | 0,37 | 3,84 | 0,37 |
| 6 | -4,17 | 4,52 | 0,36 | 4,59 | 0,36 |
| 7 | -4,44 | 5,25 | 0,40 | 5,33 | 0,40 |
| 8 | -4,83 | 5,96 | 0,42 | 6,06 | 0,43 |

Regressionsergebnisse der horizont-spezifischen Verzerrungen beim Test auf Unverzerrtheit der Prognosen anhand des ersten Modells mit $H = 8$, anhand der Regressionsgleichung $e = (i_T \otimes I_H) \alpha^H + v$. Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für jedes h sind unter beiden Annahmen die Werte der geschätzten horizont-spezifischen Verzerrungen $\hat{\alpha}_h$, ihrer Standardfehler $se(\hat{\alpha}_h)$ sowie die entsprechenden p -Werte der einzelnen t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.5.: **Ergebnisse des t -Tests für das Alternativmodell und Variante I mit $H = 8$**

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -2,86 | -2,86 |
| $se(\hat{\alpha})$ | 3,37 | 3,43 |
| p -Wert | 0,40 | 0,40 |
| $\hat{\sigma}_\varepsilon^2$ | 4,88 | – |
| $\hat{\sigma}_u^2$ | 23,75 | 24,61 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, \dots, 8$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(0,0,1)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.6.: **Ergebnisse des F -Tests für das Alternativmodell und Variante II mit $H = 8$**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,36 | 1,86 |
| p -Wert | 0,94 | 0,07 |
| $\hat{\sigma}_\varepsilon^2$ | 2,08 | – |
| $\hat{\sigma}_u^2$ | 23,86 | 24,23 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, \dots, 8$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(0,0,1)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.7.: **Regressionsergebnisse der horizont-spezifischen Verzerrungen der Prognosen anhand des bevorzugten Modells mit $H = 12$**

| h | Annahme A | | | Annahme B | |
|-----|------------------|----------------------|-----------|----------------------|-----------|
| | $\hat{\alpha}_h$ | $se(\hat{\alpha}_h)$ | p -Wert | $se(\hat{\alpha}_h)$ | p -Wert |
| 1 | -0,22 | 0,58 | 0,71 | 0,88 | 0,81 |
| 2 | -0,65 | 1,68 | 0,70 | 1,76 | 0,71 |
| 3 | -1,31 | 2,63 | 0,62 | 2,62 | 0,62 |
| 4 | -2,42 | 3,54 | 0,50 | 3,48 | 0,49 |
| 5 | -3,53 | 4,44 | 0,43 | 4,33 | 0,42 |
| 6 | -4,51 | 5,32 | 0,40 | 5,17 | 0,38 |
| 7 | -5,15 | 6,19 | 0,41 | 6,01 | 0,39 |
| 8 | -5,65 | 7,06 | 0,42 | 6,84 | 0,41 |
| 9 | -5,87 | 7,91 | 0,46 | 7,66 | 0,44 |
| 10 | -6,05 | 8,76 | 0,49 | 8,47 | 0,48 |
| 11 | -6,06 | 9,60 | 0,53 | 9,27 | 0,51 |
| 12 | -6,41 | 10,42 | 0,54 | 10,07 | 0,52 |

Regressionsergebnisse der horizont-spezifischen Verzerrungen beim Test auf Unverzerrtheit der Prognosen anhand des ersten Modells mit $H = 12$, anhand der Regressionsgleichung $e = (i_T \otimes I_H) \alpha^H + v$. Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für jedes h sind unter beiden Annahmen die Werte der geschätzten horizont-spezifischen Verzerrungen $\hat{\alpha}_h$, ihrer Standardfehler $se(\hat{\alpha}_h)$ sowie die entsprechenden p -Werte der einzelnen t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(1,0,0)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.8.: **Ergebnisse des t -Tests für das Alternativmodell und Variante I mit $H = 12$**

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -4,03 | -4,03 |
| $se(\hat{\alpha})$ | 5,70 | 5,50 |
| p -Wert | 0,48 | 0,46 |
| $\hat{\sigma}_\varepsilon^2$ | (-19,71) | – |
| $\hat{\sigma}_u^2$ | 33,55 | 31,18 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen über alle Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(0,0,1)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.9.: **Ergebnisse des F -Tests für das Alternativmodell und Variante II mit $H = 12$**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,19 | 1,42 |
| p -Wert | 0,999 | 0,15 |
| $\hat{\sigma}_\varepsilon^2$ | (-28,11) | – |
| $\hat{\sigma}_u^2$ | 34,10 | 30,73 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, \dots, 12$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise drei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(0,0,1)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.10.: **Ergebnisse des t -Tests für das Alternativmodell und Variante I mit $H = 16$**

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -4,88 | -4,88 |
| $se(\hat{\alpha})$ | 8,36 | 7,83 |
| p -Wert | 0,56 | 0,53 |
| $\hat{\sigma}_\varepsilon^2$ | (-59,34) | – |
| $\hat{\sigma}_u^2$ | 43,53 | 38,13 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen über alle Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(0,0,1)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.11.: **Ergebnisse des F -Tests für das Alternativmodell und Variante II mit $H = 16$**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,14 | 2,88 |
| p -Wert | 0,99997 | 0,00001 |
| $\hat{\sigma}_\varepsilon^2$ | (-71,79) | – |
| $\hat{\sigma}_u^2$ | 44,27 | 37,74 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, \dots, 16$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise fünf Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(0,0,1)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.12.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 1 der Prognosen mit $H = 4$ anhand des Alternativmodells**

| h | $\hat{\gamma}$ | Annahme A | | Annahme B | |
|-----|----------------|--------------------|-----------|--------------------|-----------|
| | | $se(\hat{\gamma})$ | p -Wert | $se(\hat{\gamma})$ | p -Wert |
| 2 | -0,01 | 0,27 | 0,97 | 0,22 | 0,96 |
| 3 | -0,07 | 0,26 | 0,80 | 0,21 | 0,75 |
| 4 | -0,09 | 0,16 | 0,60 | 0,14 | 0,53 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2, 3, 4$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für jedes h sind unter beiden Annahmen die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(0,0,1)(1,0,1)₄-Modells vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.13.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 1 der Prognosen mit $H = 4$ anhand des Alternativmodells**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | -0,06 | -0,06 | -0,07 | -0,07 |
| $se(\hat{\gamma})$ | 0,18 | 0,17 | 0,18 | 0,17 |
| p -Wert | 0,72 | 0,70 | 0,71 | 0,70 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 1). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(0,0,1)(1,0,1)₄-Modells mit $H = 4$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.14.: **Ergebnisse des separaten Tests auf Informationseffizienz gemäß Fall 2 der Prognosen mit $H = 4$ anhand des Alternativmodells**

| h | $\hat{\gamma}$ | Annahme A | | Annahme B | |
|-----|----------------|--------------------|-----------|--------------------|-----------|
| | | $se(\hat{\gamma})$ | p -Wert | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,06 | 0,16 | 0,70 | 0,14 | 0,64 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz für $h = 2$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t;h+1;h+2} + \delta + \omega_t$ (Fall 2). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Der Wert des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests sind für beide Annahmen auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(0,0,1)(1,0,1)₄-Modells mit $H = 4$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.15.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 1 der Prognosen mit $H = 8$ anhand des Alternativmodells**

| h | $\hat{\gamma}$ | Annahme A | | Annahme B | |
|-----|----------------|--------------------|-----------|--------------------|-----------|
| | | $se(\hat{\gamma})$ | p -Wert | $se(\hat{\gamma})$ | p -Wert |
| 2 | -0,005 | 0,26 | 0,98 | 0,22 | 0,98 |
| 3 | -0,063 | 0,25 | 0,80 | 0,22 | 0,77 |
| 4 | -0,059 | 0,16 | 0,71 | 0,14 | 0,66 |
| 5 | -0,055 | 0,17 | 0,75 | 0,14 | 0,71 |
| 6 | -0,153 | 0,18 | 0,41 | 0,16 | 0,33 |
| 7 | -0,149 | 0,18 | 0,40 | 0,15 | 0,33 |
| 8 | -0,167 | 0,13 | 0,21 | 0,11 | 0,14 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2, \dots, 8$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für jedes h sind unter beiden Annahmen die Werte des geschätzten Koeffizienten $\hat{\gamma}$ auf drei, diejenigen seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(0,0,1)(1,0,1)₄-Modells mit $H = 8$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.16.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 1 der Prognosen mit $H = 8$ anhand des Alternativmodells**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | -0,11 | -0,11 | -0,11 | -0,11 |
| $se(\hat{\gamma})$ | 0,14 | 0,14 | 0,14 | 0,14 |
| p -Wert | 0,46 | 0,45 | 0,44 | 0,44 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 1). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(0,0,1)(1,0,1)₄-Modells mit $H = 8$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.17.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 2 der Prognosen mit $H = 8$ anhand des Alternativmodells**

| h | $\hat{\gamma}$ | Annahme A | | Annahme B | |
|-----|----------------|--------------------|-----------|--------------------|-----------|
| | | $se(\hat{\gamma})$ | p -Wert | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,067 | 0,16 | 0,67 | 0,13 | 0,62 |
| 3 | -0,003 | 0,17 | 0,99 | 0,14 | 0,98 |
| 4 | 0,046 | 0,16 | 0,78 | 0,14 | 0,74 |
| 5 | 0,087 | 0,18 | 0,63 | 0,15 | 0,57 |
| 6 | 0,093 | 0,13 | 0,47 | 0,11 | 0,40 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2, \dots, 6$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t;h+1;h+2} + \delta + \omega_t$ (Fall 2). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für jedes h sind unter beiden Annahmen die Werte des geschätzten Koeffizienten $\hat{\gamma}$ auf drei, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(0,0,1)(1,0,1)₄-Modells mit $H = 8$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.18.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 2 der Prognosen mit $H = 8$ anhand des Alternativmodells**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | 0,06 | 0,06 | 0,06 | 0,06 |
| $se(\hat{\gamma})$ | 0,13 | 0,13 | 0,13 | 0,13 |
| p -Wert | 0,64 | 0,63 | 0,63 | 0,62 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 2). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(0,0,1)(1,0,1)₄-Modells mit $H = 8$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.19.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 1 der Prognosen mit $H = 8$ anhand des Alternativmodells**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,08 | 0,22 | 0,72 |
| 3 | 0,05 | 0,23 | 0,83 |
| 4 | 0,02 | 0,14 | 0,91 |
| 5 | 0,002 | 0,16 | 0,99 |
| 6 | -0,01 | 0,15 | 0,92 |
| 7 | -0,11 | 0,15 | 0,48 |
| 8 | -0,16 | 0,11 | 0,17 |
| 9 | -0,15 | 0,12 | 0,21 |
| 10 | -0,12 | 0,11 | 0,30 |
| 11 | -0,09 | 0,11 | 0,42 |
| 12 | -0,10 | 0,09 | 0,27 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2, \dots, 12$, anhand der Regressionsgleichung: $r_{t,h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für jedes h sind unter beiden Annahmen die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei bzw. drei Nachkommastellen gerundet. Die Stichprobe umfasst die Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(0,0,1)(1,0,1)₄-Modells mit $H = 12$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.20.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 1 der Prognosen mit $H = 12$ anhand des Alternativmodells**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | -0,08 | -0,08 | -0,08 | -0,08 |
| $se(\hat{\gamma})$ | 0,11 | 0,12 | 0,11 | 0,12 |
| p -Wert | 0,48 | 0,50 | 0,46 | 0,48 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 1). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(0,0,1)(1,0,1)₄-Modells mit $H = 12$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.21.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 2 der Prognosen mit $H = 12$ anhand des Alternativmodells**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,26 | 0,14 | 0,073 |
| 3 | 0,22 | 0,15 | 0,154 |
| 4 | 0,18 | 0,13 | 0,187 |
| 5 | 0,23 | 0,15 | 0,125 |
| 6 | 0,35 | 0,11 | 0,003 |
| 7 | 0,28 | 0,12 | 0,024 |
| 8 | 0,28 | 0,11 | 0,015 |
| 9 | 0,27 | 0,11 | 0,017 |
| 10 | 0,32 | 0,09 | 0,001 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2, \dots, 10$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t;h+1;h+2} + \delta + \omega_t$ (Fall 2). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für jedes h sind unter beiden Annahmen die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei bzw. drei Nachkommastellen gerundet. Die Stichprobe umfasst die Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(0,0,1)(1,0,1)₄-Modells mit $H = 12$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

Tabelle B.22.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 2 der Prognosen mit $H = 12$ anhand des Alternativmodells**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | 0,28 | 0,28 | 0,28 | 0,28 |
| $se(\hat{\gamma})$ | 0,10 | 0,11 | 0,10 | 0,11 |
| p -Wert | 0,01 | 0,01 | 0,01 | 0,01 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 2). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten Prognosen des Gesamtsteueraufkommens der BRD anhand des SARIMA(0,0,1)(1,0,1)₄-Modells mit $H = 12$ vom ersten Quartal 2001 bis zum vierten Quartal 2010 in Milliarden Euro.

C. Ergebnistabellen zur Analyse der AKS-Prognosen

Tabelle C.1.: **Regressionsergebnisse der horizont-spezifischen Verzerrungen der AKS-Prognosen mit $H = 4$**

| h | $\hat{\alpha}_h$ | Annahme A | | Annahme B | |
|-----|------------------|----------------------|-----------|----------------------|-----------|
| | | $se(\hat{\alpha}_h)$ | p -Wert | $se(\hat{\alpha}_h)$ | p -Wert |
| 1 | 0,97 | 1,32 | 0,46 | 3,83 | 0,80 |
| 2 | -4,00 | 7,99 | 0,62 | 7,60 | 0,60 |
| 3 | -11,34 | 12,86 | 0,38 | 11,32 | 0,32 |
| 4 | -16,90 | 17,46 | 0,34 | 15,00 | 0,26 |

Regressionsergebnisse der horizont-spezifischen Verzerrungen beim Test auf Unverzerrtheit der AKS-Prognosen mit $H = 4$, anhand der Regressionsgleichung $e = (i_T \otimes I_H) \alpha^H + v$. Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für jedes h sind unter beiden Annahmen die Werte der geschätzten horizont-spezifischen Verzerrungen $\hat{\alpha}_h$, ihrer Standardfehler $se(\hat{\alpha}_h)$ sowie die entsprechenden p -Werte der einzelnen t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

Tabelle C.2.: Ergebnisse des t -Tests für Variante I mit $H = 3$

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -4,79 | -4,79 |
| $se(\hat{\alpha})$ | 8,72 | 6,99 |
| p -Wert | 0,58 | 0,50 |
| $\hat{\sigma}_\varepsilon^2$ | (-609,30) | – |
| $\hat{\sigma}_u^2$ | 637,83 | 376,70 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, 2, 3$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

Tabelle C.3.: Ergebnisse des F -Tests für Variante II mit $H = 3$

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,26 | 16,58 |
| p -Wert | 0,85 | 0,0000 |
| $\hat{\sigma}_\varepsilon^2$ | (-625,19) | – |
| $\hat{\sigma}_u^2$ | 632,99 | 365,05 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, 2, 3$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

Tabelle C.4.: **Ergebnisse des t -Tests für Variante I mit $H = 2$**

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -1,52 | -1,52 |
| $se(\hat{\alpha})$ | 5,23 | 4,09 |
| p -Wert | 0,77 | 0,71 |
| $\hat{\sigma}_\varepsilon^2$ | (-384,67) | - |
| $\hat{\sigma}_u^2$ | 457,24 | 226,44 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, 2$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über beide Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

Tabelle C.5.: **Ergebnisse des F -Tests für Variante II mit $H = 2$**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,17 | 35,41 |
| p -Wert | 0,84 | 0,0000 |
| $\hat{\sigma}_\varepsilon^2$ | (-390,86) | - |
| $\hat{\sigma}_u^2$ | 457,24 | 222,73 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, 2$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

Tabelle C.6.: **Ergebnisse des t -Tests für Variante I mit $H = 5$ ohne 1990**

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -11,55 | -11,55 |
| $se(\hat{\alpha})$ | 13,04 | 11,56 |
| p -Wert | 0,38 | 0,32 |
| $\hat{\sigma}_\varepsilon^2$ | (-510,45) | – |
| $\hat{\sigma}_u^2$ | 605,62 | 466,40 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, \dots, 5$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1989 sowie 1991 bis 2010.

Tabelle C.7.: **Ergebnisse des F -Tests für Variante II mit $H = 5$ ohne 1990**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,25 | 7,82 |
| p -Wert | 0,94 | 0,0000 |
| $\hat{\sigma}_\varepsilon^2$ | (-585,49) | – |
| $\hat{\sigma}_u^2$ | 604,76 | 445,08 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, \dots, 5$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1989 sowie 1991 bis 2010.

Tabelle C.8.: **Ergebnisse des t -Tests für Variante I mit $H = 4$ ohne 1990**

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -8,43 | -8,43 |
| $se(\hat{\alpha})$ | 10,98 | 9,32 |
| p -Wert | 0,44 | 0,37 |
| $\hat{\sigma}_\varepsilon^2$ | (-577,50) | – |
| $\hat{\sigma}_u^2$ | 624,59 | 432,09 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, \dots, 4$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1989 sowie 1991 bis 2010.

Tabelle C.9.: **Ergebnisse des F -Tests für Variante II mit $H = 4$ ohne 1990**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,26 | 10,44 |
| p -Wert | 0,90 | 0,0000 |
| $\hat{\sigma}_\varepsilon^2$ | (-622,31) | – |
| $\hat{\sigma}_u^2$ | 623,17 | 415,73 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, \dots, 4$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1989 sowie 1991 bis 2010.

Tabelle C.10.: **Ergebnisse des t -Tests für Variante I mit $H = 3$ ohne 1990**

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -5,34 | -5,34 |
| $se(\hat{\alpha})$ | 8,63 | 6,91 |
| p -Wert | 0,54 | 0,44 |
| $\hat{\sigma}_\varepsilon^2$ | (-599,51) | – |
| $\hat{\sigma}_u^2$ | 624,34 | 367,40 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, 2, 3$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1989 sowie 1991 bis 2010.

Tabelle C.11.: **Ergebnisse des F -Tests für Variante II mit $H = 3$ ohne 1990**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,29 | 15,55 |
| p -Wert | 0,83 | 0,0000 |
| $\hat{\sigma}_\varepsilon^2$ | (-616,90) | – |
| $\hat{\sigma}_u^2$ | 619,92 | 355,54 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, 2, 3$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1989 sowie 1991 bis 2010.

Tabelle C.12.: **Ergebnisse des t -Tests für Variante I mit $H = 2$ ohne 1990**

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -2,04 | -2,04 |
| $se(\hat{\alpha})$ | 5,13 | 4,01 |
| p -Wert | 0,69 | 0,61 |
| $\hat{\sigma}_\varepsilon^2$ | (-370,92) | – |
| $\hat{\sigma}_u^2$ | 440,46 | 217,92 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, 2$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1989 sowie 1991 bis 2010.

Tabelle C.13.: **Ergebnisse des F -Tests für Variante II mit $H = 2$ ohne 1990**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,24 | 34,24 |
| p -Wert | 0,79 | 0,0000 |
| $\hat{\sigma}_\varepsilon^2$ | (-377,62) | – |
| $\hat{\sigma}_u^2$ | 440,46 | 213,89 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, 2$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1989 sowie 1991 bis 2010.

Tabelle C.14.: **Ergebnisse des t -Tests für Variante I mit $H = 5$ ohne die Jahre 1991 bis 1995**

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -14,28 | -14,28 |
| $se(\hat{\alpha})$ | 10,28 | 9,88 |
| p -Wert | 0,17 | 0,15 |
| $\hat{\sigma}_\varepsilon^2$ | (-113,78) | – |
| $\hat{\sigma}_u^2$ | 372,02 | 340,99 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, \dots, 5$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1976 bis 1990 sowie 1996 bis 2010.

Tabelle C.15.: **Ergebnisse des F -Tests für Variante II mit $H = 5$ ohne die Jahre 1991 bis 1995**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,72 | 17,72 |
| p -Wert | 0,61 | 0,0000 |
| $\hat{\sigma}_\varepsilon^2$ | (-224,89) | – |
| $\hat{\sigma}_u^2$ | 359,93 | 298,60 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, \dots, 5$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1976 bis 1990 sowie 1996 bis 2010.

Tabelle C.16.: **Ergebnisse des t -Tests für Variante I mit $H = 4$ ohne die Jahre 1991 bis 1995**

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -9,79 | -9,79 |
| $se(\hat{\alpha})$ | 9,24 | 8,23 |
| p -Wert | 0,29 | 0,24 |
| $\hat{\sigma}_\varepsilon^2$ | (-302,47) | – |
| $\hat{\sigma}_u^2$ | 438,04 | 337,21 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, \dots, 4$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1976 bis 1990 sowie 1996 bis 2010.

Tabelle C.17.: **Ergebnisse des F -Tests für Variante II mit $H = 4$ ohne die Jahre 1991 bis 1995**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,59 | 22,05 |
| p -Wert | 0,67 | 0,0000 |
| $\hat{\sigma}_\varepsilon^2$ | (-361,91) | – |
| $\hat{\sigma}_u^2$ | 428,44 | 307,81 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, \dots, 4$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1976 bis 1990 sowie 1996 bis 2010.

Tabelle C.18.: **Ergebnisse des t -Tests für Variante I mit $H = 3$ ohne die Jahre 1991 bis 1995**

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -5,54 | -5,54 |
| $se(\hat{\alpha})$ | 7,75 | 6,34 |
| p -Wert | 0,48 | 0,38 |
| $\hat{\sigma}_\varepsilon^2$ | (-447,10) | – |
| $\hat{\sigma}_u^2$ | 501,10 | 309,49 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, 2, 3$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1976 bis 1990 sowie 1996 bis 2010.

Tabelle C.19.: **Ergebnisse des F -Tests für Variante II mit $H = 3$ ohne die Jahre 1991 bis 1995**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,48 | 31,92 |
| p -Wert | 0,70 | 0,0000 |
| $\hat{\sigma}_\varepsilon^2$ | (-467,58) | – |
| $\hat{\sigma}_u^2$ | 491,85 | 291,45 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, 2, 3$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1976 bis 1990 sowie 1996 bis 2010.

Tabelle C.20.: **Ergebnisse des t -Tests für Variante I mit $H = 2$ ohne die Jahre 1991 bis 1995**

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -1,45 | -1,45 |
| $se(\hat{\alpha})$ | 4,87 | 3,82 |
| p -Wert | 0,77 | 0,71 |
| $\hat{\sigma}_\varepsilon^2$ | (-331,31) | – |
| $\hat{\sigma}_u^2$ | 395,82 | 192,09 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, 2$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1976 bis 1990 sowie 1996 bis 2010.

Tabelle C.21.: **Ergebnisse des F -Tests für Variante II mit $H = 2$ ohne die Jahre 1991 bis 1995**

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,23 | 58,97 |
| p -Wert | 0,80 | 0,0000 |
| $\hat{\sigma}_\varepsilon^2$ | (-339,55) | – |
| $\hat{\sigma}_u^2$ | 395,82 | 192,09 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, 2$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1976 bis 1990 sowie 1996 bis 2010.

Tabelle C.22.: Ergebnisse des t -Tests für Variante I mit $H = 5$ ohne 1991

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -12,43 | -12,43 |
| $se(\hat{\alpha})$ | 12,55 | 11,17 |
| p -Wert | 0,32 | 0,27 |
| $\hat{\sigma}_\varepsilon^2$ | (-459,28) | – |
| $\hat{\sigma}_u^2$ | 561,16 | 435,90 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, \dots, 5$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1990 sowie 1992 bis 2010.

Tabelle C.23.: Ergebnisse des F -Tests für Variante II mit $H = 5$ ohne 1991

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,30 | 10,76 |
| p -Wert | 0,91 | 0,0000 |
| $\hat{\sigma}_\varepsilon^2$ | (-551,10) | – |
| $\hat{\sigma}_u^2$ | 562,43 | 412,13 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, \dots, 5$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1990 sowie 1992 bis 2010.

Tabelle C.24.: Ergebnisse des t -Tests für Variante I mit $H = 4$ ohne 1991

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -9,17 | -9,17 |
| $se(\hat{\alpha})$ | 10,50 | 8,96 |
| p -Wert | 0,38 | 0,31 |
| $\hat{\sigma}_\varepsilon^2$ | (-509,47) | – |
| $\hat{\sigma}_u^2$ | 569,94 | 400,12 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, \dots, 4$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1990 sowie 1992 bis 2010.

Tabelle C.25.: Ergebnisse des F -Tests für Variante II mit $H = 4$ ohne 1991

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,33 | 14,60 |
| p -Wert | 0,86 | 0,0000 |
| $\hat{\sigma}_\varepsilon^2$ | (-565,80) | – |
| $\hat{\sigma}_u^2$ | 569,78 | 381,18 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, \dots, 4$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1990 sowie 1992 bis 2010.

Tabelle C.26.: Ergebnisse des t -Tests für Variante I mit $H = 3$ ohne 1991

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -5,85 | -5,85 |
| $se(\hat{\alpha})$ | 8,23 | 6,64 |
| p -Wert | 0,48 | 0,38 |
| $\hat{\sigma}_\varepsilon^2$ | (-529,88) | – |
| $\hat{\sigma}_u^2$ | 566,49 | 339,39 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, 2, 3$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1990 sowie 1992 bis 2010.

Tabelle C.27.: Ergebnisse des F -Tests für Variante II mit $H = 3$ ohne 1991

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,38 | 21,86 |
| p -Wert | 0,76 | 0,0000 |
| $\hat{\sigma}_\varepsilon^2$ | (-553,63) | – |
| $\hat{\sigma}_u^2$ | 562,48 | 325,21 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, 2, 3$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1990 sowie 1992 bis 2010.

Tabelle C.28.: Ergebnisse des t -Tests für Variante I mit $H = 2$ ohne 1991

| t -Test mit $H_0 : \alpha = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|-------------------------------------|--|---|
| $\hat{\alpha}$ | -2,27 | -2,27 |
| $se(\hat{\alpha})$ | 4,86 | 3,86 |
| p -Wert | 0,64 | 0,56 |
| $\hat{\sigma}_\varepsilon^2$ | (-316,71) | – |
| $\hat{\sigma}_u^2$ | 391,56 | 201,54 |

Ergebnisse des t -Tests der Nullhypothese der Unverzerrtheit der Prognosen für $h = 1, 2$ Prognosehorizonte gegen die Alternativhypothese einer gleich großen Verzerrung über alle Prognosehorizonte (Variante I). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind die Werte der geschätzten Verzerrung $\hat{\alpha}$, ihres Standardfehlers $se(\hat{\alpha})$, der entsprechende p -Wert des t -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1990 sowie 1992 bis 2010.

Tabelle C.29.: Ergebnisse des F -Tests für Variante II mit $H = 2$ ohne 1991

| F -Test mit $H_0 : \alpha^H = 0$ | Annahme A (Modell mit ε_{th}) | Annahme B (Modell ohne ε_{th}) |
|---------------------------------------|--|---|
| F -Statistik | 0,34 | 51,39 |
| p -Wert | 0,71 | 0,0000 |
| $\hat{\sigma}_\varepsilon^2$ | (-325,75) | – |
| $\hat{\sigma}_u^2$ | 391,56 | 196,12 |

Ergebnisse des F -Tests der Nullhypothese der Unverzerrtheit der Prognosen für alle $h = 1, 2$ Prognosehorizonte gegen die Alternativhypothese horizont-spezifischer Verzerrungen (Variante II). Annahme A bezeichnet die Modellvariante mit idiosynkratischem Störterm ε_{th} , Annahme B diejenige ohne idiosynkratischen Störterm. Für beide Annahmen sind der Wert der F -Statistik, der entsprechende p -Wert des F -Tests sowie die beiden Varianzschätzer $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_u^2$ auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1990 sowie 1992 bis 2010.

Tabelle C.30.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 4$**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,29 | 0,21 | 0,19 |
| 3 | 0,32 | 0,19 | 0,10 |
| 4 | 0,32 | 0,17 | 0,07 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2, 3, 4$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für jedes h sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 4$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

Tabelle C.31.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 4$**

| | Variante I | | Variante II | |
|--------------------|-------------------|------------------|--------------------|------------------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | 0,31 | 0,31 | 0,31 | 0,31 |
| $se(\hat{\gamma})$ | 0,15 | 0,18 | 0,14 | 0,18 |
| p -Wert | 0,04 | 0,09 | 0,03 | 0,08 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 1). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 4$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

Tabelle C.32.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 3$**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,29 | 0,20 | 0,15 |
| 3 | 0,32 | 0,17 | 0,07 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2$ und $h = 3$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für jedes h sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 3$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

Tabelle C.33.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 3$**

| | Variante I | | Variante II | |
|--------------------|-------------------|------------------|--------------------|------------------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | 0,31 | 0,31 | 0,31 | 0,31 |
| $se(\hat{\gamma})$ | 0,09 | 0,18 | 0,08 | 0,18 |
| p -Wert | 0,0013 | 0,09 | 0,0004 | 0,09 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 1). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 3$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

Tabelle C.34.: **Ergebnis des separaten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 2$**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,29 | 0,15 | 0,07 |

Ergebnis des t -Tests der Nullhypothese der Informationseffizienz für $h = 2$ anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für $h = 2$ sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 2$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

Tabelle C.35.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 5$ ohne 1990**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,27 | 0,22 | 0,22 |
| 3 | 0,32 | 0,19 | 0,11 |
| 4 | 0,32 | 0,17 | 0,08 |
| 5 | 0,37 | 0,18 | 0,04 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2, \dots, 5$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für jedes h sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 5$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1989 sowie 1991 bis 2010.

Tabelle C.36.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 5$ ohne 1990**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | 0,32 | 0,32 | 0,32 | 0,32 |
| $se(\hat{\gamma})$ | 0,16 | 0,18 | 0,16 | 0,17 |
| p -Wert | 0,05 | 0,07 | 0,04 | 0,06 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 1). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 5$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1989 sowie 1991 bis 2010.

Tabelle C.37.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 4$ ohne 1990**

| h | Annahme B | | |
|-----|----------------|--------------------|-----------|
| | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,27 | 0,21 | 0,21 |
| 3 | 0,32 | 0,19 | 0,10 |
| 4 | 0,32 | 0,17 | 0,07 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2, 3, 4$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für jedes h sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 4$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1989 sowie 1991 bis 2010.

Tabelle C.38.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 4$ ohne 1990**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | 0,30 | 0,30 | 0,31 | 0,31 |
| $se(\hat{\gamma})$ | 0,15 | 0,18 | 0,14 | 0,18 |
| p -Wert | 0,04 | 0,09 | 0,04 | 0,09 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 1). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 4$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1989 sowie 1991 bis 2010.

Tabelle C.39.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 3$ ohne 1990**

| h | Annahme B | | |
|-----|----------------|--------------------|-----------|
| | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,27 | 0,20 | 0,17 |
| 3 | 0,32 | 0,17 | 0,08 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2$ und $h = 3$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für jedes h sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 3$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1989 sowie 1991 bis 2010.

Tabelle C.40.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 3$ ohne 1990**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | 0,30 | 0,30 | 0,30 | 0,30 |
| $se(\hat{\gamma})$ | 0,09 | 0,18 | 0,08 | 0,17 |
| p -Wert | 0,0016 | 0,10 | 0,0004 | 0,09 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 1). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 3$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1989 sowie 1991 bis 2010.

Tabelle C.41.: **Ergebnis des separaten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 2$ ohne 1990**

| h | Annahme B | | |
|-----|----------------|--------------------|-----------|
| | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,27 | 0,15 | 0,08 |

Ergebnis des t -Tests der Nullhypothese der Informationseffizienz für $h = 2$ anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für $h = 2$ sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 2$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1989 sowie 1991 bis 2010.

Tabelle C.42.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 5$ ohne 1991**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,22 | 0,23 | 0,35 |
| 3 | 0,27 | 0,19 | 0,17 |
| 4 | 0,31 | 0,17 | 0,08 |
| 5 | 0,37 | 0,17 | 0,04 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2, \dots, 5$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für jedes h sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 5$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1990 sowie 1992 bis 2010.

Tabelle C.43.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 5$ ohne 1991**

| | Variante I | | Variante II | |
|--------------------|-------------------|------------------|--------------------|------------------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | 0,30 | 0,30 | 0,30 | 0,30 |
| $se(\hat{\gamma})$ | 0,16 | 0,17 | 0,15 | 0,17 |
| p -Wert | 0,06 | 0,08 | 0,04 | 0,07 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 1). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 5$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1990 sowie 1992 bis 2010.

Tabelle C.44.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 4$ ohne 1991**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,22 | 0,22 | 0,33 |
| 3 | 0,27 | 0,18 | 0,16 |
| 4 | 0,31 | 0,16 | 0,07 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2, 3, 4$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für jedes h sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 4$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1990 sowie 1992 bis 2010.

Tabelle C.45.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 4$ ohne 1991**

| | Variante I | | Variante II | |
|--------------------|-------------------|------------------|--------------------|------------------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | 0,27 | 0,27 | 0,27 | 0,27 |
| $se(\hat{\gamma})$ | 0,14 | 0,17 | 0,13 | 0,17 |
| p -Wert | 0,06 | 0,12 | 0,05 | 0,11 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 1). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 4$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1990 sowie 1992 bis 2010.

Tabelle C.46.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 3$ ohne 1991**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,22 | 0,20 | 0,29 |
| 3 | 0,27 | 0,17 | 0,13 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2$ und $h = 3$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für jedes h sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 3$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1990 sowie 1992 bis 2010.

Tabelle C.47.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 3$ ohne 1991**

| | Variante I | | Variante II | |
|--------------------|-------------------|------------------|--------------------|------------------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | 0,24 | 0,24 | 0,25 | 0,25 |
| $se(\hat{\gamma})$ | 0,08 | 0,17 | 0,06 | 0,17 |
| p -Wert | 0,0024 | 0,17 | 0,0001 | 0,15 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 1). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei beziehungsweise vier Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 3$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1990 sowie 1992 bis 2010.

Tabelle C.48.: **Ergebnis des separaten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 2$ ohne 1991**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,22 | 0,16 | 0,18 |

Ergebnis des t -Tests der Nullhypothese der Informationseffizienz für $h = 2$ anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für $h = 2$ sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 2$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1990 sowie 1992 bis 2010.

Tabelle C.49.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 4$ ohne die Jahre 1991 bis 1995**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,17 | 0,20 | 0,40 |
| 3 | 0,21 | 0,19 | 0,28 |
| 4 | 0,17 | 0,17 | 0,33 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2, 3, 4$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme B bezeichnet die Modellvariante ohne idiosynkratischem Störterm ε_{th} . Für jedes h sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 4$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1976 bis 1990 sowie 1996 bis 2010.

Tabelle C.50.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 4$ ohne die Jahre 1991 bis 1995**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | 0,18 | 0,18 | 0,18 | 0,18 |
| $se(\hat{\gamma})$ | 0,15 | 0,18 | 0,14 | 0,17 |
| p -Wert | 0,22 | 0,30 | 0,19 | 0,29 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 1). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 4$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1976 bis 1990 sowie 1996 bis 2010.

Tabelle C.51.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 3$ ohne die Jahre 1991 bis 1995**

| h | Annahme B | | |
|-----|----------------|--------------------|-----------|
| | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,17 | 0,19 | 0,38 |
| 3 | 0,21 | 0,18 | 0,26 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2$ und $h = 3$, anhand der Regressionsgleichung: $r_{t,h-1,h} = \gamma r_{t-1,h-1,h} + \delta + \omega_t$ (Fall 1). Annahme A bezeichnet die Modellvariante ohne idiosynkratischem Störterm ε_{th} . Für jedes h sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 3$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1976 bis 1990 sowie 1996 bis 2010.

Tabelle C.52.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 3$ ohne die Jahre 1991 bis 1995**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | 0,19 | 0,19 | 0,19 | 0,19 |
| $se(\hat{\gamma})$ | 0,09 | 0,18 | 0,07 | 0,18 |
| p -Wert | 0,04 | 0,29 | 0,01 | 0,28 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 1). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 3$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1976 bis 1990 sowie 1996 bis 2010.

Tabelle C.53.: **Ergebnis des separaten Tests auf Informationseffizienz gemäß Fall 1 der AKS-Prognosen mit $H = 2$ ohne die Jahre 1991 bis 1995**

| h | Annahme B | | |
|-----|----------------|--------------------|-----------|
| | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | 0,17 | 0,16 | 0,28 |

Ergebnis des t -Tests der Nullhypothese der Informationseffizienz für $h = 2$ anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t-1;h-1;h} + \delta + \omega_t$ (Fall 1). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für $h = 2$ sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 2$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1976 bis 1990 sowie 1996 bis 2010.

Tabelle C.54.: **Ergebnis des separaten Tests auf Informationseffizienz gemäß Fall 2 der AKS-Prognosen mit $H = 4$**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | -0,12 | 0,17 | 0,47 |

Ergebnis des t -Tests der Nullhypothese der Informationseffizienz für $h = 2$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t;h+1;h+2} + \delta + \omega_t$ (Fall 2). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für $h = 2$ sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 4$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1981 bis 2010.

Tabelle C.55.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 2 der AKS-Prognosen mit $H = 5$ ohne 1990**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | -0,13 | 0,17 | 0,47 |
| 3 | -0,26 | 0,16 | 0,11 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2$ und $h = 3$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t;h+1;h+2} + \delta + \omega_t$ (Fall 2). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für beide Prognosehorizonte h sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 5$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1989 sowie 1991 bis 2010.

Tabelle C.56.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 2 der AKS-Prognosen mit $H = 5$ ohne 1990**

| | Variante I | | Variante II | |
|--------------------|------------|-----------|-------------|-----------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | -0,20 | -0,20 | -0,20 | -0,20 |
| $se(\hat{\gamma})$ | 0,10 | 0,16 | 0,08 | 0,16 |
| p -Wert | 0,05 | 0,22 | 0,01 | 0,21 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 2). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 5$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1989 sowie 1991 bis 2010.

Tabelle C.57.: **Ergebnis des separaten Tests auf Informationseffizienz gemäß Fall 2 der AKS-Prognosen mit $H = 4$ ohne 1990**

| h | Annahme B | | |
|-----|----------------|--------------------|-----------|
| | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | -0,13 | 0,17 | 0,45 |

Ergebnis des t -Tests der Nullhypothese der Informationseffizienz für $h = 2$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t;h+1;h+2} + \delta + \omega_t$ (Fall 2). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für $h = 2$ sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 4$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1989 sowie 1991 bis 2010.

Tabelle C.58.: **Ergebnisse der separaten Tests auf Informationseffizienz gemäß Fall 2 der AKS-Prognosen mit $H = 5$ ohne 1991**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | -0,15 | 0,17 | 0,38 |
| 3 | -0,26 | 0,15 | 0,09 |

Ergebnisse des t -Tests der Nullhypothese der Informationseffizienz, separat für $h = 2$ und $h = 3$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t;h+1;h+2} + \delta + \omega_t$ (Fall 2). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für beide Prognosehorizonte h sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 5$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1990 sowie 1992 bis 2010.

Tabelle C.59.: **Ergebnisse des gepoolten Tests auf Informationseffizienz gemäß Fall 2 der AKS-Prognosen mit $H = 5$ ohne 1991**

| | Variante I | | Variante II | |
|--------------------|-------------------|------------------|--------------------|------------------|
| | Annahme A | Annahme B | Annahme A | Annahme B |
| $\hat{\gamma}$ | -0,21 | -0,21 | -0,21 | -0,21 |
| $se(\hat{\gamma})$ | 0,10 | 0,16 | 0,07 | 0,15 |
| p -Wert | 0,04 | 0,19 | 0,01 | 0,17 |

Ergebnisse des Tests der Nullhypothese der Informationseffizienz, anhand der Regressionsgleichung: $r = \gamma r_{-1} + \delta_d + \omega$ (Fall 2). Für Variante I (gleich große Verzerrung über alle Prognosehorizonte) und Variante II (horizont-spezifische Verzerrungen) sind jeweils unter Annahme A (Modell mit idiosynkratischem Störterm ε_{th}) sowie unter Annahme B (Modell ohne ε_{th}) die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie die entsprechenden p -Werte des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 5$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1990 sowie 1992 bis 2010.

Tabelle C.60.: **Ergebnis des separaten Tests auf Informationseffizienz gemäß Fall 2 der AKS-Prognosen mit $H = 4$ ohne 1991**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | -0,15 | 0,16 | 0,36 |

Ergebnis des t -Tests der Nullhypothese der Informationseffizienz für $h = 2$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t;h+1;h+2} + \delta + \omega_t$ (Fall 2). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für $h = 2$ sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 4$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1980 bis 1990 sowie 1992 bis 2010.

Tabelle C.61.: **Ergebnis des separaten Tests auf Informationseffizienz gemäß Fall 2 der AKS-Prognosen mit $H = 4$ ohne die Jahre 1991 bis 1995**

| Annahme B | | | |
|------------------|----------------|--------------------|-----------|
| h | $\hat{\gamma}$ | $se(\hat{\gamma})$ | p -Wert |
| 2 | -0,15 | 0,16 | 0,36 |

Ergebnis des t -Tests der Nullhypothese der Informationseffizienz für $h = 2$, anhand der Regressionsgleichung: $r_{t;h-1;h} = \gamma r_{t;h+1;h+2} + \delta + \omega_t$ (Fall 2). Annahme B bezeichnet die Modellvariante ohne idiosynkratischen Störterm ε_{th} . Für $h = 2$ sind die Werte des geschätzten Koeffizienten $\hat{\gamma}$, seines Standardfehlers $se(\hat{\gamma})$ sowie der entsprechende p -Wert des t -Tests auf zwei Nachkommastellen gerundet. Die Stichprobe umfasst die gepoolten AKS-Prognosen mit $H = 4$ des Gesamtsteueraufkommens der BRD in Milliarden Euro für die Jahre 1976 bis 1990 sowie 1996 bis 2010.

D. Ergebnistabellen zur deskriptiven Analyse der Prognosegüte

Tabelle D.1.: **Prozentfehler PE_{th} der Prognosen anhand des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells gemäß Methode 1**

| Prognoseziel t | Prognosehorizont h: | | |
|----------------|---------------------|--------|--------|
| | 2 | 3 | 4 |
| 2001 | -9,37 | -13,47 | 4,48 |
| 2002 | -6,56 | -15,40 | -21,89 |
| 2003 | 8,01 | -7,90 | -20,67 |
| 2004 | 4,28 | 11,58 | -9,38 |
| 2005 | 1,01 | 7,70 | 16,60 |
| 2006 | 4,13 | 7,16 | 15,55 |
| 2007 | 2,96 | 3,56 | 8,38 |
| 2008 | 4,31 | 5,30 | 4,80 |
| 2009 | -17,87 | -18,12 | -20,23 |
| 2010 | -2,42 | -24,86 | -23,36 |

Werte des auf zwei Nachkommastellen gerundeten Prozentfehlers PE_{th} der Prognosen des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells des Gesamtsteueraufkommens der BRD für die Prognoseziele $t = 2001$ bis 2010 und Prognosehorizonte $h = 2, 3, 4$. Die Quartalsprognosen wurden gemäß Methode 1 kumuliert (vgl. Kapitel 4.4.1).

Tabelle D.2.: **Prozentfehler PE_{th} der Prognosen anhand des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells gemäß Methode 1**

| Prognoseziel t | Prognosehorizont h: | | |
|----------------|---------------------|--------|--------|
| | 2 | 3 | 4 |
| 2001 | -9,39 | -13,46 | 4,35 |
| 2002 | -6,63 | -15,44 | -21,88 |
| 2003 | 7,95 | -8,00 | -20,73 |
| 2004 | 4,21 | 11,48 | -9,52 |
| 2005 | 0,97 | 7,60 | 16,48 |
| 2006 | 4,14 | 7,12 | 15,43 |
| 2007 | 2,90 | 3,56 | 8,33 |
| 2008 | 4,30 | 5,20 | 4,82 |
| 2009 | -17,83 | -18,10 | -20,07 |
| 2010 | -2,53 | -24,79 | -23,32 |

Werte des auf zwei Nachkommastellen gerundeten Prozentfehlers PE_{th} der Prognosen des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells des Gesamtsteueraufkommens der BRD für die Prognoseziele $t = 2001$ bis 2010 und Prognosehorizonte $h = 2, 3, 4$. Die Quartalsprognosen wurden gemäß Methode 1 kumuliert (vgl. Kapitel 4.4.1).

Tabelle D.3.: **Prozentfehler PE_{th} der Prognosen anhand des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells gemäß Methode 2**

| Prognoseziel t | Prognosehorizont h: | | | |
|----------------|---------------------|--------|--------|--------|
| | 2 | 3 | 4 | 5 |
| 2001 | -7,11 | -13,59 | -10,65 | 2,62 |
| 2002 | 1,89 | -11,25 | -21,84 | -17,82 |
| 2003 | 1,22 | 4,77 | -14,65 | -29,50 |
| 2004 | -1,66 | 1,26 | 7,37 | -18,38 |
| 2005 | 1,48 | -0,97 | 3,05 | 11,56 |
| 2006 | 4,37 | 7,88 | 4,83 | 9,52 |
| 2007 | -4,29 | 3,47 | 9,34 | 5,56 |
| 2008 | 0,56 | -6,90 | 4,35 | 12,40 |
| 2009 | -9,78 | -10,14 | -22,38 | -5,92 |
| 2010 | 5,50 | -11,50 | -12,60 | -29,21 |

Werte des auf zwei Nachkommastellen gerundeten Prozentfehlers PE_{th} der Prognosen des bevorzugten SARIMA(1,0,0)(1,0,1)₄-Modells des Gesamtsteueraufkommens der BRD für die Prognoseziele $t = 2001$ bis 2010 und Prognosehorizonte $h = 2, 3, 4, 5$. Die Quartalsprognosen wurden gemäß Methode 2 kumuliert (vgl. Kapitel 4.4.2).

Tabelle D.4.: **Prozentfehler PE_{th} der Prognosen anhand des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells gemäß Methode 2**

| Prognoseziel t | Prognosehorizont h: | | | |
|----------------|---------------------|--------|--------|--------|
| | 2 | 3 | 4 | 5 |
| 2001 | -7,15 | -13,59 | -10,67 | 2,49 |
| 2002 | 1,81 | -11,32 | -21,84 | -17,85 |
| 2003 | 1,18 | 4,64 | -14,75 | -29,50 |
| 2004 | -1,70 | 1,19 | 7,20 | -18,52 |
| 2005 | 1,44 | -1,04 | 2,96 | 11,35 |
| 2006 | 4,35 | 7,82 | 4,74 | 9,42 |
| 2007 | -4,27 | 3,45 | 9,25 | 5,45 |
| 2008 | 0,54 | -6,88 | 4,32 | 12,29 |
| 2009 | -9,82 | -10,18 | -22,34 | -5,95 |
| 2010 | 5,41 | -11,56 | -12,64 | -29,15 |

Werte des auf zwei Nachkommastellen gerundeten Prozentfehlers PE_{th} der Prognosen des alternativen SARIMA(0,0,1)(1,0,1)₄-Modells des Gesamtsteueraufkommens der BRD für die Prognoseziele $t = 2001$ bis 2010 und Prognosehorizonte $h = 2, 3, 4, 5$. Die Quartalsprognosen wurden gemäß Methode 2 kumuliert (vgl. Kapitel 4.4.2).

Literaturverzeichnis

- Akaike, H. (1974): A New Look at Statistical Model Identification. *IEEE Transactions on Automobile Control*, 19(6):716–723.
- Anderson, R. L. (1942): Distribution of the Serial Correlation Coefficient. *Annals of Mathematical Statistics*, 13:1–13.
- Auerbach, A. J. (1996): Dynamic Revenue Estimation. *Journal of Economic Perspectives*, 10(1):141–157.
- Bartlett, M. S. (1946): On the Theoretical Specification of Sampling Properties of Autocorrelated Time Series. *Journal of the Royal Statistical Society, Series B*, 8:27.
- Becker, I. und T. Buettner (2007): *Are German Tax-Revenue Forecasts Flawed?* Paper präsentiert auf der Jahrestagung des Vereins für Socialpolitik 2007.
- Bowen, E. K. und M. K. Starr (1982): *Basic Statistics for Business and Economics*. McGraw-Hill Book Company.
- Box, G. E. P. und D. R. Cox (1982): An Analysis of Transformations Revisted, Rebutted. *Journal of the American Statistical Association*, 77(377):209–210.
- Box, G. E. P. und G. M. Jenkins (1970): *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.
- Box, G. E. P. und D. A. Pierce (1970): Distribution of Residual Autocorrelations in Autoregressive Moving Average Time Series Models. *Journal of the American Statistical Association*, 65:1509–1526.

- Bretschneider, St. und W. Gorr (1987): State and Local Government Revenue Forecasting. In: *The Handbook of Forecasting: A Manager's Guide*, Hg. S. Makridakis und C. Wheelwright St.
- Bretschneider, St.; W. Gorr; G. Grizzle und E. Klay (1989): Political and Organizational Influences on the Accuracy of Forecasting State Government Revenues. *International Journal of Forecasting*, 5:307–319.
- Büttner, T. und B. Kauder (2008a): Steuerschätzung im internationalen Vergleich. In: *Monatsbericht des BMF Juni 2008*. S. 55–65.
- Büttner, T. und B. Kauder (2008b): Qualität der Steuerschätzung im internationalen Vergleich. *ifo Schnelldienst*, 14:29–35.
- Bundesministerium der Finanzen, BMF (2005): *50 Jahre Arbeitskreis „Steuerschätzungen“*.
- Bundesministerium der Finanzen, BMF (2011a): Ergebnisse des Arbeitskreises „Steuerschätzungen“ (Pressemitteilungen) seit 1971. Homepage des Bundesministeriums der Finanzen: http://www.bundesfinanzministerium.de/nn_53848/DE/Wirtschaft__und__Verwaltung/Steuern/Steuerschaetzung__einnahmen/Ergebnis__der__Steuerschaetzung/19761.html?__nnn=true. Zugegriffen am 15.01.2012.
- Bundesministerium der Finanzen, BMF (2011b): Haushaltsrechnung des Bundes für das Haushaltsjahr 2010. Homepage des Bundesministeriums der Finanzen: http://www.bundesfinanzministerium.de/nn_4314/DE/Wirtschaft__und__Verwaltung/Finanz__und__Wirtschaftspolitik/Bundeshaushalt/Haushalts-und-Vermögensrechnung/Haushaltsrechnung_202010,templateId=raw,property=publicationFile.pdf. Zugegriffen am 08.02.2012.
- Bundesregierung (1996): Methoden und Ergebnisse der Steuerschätzung - Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Dr. Barbara Höll und der Gruppe der PDS. Deutscher Bundestag: <http://dip21.bundestag.de/dip21/btd/13/049/1304958.pdf>. Zugegriffen am: 19.03.2007.

- Clemente, J.; A. Montanés und M. Reyes (1998): Testing for a Unit Root in Variables with a Double Change in the Mean. *Economics Letters*, 59:175–182.
- Clements, M.; F. Joutz und H. Stekler (2007): An Evaluation of the Forecasts of the Federal Reserve: A Pooled Approach. *Journal of Applied Economics*, 22:121–136.
- Clements, M. P. (1995): Rationality and the Role of Judgement in Macroeconomic Forecasting. *Economic Journal*, 105:410–420.
- Clements, M. P. (1997): Evaluating the Rationality of Fixed-Event Forecasts. *Journal of Forecasting*, 16:225–239.
- Conover, W. J. (1999): *Practical Nonparametric Statistics*. Wiley, 3. Aufl.
- Davies, A. und K. Lahiri (1995): A New Framework for Analyzing Survey Forecasts Using Three-dimensional Panel Data. *Journal of Econometrics*, 68:205–227.
- Davies, A. und K. Lahiri (1999): Re-examining the Rational Expectations Hypothesis Using Panel Data on Multiperiod Forecasts. In: *Analysis of Panels and Limited Dependent Variable Models*, Hg. K.; Lee L. F. Hsiao, C.; Lahiri und M. H. Pesaran. Cambridge University Press, Cambridge, UK, S. 226–254.
- Davis, E. P. und G. Fagan (1997): Are Financial Spreads Useful Indicators of Future Inflation and Output Growth in EU Countries? *Journal of Applied Econometrics*, 12:701–714.
- Depalo, D. (2009): A Seasonal Unit-Root Test with Stata. *Stata Journal*, 9(3):422–438.
- Dickey, D. A. und W. A. Fuller (1979): Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, 74:427–431.
- Diebold, F. X. (1998): *Elements of Forecasting*. South-Western College Publishing.
- Diebold, F. X. und R. S. Mariano (1995): Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, 13(3):253–263.

- Diller, D. (1987): Steuerschätzung - ein Prognoseinstrument der staatlichen Finanzplanung. *Das Wirtschaftsstudium: wisu*, Heft 6:293–295.
- Doornik, J. A. und H. Hansen (2008): An Omnibus Test for Univariate and Multivariate Normality. *Oxford Bulletin of Economics and Statistics*, 70:927–939.
- Elliott, G.; T. J. Rothenberg und J. H. Stock (July 1996): Efficient Tests for an Autoregressive Unit Root. *Econometrica*, 64(4):813–836.
- Enders, W. (1995): *Applied Econometric Time Series*. Wiley Series in Probability and Mathematical Statistics.
- Fildes, R. (1995): *World Index of Economic Forecasts*. Gower Publishing Limited, Hampshire, 4. Aufl.
- Flascha, K. (1985): *Probleme und Methoden der Steuerschätzung*. Inaugural-Dissertation zur Erlangung der wirtschaftswissenschaftlichen Doktorwürde des Fachbereichs Wirtschaftswissenschaften der Philipps-Universität zu Marburg.
- Fox, K.-P. (2005): 50 Jahre Steuerschätzung: Die Notwendigkeit einer undankbaren Aufgabe. *Wirtschaftsdienst: Zeitschrift für Wirtschaftspolitik*, 85, Nr. 4:244–248.
- Gebhardt, H. (2001): Methoden, Probleme und Ergebnisse der Steuerschätzung. *RWI-Mitteilungen*, 52(2):127–147.
- Gourieroux, C. und A. Monfort (1997): *Time Series and Dynamic Models*. Cambridge University Press.
- Gravelle, J. G. (1995): Behavioral Feedback and the Revenue-Estimating Process. *National Tax Journal*, 48(3):463–477.
- Greene, W. H. (2003): *Econometric Analysis*. Pearson Education Inc., New Jersey, 5. Aufl.
- Hamilton, J. D. (1994): *Time Series Analysis*. Princeton University Press, Princeton, New Jersey.
- Hartung, J. (2009): *Statistik*. Oldenbourg Verlag, München, 15. Aufl.

- Harvey, A. C. (1989): *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.
- Harvey, D.; S. Leybourne und P. Newbold (1997): Testing the Equality of Prediction Mean Squared Errors. *International Journal of Forecasting*, 13:281–291.
- Hayashi, F. (2000): *Econometrics*. Princeton University Press.
- Holden, K. und D. A. Peel (1990): On Testing for Unbiasedness and Efficiency of Forecasts. *Manchester School*, 58:120–127.
- Hyllenberg, S.; R. F. Engle; C. W. J. Granger und B. S. Yoo (1990): Seasonal Integration and Cointegration. *Journal of Econometrics*, 44:215–238.
- Institut für Finanzen und Steuern (1988): *Zur Problematik von Steuerschätzungen bei gesetzlichen Steueränderungen*. Grüne Briefe 281, Institut für Finanzen und Steuern, Bonn.
- Jarque, C. M. und A. K. Bera (1980): Efficient Tests for Normality, Heteroskedasticity, and Serial Independence of Regression Residuals. *Economics Letters*, 6:255–259.
- Jarque, C. M. und A. K. Bera (1987): A Test for Normality of Observations and Regression Residuals. *International Statistical Review*, 55(2):163–172.
- Karrenberg, H. (Juli 2001): Neue Steuerschätzung für 2001 bis 2005. *Kommunale Steuer-Zeitschrift*, 50(7):121–126.
- Keane, M. P. und D. E. Runkle (1990): Testing the Rationality of Price Forecasts: New Evidence form Panel Data. *American Economic Review*, 80:714–735.
- Kirchgässner, G. und J. Wolters (2007): *Introduction to Modern Time Series Analysis*. Springer Verlag, Berlin Heidelberg.
- Kolmogorov, A. N. (1933): Sulla determinazione empirica di una legge di distribuzione. *Giornale dell' Istituto Italiano degli Attuari*, 4:83–91.

- Körner, J. (1983): Probleme der Steuerschätzung. *Staatsfinanzierung im Wandel: Verhandlungen auf der Jahrestagung des Vereins für Socialpolitik, Gesellschaft für Wirtschafts- und Sozialwissenschaften in Köln 1982*, -:215–252.
- Körner, J. (1995): Steuerschätzungen: keine weiteren Mindereinnahmen. *ifo Schnelldienst*, 48(31):14–17.
- Kwiatkowski, D.; P. C. B. Phillips; P. Schmidt und Y. Shin (1992): Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root. *Journal of Econometrics*, 54:159–178.
- Lehmann, R. (2010): Die Steuerschätzung in Deutschland - eine Erfolgsgeschichte? *ifo Dresden*, 3:34–37.
- Leibrecht, M. (2003): Determinanten einer präzisen Steuerschätzung. *Das Öffentliche Haushaltswesen*, 44(3/4):132–155.
- Leibrecht, M. (2004): *Steuerschätzung in Österreich: Ablauf, Methoden und Präzision der Prognose*. DUV.
- von der Lippe, P. (1986): Steuerschätzungen und ihre wirtschaftspolitische Bedeutung - Der Steuerschätzung und seine Tätigkeit. *Steuer und Studium*, 11:334–340.
- Ljung, G. M. und G. E. P. Box (1978): On a Measure of Lack of Fit in Time Series Models. *Biometrika*, 65:297–303.
- Lütkepohl, H. (2004): *Applied Time Series Econometrics*. Cambridge University Press.
- Makridakis, S.; S. C. Wheelwright und R. J. Hyndman (1998): *Forecasting: Methods and Applications*. Wiley & Sons.
- Mincer, J. und V. Zarnowitz (1969): The Evaluation of Economic Forecasts. In: *Economic Forecasts and Expectations*, Hg. J. Mincer. National Bureau of Economic Research, New York, S. 3–46.
- Nordhaus, W. D. (1987): Forecasting Efficiency: Concepts and Applications. *Review of Economics and Statistics*, 69:667–674.

- Osterloh, S. (2008): Accuracy and Properties of German Business Cycle Forecasts. *Applied Economics Quarterly*, 54:27–57.
- Perron, P. (1989): The Great Crash, The Oil Price Shock, and the Unit Root Hypothesis. *Econometrica*, 57:1361–1401.
- Perron, P. (1994): Trend, Unit Root and Structural Change in Macroeconomic Time Series. In: *Cointegration for the Applied Economist*, Hg. B. Bhaskara Rao. St. Martins Press, New York, S. 113–146.
- Perron, P. (1997): Further Evidence on Breaking Trend Functions in Macroeconomic Variables. *Journal of Econometrics*, 80:355–385.
- Pindyck, R. S. und D. L. Rubinfeld (1998): *Econometric Models and Economic Forecasts*. McGraw-Hill, Boston, 4. Aufl.
- Quenouille, M. H. (1949): The Joint Distribution of Serial Correlation Coefficients. *Annals of Mathematical Statistics*, 20:561–571.
- Rürup, B. und M. Schachler (1983): Methoden und Probleme der Steuerschätzung. *WiSt Wirtschaftswissenschaftliches Studium*, 12. Jahrgang, Heft 5:217–223.
- Schaft, W. (1981): Risiken der Steuerschätzung. *Wirtschaftsdienst*, 61(10):488–490.
- Schäfers, M. (2005): Steuerschätzung – der Mai ist kein Wonnemonat für Finanzminister. *FAZ*, 107:16.
- Schoof, D. (1998): Von den Schwierigkeiten der Steuerschätzung. *DStZ*, Nr. 6:209–214.
- Schwarz, G. (1978): Estimating the Dimensions of a Model. *Annals of Statistics*, 6:461–464.
- Shapiro, S. S. und M. B. Wilk (1965): An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3/4):591–611.
- Smirnov, N. V. (1939): Estimate of Deviation Between Empirical Distribution Functions in Two Independent Samples. *Bulletin Moscow University*, 2:3–16.
- StataCorp (2009): *Stata Time-Series Reference Manual - Release 11*. Stata Press.

-
- Theil, H. (1966): *Applied Economic Forecasting*. North-Holland Pub. Co.
- Wagenhals, G. (2004): Tax-benefit Microsimulation Models for Germany: A Survey. *IAW-Report*, 32(1):55–74.
- Wagenhals, G. (2011): Dual Income Tax Reform in Germany. A Microsimulation Approach. *International Journal of Microsimulation*, 4(2):3–13.
- Wold, H. (1938): *A Study in the Analysis of Stationary Time Series*. Almqvist and Wicksell, Stockholm.
- Wooldridge, J. M. (2009): *Introductory Econometrics - A Modern Approach*. Thomson South-Western.
- Zivot, E. und D. W. K. Andrews (July 1992): Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis. *Journal of Business and Economic Statistics*, 10(3):251–270.