

Dynamische Modellierung der Zinsstruktur von Staatsanleihen

Bachelorarbeit zur Erlangung des Grades

Bachelor of Science (B.Sc.)

im Studiengang Volkswirtschaftslehre

an der Rheinischen Friedrich-Wilhelms-Universität Bonn

Themensteller: Prof. Dr. Dominik Liebl

vorgelegt im Januar 2022 von:

Justin Franken

Matrikelnummer: 3233058

Inhaltsverzeichnis

1	Einleitung	1
2	Modellierung der Zinsstruktur durch die Faktorenanalyse	2
2.1	Das Modell der Faktorenanalyse	2
2.2	Herleitung der Faktorenanalyse	3
2.3	Bestimmung der Anzahl an relevanten Faktoren	5
3	Weitere Methoden zur Modellierung der Zinsstruktur	6
3.1	Das Nelson und Siegel Modell	7
3.2	Das Nelson und Siegel Modell nach Diebold und Li	9
3.3	Die Erweiterung des Nelson und Siegel Modells nach Svensson	11
4	Dynamische Vorhersage von Zinsstrukturkurven	12
4.1	Dynamisierung unserer Modellgleichungen	12
4.2	Hinzunahme weiterer Variablen für die dynamischen Vorhersage	13
5	Anwendung	15
5.1	Verwendete Daten	15
5.1.1	Anzahl an verwendeten Faktoren in der Faktorenanalyse	16
5.2	Modellierung von US-Staatsanleihen	17
5.3	Vorhersage von US-Staatsanleihen	19
6	Schlussbemerkung	20
7	Literaturverzeichnis	i
8	Anhang	ii
9	Schriftliche Versicherung	v

1 Einleitung

In der Finanzliteratur hat sich die Zinsstrukturkurve etabliert, um das Verhältnis verschiedener Zinssätze bezüglich der kurzen, mittleren oder langen Laufzeiten von Anleihen grafisch darzustellen. Jedoch sollte eine aussagekräftige Zinsstrukturkurve die verschiedenen Einflussfaktoren im Wirtschaftsgeschehen genau abbilden. Denn makroökonomische Einflussfaktoren, wie die Inflation, die reale ökonomische Aktivität, Leitzinsen der Zentralbanken oder Steuersätze der Fiskalpolitik, wirken sich auf die Form der Zinsstruktur unterschiedlich stark aus. Und während insbesondere die geldpolitischen Entscheidungen der Zentralbanken die kurzfristigen Zinsen entscheidend beeinflussen, wird der langfristige Zins deutlicher durch die Erwartung der Marktteilnehmer geprägt. Damit diese komplexe Datenlage mit der Zinsstrukturkurve adäquat erfasst werden kann, bietet sich als Verfahren die Faktorenanalyse an. Die Faktorenanalyse macht unbeobachtete Zusammenhänge in der Struktur von Daten sichtbar und zeigt diese in einer komprimierten Fassung auf. Dabei wird häufig die Dimension der ursprünglichen Daten erheblich reduziert. So hat [Bliss \(1997\)](#) durch eine Faktorenanalyse die Zinsstruktur mit lediglich drei Faktoren beschreiben können. Eine leicht andere Herangehensweise ist in [Diebold und Li \(2006\)](#) oder [Svensson \(1995\)](#) beschrieben. Hier werden auch Faktoren herangezogen, aber die Faktorgewichte sind jeweils durch die Methode von Nelson und Siegel oder derjenigen von Svensson vorab festgelegt. Dadurch, dass die jeweiligen Gewichte nicht mehr diskret, sondern stetig sind, lässt sich die Zinsstrukturkurve zusätzlich stetig darstellen. Und weil sich mit den ermittelten Faktoren die Struktur der Daten in kompakter Form aufbereiten lässt, können diese auch für die Prognose zukünftiger Zinsstrukturkurven genutzt werden. Für die Bestimmung von Faktoren zukünftiger Perioden wird häufig auch auf autoregressive oder vektorautoregressive Modelle zurückgegriffen. Die Prognosekraft dieser Modelle lässt sich durch die Aufnahme zusätzlicher Variablen optimieren. Für die Vorhersage von Staatsanleihen sollten am besten makroökonomische Variablen genutzt werden, weil diese sehr wahrscheinlich einen starken Einfluss auf die Zinsstruktur haben. [Diebold, Rudebusch und Aruoba \(2006\)](#) betrachteten dazu die Inflationsrate, die Leitzinsen und die Produktionsauslastung in einer Volkswirtschaft. Aufgrund dessen sollten sich zukünftige Zinsstrukturkurven dynamisch besser vorhersagen lassen. Um die Relevanz der Faktorenanalyse und der modifizierten Ansätze für die Berechnung einer aussagekräftigen Zinsstruktur zu verdeutlichen, ist die Bachelorarbeit wie folgt aufgebaut. Im zweiten Abschnitt werden wir die Faktorenanalyse herleiten, um damit die Zinsstruktur zu erklären. Anschließend wird im dritten Abschnitt die Nelson- und Siegel-Methode nach [Diebold und Li \(2006\)](#) sowie das Verfahren von [Svensson \(1995\)](#) betrachtet. Dabei gehen wir insbesondere auf die jeweils festgelegten Faktorladungen ein und wie diese die unterschiedlichen Laufzeiten der Zinsen von Staatsanleihen beeinflussen. Im vierten Abschnitt schauen wir uns an, wie die modellierten Zinsstrukturkurven mittels autoregressiver Prozesse AR(1) oder vektorautoregressiver Prozesse VAR(1) dynamisch vorhergesagt werden können. Diese Konzepte erweitern wir anschließend um vordefinierte makroökonomische Variablen. Im fünften Abschnitt testen wir, wie gut die vorher beschriebenen Verfahren Zinsstrukturkurven modellieren können. Zum Schluss werden wir im sechsten Abschnitt ein Fazit zu dieser Bachelorarbeit ziehen.

2 Modellierung der Zinsstruktur durch die Faktorenanalyse

2.1 Das Modell der Faktorenanalyse

Um die Faktorenanalyse herzuleiten, sei unser Set an Beobachtungen repräsentiert durch die $(T \times N)$ Matrix X . T ist die totale Anzahl an Beobachtungen und N ist die totale Anzahl an betrachteten Laufzeiten der Staatsanleihe, pro Periode.

Im Folgenden wollen wir nur die zentrierten Beobachtungen Y nutzen, wobei Y definiert wird, als $Y_i = X_i - \mu_i$, wobei $\mu_i = \frac{1}{T} \sum_{t=1}^T x_{it}$ mit $i = 1, \dots, N$ und $t = 1, \dots, T$. Anschließend wird angenommen, dass die N Variablen der zentrierten Beobachtungsmatrix Y , einen linearen Zusammenhang zu r , $r < N$, unbeobachteten Faktoren wie folgt haben:

$$y_{it} = \underset{1 \times r}{\lambda_i'} \times \underset{r \times 1}{f_t} + u_{it} \quad (1)$$

In Matrix Notation erhalten wir:

$$\underset{N \times 1}{y_t} = \underset{N \times r}{\Lambda} \times \underset{r \times 1}{f_t} + \underset{N \times 1}{u_t} \quad (2)$$

$$\underset{T \times N}{Y} = \underset{T \times r}{F} \times \underset{r \times N}{\Lambda'} + \underset{T \times N}{U} \quad (3)$$

Man beachte, dass neben der linken Seite in Gleichung (1), (2) und (3) zu Beginn keine der Parameter bekannt ist. Wir müssen somit jeden der Parameter auf der rechten Seite beschreiben und identifizieren.

F beziehungsweise f_t sind die jeweiligen Faktoren, die die zugrundeliegende Struktur in den zentrierten Daten von Y , y_t oder y_{it} in eine komprimierte Form, mit $r < N$, erfassen sollen. Λ sind die Faktorladungen, die den berechneten Faktoren ihr jeweiliges Gewicht zuordnet, damit die Faktoren entsprechend auf die jeweiligen Variablen geladen werden. U , u_t und u_{it} sind zu guter Letzt die Fehler der Schätzgleichungen.

Wir nehmen zur Identifizierung des Modells an, dass $u_{it} \stackrel{iid}{\sim} N(0, \sigma_u^2)$, $f_t \stackrel{iid}{\sim} N(0, \Sigma_f)$ und $E(f_t u_{it}) = 0$. Betrachten wir an dieser Stelle die Kovarianz eines Querschnitts zweier Variablen:

$$\begin{aligned} cov(y_{it} y_{jt}) &= E(y_{it} y_{jt}) = E(\lambda_i' f_t f_t' \lambda_j) = \lambda_i' E(f_t f_t') \lambda_j \\ &= \lambda_i' \Sigma_f \lambda_j \end{aligned}$$

Aus dieser Darstellung lässt sich erkennen, dass die Beziehung zwischen den Variablen i und j im Querschnitt zum Zeitpunkt t , durch die Kovarianz der gemeinsamen Faktoren Σ_f und die jeweiligen Faktorgewichte λ_i und λ_j bestimmt wird. Somit ist der Zusammenhang der Beobachtungen durch die gemeinsamen Faktoren gekennzeichnet, da die jeweiligen Ladungen genutzt werden, die Faktoren entsprechend auf y_{it} und y_{jt} zu laden.

Sei an dieser Stelle darauf verwiesen, dass wir statt f_t auch $f_t^* = Q' f_t$ und statt Λ auch $\Lambda^* = \Lambda Q$ schreiben können, wobei Q irgendeine orthogonale ($r \times r$) Matrix mit $Q'Q = QQ' = I_r$ darstellt. Aus unserer Modellgleichung erschließt sich zum Zeitpunkt t anschließend:

$$y_t = \Lambda^* f_t^* + u_t = \Lambda Q Q' f_t + u_t = \Lambda f_t + u_t$$

Daraus lässt sich schließen, dass das Modell durch eine beliebige orthogonale Matrix Q verändert werden kann. Wenn wir nun zusätzlich annehmen, dass $\Sigma_f = I_r$, folgt für unsere Annahmen an das Modell:

$$E(f_t^*) = Q' E(f_t) = 0 \quad \text{und für} \quad \text{cov}(f_t^*) = Q' \text{cov}(f_t) Q = I_r Q' Q = I_r$$

Somit können wir die Matrix Q Annahmen konsistent in das Modell hinzufügen. Häufig wird dann versucht Q dazu zu nutzen, dass die i -te Ladung Λ_i , den Faktor i , hoch auf einige wenige Variablen und klein bis gar nicht auf die anderen Variablen lädt. Dadurch ist anschließend eine ökonomisch wertvolle Interpretation der Faktoren möglich.

In dieser Arbeit liegt der Fokus jedoch nicht auf der Interpretation der Faktoren, sondern auf der Modellierung der Zinsstruktur, weswegen wir unsere Lösungen nicht rotieren werden und darum auf die Einschränkung $\Sigma_f = I_r$ verzichten.

2.2 Herleitung der Faktorenanalyse

Nun müssen wir noch die freien Parameter Λ und F aus der Modellgleichung schätzen. Betrachten wir dazu Gleichung (2). Dadurch, dass uns y_t bei der Schätzung bekannt ist und die Residuen u_t unabhängig voneinander und normalverteilt sind, möchten wir f_t und Λ durch das Minimieren des kleinsten quadrierten Fehlers zum Zeitpunkt t finden. Der kleinste quadrierte Fehler ergibt sich durch: $u_t' u_t = (y_t - \Lambda f_t)' (y_t - \Lambda f_t)$

Dabei wollen wir die Summe der quadratischen Residuen, unter der Nebenbedingung $\Lambda \Lambda' = \Lambda' \Lambda = I_r$ minimieren. Somit sind die jeweiligen Faktorladungen voneinander unabhängig und auf Varianz eins normiert. Wir erhalten folgendes Minimierungsproblem:

$$\min_{\Lambda, f_t} \sum_{t=1}^T u_t' u_t = \min_{\Lambda, f_t} \sum_{t=1}^T (y_t - \Lambda f_t)' (y_t - \Lambda f_t) \quad \text{s.d.} \quad \Lambda' \Lambda = I \quad (4)$$

In Matrix-Notation, für alle Perioden T , erhalten wir:

$$\min_{\Lambda, F} \text{tr}[(Y - F \Lambda')' (Y - F \Lambda')] \quad \text{s.d.} \quad \Lambda' \Lambda = I$$

Dadurch, dass wir F auch durch den multiplen Regressionsschätzer bestimmen können, schreiben wir F als: $F = \underbrace{(\Lambda' \Lambda)^{-1}}_{=I} \Lambda' Y = \Lambda' Y$. Daraus folgt:

$$\begin{aligned}
& \min_{\Lambda, F} tr[(Y - F\Lambda')'(Y - F\Lambda')] \quad \text{s.d.} \quad \Lambda'\Lambda = I \\
& = \min_{\Lambda} tr[(Y - \Lambda'Y\Lambda')'(Y - \Lambda'Y\Lambda')] \quad \text{s.d.} \quad \Lambda'\Lambda = I \\
& = \min_{\Lambda} tr[Y'Y] - tr[\Lambda'Y'Y\Lambda] \quad \text{s.d.} \quad \Lambda'\Lambda = I
\end{aligned}$$

Wenn $tr[Y'Y] - tr[\Lambda'Y'Y\Lambda]$ über Λ minimiert werden soll, können wir dies auch durch das Maximieren von $tr[\Lambda'Y'Y\Lambda]$ über Λ erreichen. Für den Fall, indem wir nur einen Faktor betrachten, ergibt sich nun das Minimierungsproblem zu:

$$\max_{\Lambda} \Lambda' \hat{\Sigma}_Y \Lambda - \gamma(\Lambda' \Lambda - 1)$$

wobei γ der Lagrange-Multiplikator für unsere Nebenbedingung ist und $\hat{\Sigma}_Y = T^{-1}Y'Y$ die Kovarianzmatrix der zentrierten Beobachtungen darstellt. Die erste Ableitung ergibt:

$$(\gamma I_N - \hat{\Sigma}_Y)\hat{\Lambda} = 0 \quad \Leftrightarrow \quad \hat{\Sigma}_Y \hat{\Lambda} = \hat{\Lambda} \gamma \quad (5)$$

Aus Gleichung (5) erkennen wir, dass $(\gamma I_N - \hat{\Sigma}_Y)$ groß wird, durch den größten Eigenwert γ der berechneten Kovarianzmatrix $\hat{\Sigma}_Y$. Somit gilt für $\hat{\Lambda}$, dass dieser der Eigenvektor, zu dem größten Eigenwert γ der ermittelten Kovarianzmatrix $\hat{\Sigma}_Y$ sein muss. Der Faktor ergibt sich nachfolgend aus dem kleinsten Quadrate Schätzer der einfachen Regression, wobei Λ in Gleichung (3) durch $\hat{\Lambda}$ ersetzt wurde:

$$\hat{F} = Y \hat{\Lambda} \underbrace{(\hat{\Lambda}' \hat{\Lambda})^{-1}}_{=I} = Y \hat{\Lambda}$$

Betrachten wir an dieser Stelle die Kovarianz des geschätzten Faktors:

$$cov(\hat{F}) = cov(Y \hat{\Lambda}) = \hat{\Lambda}' cov(Y) \hat{\Lambda} = \hat{\Lambda}' \hat{\Sigma}_Y \hat{\Lambda}$$

Aus Gleichung (5) wissen wir, dass $\hat{\Sigma}_Y = \hat{\Lambda} \gamma \hat{\Lambda}'$. Somit gilt für die Kovarianz von \hat{F} :

$$cov(\hat{F}) = \underbrace{\hat{\Lambda}' \hat{\Lambda}}_{=1} \gamma \underbrace{\hat{\Lambda} \hat{\Lambda}'}_{=1} = \gamma \quad (6)$$

Wir sehen also, dass die Varianz des geschätzten Faktors \hat{F} , dem größten Eigenwert γ entspricht. Weiteren wir das oben beschriebene Konzept, auf den Fall $r > 1$ aus. Wenn Gleichung (5) ein wenig umgeformt und in Matrix Notation geschrieben wird, erhalten wir:

$$\begin{array}{ccccccc}
\hat{\Sigma}_Y & = & \hat{\Lambda} & \times & \underbrace{\Gamma}_{(\tilde{\gamma} \times I_N)} & \times & \hat{\Lambda}' & (\tilde{\gamma} = (\gamma_1, \dots, \gamma_N)) & (7) \\
(N \times N) & & (N \times N) & & (N \times N) & & (N \times N) & &
\end{array}$$

Aus Gleichung (5) und (7) wird deutlich, dass wir das Minimierungsproblem der Gleichung (4) lösen können, durch die ersten r Spalten der Eigenvektormatrix $\hat{\Lambda}$, die durch die r korrespondierenden, größten Eigenwerte der diagonalen Eigenwertmatrix Γ , mit $\gamma_1 \geq \gamma_2 \geq \dots \geq \gamma_r$ bestimmt werden. Wobei $\hat{\Lambda}$ und Γ die Eigenvektoren und Eigenwerte der Kovarianzmatrix unserer zentrierten Beobachtung sind. Somit sind die geschätzten Parameter der Modellgleichung (3) gegeben durch $\hat{\Lambda}_r$ und \hat{F}_r . Die Faktoren \hat{F}_r ergibt sich nun aus dem multiplen Regressions-schätzer:

$$\hat{F}_r = \underbrace{(\hat{\Lambda}_r' \hat{\Lambda}_r)^{-1}}_{=I} \hat{\Lambda}_r' Y = \hat{\Lambda}_r' Y \quad (8)$$

Auch hier ergibt sich die Kovarianz von \hat{F}_r zu $cov(\hat{F}_r) = \underbrace{\hat{\Lambda}_r' \hat{\Lambda}_r}_{=1} \Gamma_r \underbrace{\hat{\Lambda}_r \hat{\Lambda}_r'}_{=1} = \Gamma_r$ und somit wird klar, dass die geschätzten Faktoren unabhängig voneinander sind und Varianz entsprechend des jeweiligen Eigenwerts haben. Also erklärt der erste Faktor \hat{F}_1 bereits $\frac{\gamma_1}{\sum_{i=1}^r \gamma_i}$ -Prozent der gesamten Faktoren. Zum Schluss können wir nun unsere zentrierte Zinsstruktur, durch die geschätzten Faktoren \hat{F}_r , gewichtet mit den geschätzten Faktorladungen $\hat{\Lambda}_r$, beschreiben durch:

$$\begin{array}{ccccc} \hat{Y} & = & \hat{F}_r & \times & \hat{\Lambda}_r' \\ (T \times N) & & (T \times r) & & (r \times N) \end{array}$$

Die geschätzten Zinsen zum Zeitpunkt t sind dann gegeben durch

$$\hat{y}_t = \hat{\Lambda}_r' \hat{f}_{rt}$$

und der geschätzte Zins zum Zeitpunkt t , für die Laufzeit i wird dann berechnet mittels

$$\hat{x}_{it} = \hat{\mu}_i + \hat{\lambda}_i' \hat{f}_{rt} \quad (9)$$

wobei $\hat{\mu}_i$ der Mittelwert der beobachteten Laufzeit i , über alle Perioden T ist.

2.3 Bestimmung der Anzahl an relevanten Faktoren

Bisher haben wir angenommen, dass es r Faktoren gibt, die einen linearen Zusammenhang zu den N beobachteten Variablen von Y haben. Aus der Herleitung unserer geschätzten Faktoren, haben wir jedoch insbesondere durch Gleichung (7) sehen können, dass N Eigenvektoren $\hat{\Lambda}_N$ aus der Kovarianzmatrix $\hat{\Sigma}_Y$ ermittelt werden können. Aus Gleichung (8) wird dann klar, dass es somit potenziell auch N Faktoren \hat{F}_N geben muss, die die Zinsstruktur beschreiben können. Zur Identifizierung der $r < N$ Faktoren existieren mehrere Herangehensweisen. Aus der Herleitung ist uns bereits bekannt, dass der i -te Eigenwert γ_i , bereits $\frac{\gamma_i}{\sum_{j=1}^N \gamma_j}$ -Prozent aller N Faktoren erklären kann. Da alle N Faktoren das gesamte Modell beschreiben, beschreiben somit die ersten r Faktoren $\frac{\sum_{i=1}^r \gamma_i}{\sum_{j=1}^N \gamma_j}$ -Prozent des Modells. Um nun r zu bestimmen, werden so viele Faktoren

r berücksichtigt, bis für einen vordefinierter Schwellenwert gilt: $\tau \leq \frac{\sum_{i=1}^r \gamma_i}{\sum_{j=1}^N \gamma_j}$. So wird häufig in ökonomischen Kontexten, ab $\tau = 90\%$ der erklärten Varianz, kein weiterer Faktor hinzugezogen. Eine weitere Methode ist der "Screeplot". In diesem werden die geordneten N Eigenwerte $\gamma_1, \dots, \gamma_N$ in einer zweidimensionalen Grafik dargestellt. Da die Eigenwerte der Größe nach geordnet sind, erhält man eine konvex verlaufende Kurve. Nun sucht man im Screeplot nach einem Knick¹ in der Kurve. An dieser Stelle wird dann aufgehört weitere Faktoren zu berücksichtigen, da der Erklärungsgehalt des betrachteten Eigenwerts an dieser Stelle stark nachlässt. Um die relevante Anzahl an Faktoren r formaler zu bestimmen, können wir die Methode von [Ahn und Horenstein \(2013\)](#) verwenden. Die Intuition ist eine ähnliche, wie die im Screeplot. Man möchte einen Punkt in den Eigenwerten finden, an dem der Abstand zwischen dem i -ten und dem $i+1$ -ten Eigenwert am größten ist, da an dieser Stelle der Erklärungsgehalt des $i+1$ -ten Eigenwertes am stärksten nachlässt. Sei $ER(i) = \frac{\gamma_i}{\gamma_{i+1}}$, für $i = 1, \dots, N$ suchen wir dann jenes \hat{r}_{ER} , welches

$$\hat{r}_{ER} = \max_{1 \leq i \leq N} ER(i) = \max_{1 \leq i \leq N} \frac{\gamma_i}{\gamma_{i+1}} \quad \text{erfüllt.}$$

Eine weitere häufig verwendete Methode zur Identifizierung des optimalen r 's, ist das Informationskriterium. Ein beliebter Aufsatz dazu ist von [Bai und Ng \(2002\)](#). Dabei betrachtet man zunächst die Summe der quadrierten Fehler $V(r) = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (y_{it} - \hat{y}_{it})^2$. Das Ziel ist es ein r zu finden, bei dem $V(r)$ klein ist. $V(r)$ wird aber immer kleiner für $r \rightarrow N$. Somit müssen wir $V(r)$ dafür bestrafen, wenn zu viele Faktoren r berücksichtigt werden, da wir sonst die ursprüngliche Dimension nicht reduzieren würden. Somit betrachten wir folgende Gleichung:

$$IC(r) = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (y_{it} - \hat{y}_{it})^2 + rg(N, T) = V(r) + rg(N, T)$$

Die beiden klassischsten Strafterme sind das *Bayesian Information Criterion* mit $g(N, T) = \ln(T)/T$ und das *Akaike Information Criterion* mit $g(N, T) = 2/T$. Bai und Ng empfehlen, dass man den Strafterm $g(N, T)$ so gestalten sollte, sodass $g(N, T) \rightarrow 0$ und $(\min\{\sqrt{N}, \sqrt{T}\})^2 \cdot g(N, T) \rightarrow \infty$ für N und $T \rightarrow \infty$, damit ein Modell ausgewählt wird, welches nicht über oder unter parametrisiert ist. Für weitere Ausführungen sei auf das Papier von Bai und Ng verwiesen. Wie viele Faktoren bei der Modellierung von Staatsanleihen benötigt werden, wurde in [Bliss \(1997\)](#) behandelt. Dabei wurden amerikanische Staatsanleihen in dem Zeitraum von 1970 bis 1995 durch nur drei der ursprünglichen 20 möglichen Faktoren beschrieben und erklärten bereits 95,3% der Varianz, wobei der erste Faktor bereits 80,6% der Varianz erklären konnte.

3 Weitere Methoden zur Modellierung der Zinsstruktur

Wir konnten nun sehen, dass die Faktorenanalyse einen hochdimensionalen Datensatz in kompakter Form beschreiben kann und dabei einen Großteil der Varianz erklärt. Dadurch, dass wir uns in einer Welt befinden, in der die Größe an gesammelten Datenmengen immer weiter zu-

¹Häufig wird dieser Knick auch mit einem Ellenbogen verglichen.

nimmt und deren Interpretation zunehmend schwerer wird, findet die Faktorenanalyse nicht nur Anwendung in der Modellierung einer Zinsstruktur, sondern auch in anderen Bereichen, in denen man die Komplexität des Datensatzes reduzieren oder zusammenfassen möchte.

Es existieren jedoch auch Methoden, die ausschließlich für die Modellierung der Zinsstruktur geschaffen wurden. Diese unterscheiden sich zu der Faktorenanalyse zum Beispiel dadurch, dass diese eine stetige Kurve beschreiben, in der jede erdenkliche Laufzeit t_i , zwischen der kleinsten und größten Laufzeit der Zinsstrukturkurve ($t_1 \leq t_i \leq t_M$), durch einen entsprechenden Zins definiert ist. Dabei muss zunächst ein Weg gefunden werden, den Raum zwischen den beobachteten Zinsen einer Staatsanleihe genau zu definieren.

So hat [McCulloch \(1971, 1975\)](#) die Zinsstruktur mittels polynomischer Splines definiert. Dabei wird der Raum zwischen zwei beobachteten Zinspunkten durch ein Polynom n-ten Grades beschrieben. Dies wird dann so oft wiederholt, bis jeder beobachtete Zinspunkt miteinander verbunden wurde. Eine wichtige Annahme ist dann, dass an dem beobachteten Zinspunkt, an dem zwei berechnete Polynome miteinander verbunden werden, die Polynome dieselbe Ableitung besitzen, wodurch ein nahtloser und stetiger Übergang, zum nächsten berechneten Polynom entsteht. Alle Polynome (n-ten Grades) zusammen, beschreiben anschließend die Zinsstrukturkurve, welche für jeden Zins definiert ist. Das Problem dieser Herangehensweise ist, dass ein Polynom n-ten Grades, gerade Kurven schlecht beschreiben kann und dass das Polynom welches die zwei Zinspunkte miteinander verbunden hat, nicht gegen einen bestimmten Wert konvergiert, sondern in das Unendliche divergiert. Dadurch werden monotone Zinsstrukturkurven und Zinsen zwischen zwei weiter entfernten, beobachteten Laufzeiten, schlecht beschrieben und vorhergesagt.

3.1 Das Nelson und Siegel Modell

Eine Möglichkeit dem Problem der Polynomial Splines entgegen zu wirken, ist der Ansatz welcher von [Nelson und Siegel \(1987\)](#) entwickelt wurde. Um das Modell von Nelson und Siegel zu verstehen, müssen wir zunächst begreifen, wie die Spot Raten der am Markt beobachteten Preise von Staatsanleihen bestimmt werden. Da stehen wir nämlich vor dem Problem, dass die Spot Raten am Markt nicht direkt beobachtbar sind. Deswegen müssen wir diese zunächst implizit aus bestehenden Staatsanleihen Preisen berechnen. Sei $i(t, T)$ der Spot Zins für eine Staatsanleihe, zum Zeitpunkt t , welche zum Zeitpunkt T ausläuft. Dies ist der tatsächliche Zins, den ein Anleger für eine Investition in die Staatsanleihe zum Zeitpunkt t für die gesamte Laufzeit erhält. Die Laufzeit ist dann gegeben durch $m = T - t$. Der Preis der betrachteten Staatsanleihe zum Zeitpunkt t , für eine Investition von einem Euro ist formal:

$$S(t) = \exp \left(-\frac{i(t, T)}{100} (T - t) \right) \quad (10)$$

Da uns aus Gleichung (10) der aktuelle Spot-Preis der Anleihe $S(t)$, als auch die Laufzeit m bekannt ist, können wir den impliziten Spot Zins $\hat{i}(t, T)$ aus Gleichung (10) berechnen.

An dieser Stelle sei darauf verwiesen, dass der betrachtete Spot-Preis der Staatsanleihe $S(t)$,

der Preis einer Nullkuponanleihe ist. Für implizite Spot Zinsen von Staatsanleihen, deren Laufzeiten länger als 12 Monate sind, wird es zunehmend schwerer gehandelte Nullkuponanleihen am Markt zu finden, weswegen dann auf Kuponanleihen zurückgegriffen werden muss. Dann ist der Spot-Preis der Kuponanleihe jedoch nicht mehr durch Gleichung (10) beschrieben, da nun zusätzlich regelmäßige Kuponzahlungen anfallen. Wie aus Kuponanleihen der Nullkuponpreis berechnet werden kann, wird unter anderem in [Svensson \(1995\)](#) erklärt. In diesem Teil der Bachelorarbeit nehmen wir an, dass der implizite Spot Zins nur aus Nullkuponstaatsanleihen entnommen wird. Die implizite Forward Rate für eine Nullkuponstaatsanleihe zum Zeitpunkt t , welche zum Zeitpunkt $t' > t$ gekauft und bis $T > t'$ gehandelt wird, ist dann beschrieben durch:

$$\hat{f}(t, t', T) = \frac{(T - t) \hat{i}(t, T) - (t' - t) \hat{i}(t, t')}{T - t'}$$

Durch diese Gleichung sind wir nun in der Lage, den Zins eines Produktes zum Zeitpunkt t , welches für den Zeitraum t' bis T gehandelt wird, aus den impliziten Spot Raten $\hat{i}(t, T)$ und $\hat{i}(t, t')$ zu berechnen. Um daraus die echten Spot Raten zum Zeitpunkt t mit Laufzeit $m = T - t$ zu finden, betrachten wir zunächst die kurzfristige Forward Rate:

$$f(t, t') = \lim_{T \rightarrow t'} f(t, t', T)$$

Dabei begutachten wir eine Forward Rate, dessen Laufzeit infinitesimal klein wird. Die kurzfristige Forward Rate kann dann als ein marginaler Anstieg der Rendite einer Anlage, aufgrund eines marginalen Anstiegs der Laufzeit verstanden werden. Somit sind Spot Raten und kurzfristige Forward Raten durch die folgende Gleichung miteinander verwandt:

$$i(t, T) = \frac{\int_{\tau=t}^T f(t, \tau) d\tau}{T - t} \quad (11)$$

Ebenso können wir mit derselben Logik, die Forward Rate zum Zeitpunkt t , welche von t' bis T gehandelt werden, durch folgende Gleichung berechnen:

$$f(t, t', T) = \frac{\int_{\tau=t'}^T f(t, \tau) d\tau}{T - t'} \quad (12)$$

Gleichung (11) und (12) können auch als Durchschnitte der kurzfristigen Forward Raten, mit Laufzeiten t bis T oder t' bis T zum Zeitpunkt t verstanden werden. Nelson und Siegel nahmen nun an, dass die kurzfristige Forward Rate der Zinsstrukturkurve durch die Lösung einer Differentialgleichung 2. Ordnung mit geraden und realen Wurzeln definiert ist. Diese Lösung kann ebenfalls als eine Laguerre Funktion verstanden werden, welche eine beliebige Approximationsfunktion in der Mathematik darstellt. Die Lösung ist durch die folgende Gleichung charakterisiert:

$$f_{tm}(b) = \beta_{0t} + \beta_{1t} \exp\left(-\frac{m}{\tau}\right) + \beta_{2t} \frac{m}{\tau} \exp\left(-\frac{m}{\tau}\right) \quad (13)$$

Wobei $b = (\beta_{0t}, \beta_{1t}, \beta_{2t}, \tau)$ der Vektor an zu berechnenden Parametern der Modellgleichung, zum Zeitpunkt t ist. Diese Funktionale Form dient dann der Beschreibung der kurzfristigen Forward Raten $f(t, t')$ zum Zeitpunkt t , mit Laufzeit m ($m = T \rightarrow t'$) und Handelstag t' .

Bei genauerer Betrachtung lässt Gleichung (13) erkennen, wie die Parameter β_0 , β_1 und β_2 , aufgrund der jeweiligen Gewichtungen, für die Beschreibung einer Zinsstrukturkurve interpretiert werden können. Wenn die Laufzeit $m \rightarrow \infty$ steuert, dann wird der langfristige Zins fast ausschließlich durch β_0 erklärt, da die Gewichte vor β_1 und β_2 gegen null zielen. Daraus lässt sich schließen, dass β_0 für die Beschreibung der langfristigen Zinsen verantwortlich ist.

Wenn in Gleichung (13) hingegen die Laufzeit $m \rightarrow 0$ strebt, wird der sehr kurzfristige Zins ebenfalls durch β_0 aber auch durch β_1 bestimmt. Da für sehr große m das Gewicht von β_1 sehr klein wird, beschreibt β_1 fast ausschließlich die kurzfristigen Zinsen der Zinsstrukturkurve.

Der Parameter β_2 , erklärt am stärksten die Zinsen der Laufzeiten m , die zwischen ($0 < m_i, m_j < \infty$) liegen, da für $m = 0$ und für $m \rightarrow \infty$, das Gewicht von β_2 gegen null steuert. Dadurch lässt sich schließen, dass β_2 die mittelfristigen Zinsen der Zinsstrukturkurve beschreibt. Insbesondere in Abbildung 3.2 kann der unterschiedliche Einfluss der Faktorladungen, im Bezug auf die jeweilige Laufzeit m , gut erkannt werden.

Durch Gleichung (11) wissen wir nun auch, wie wir die Spot Zinsen einer Staatsanleihe aus den kurzfristigen Forward Raten (13), für die Laufzeit m , mit Handelstag t berechnen können:

$$i_{tm}(b) = \beta_{0t} + \beta_{1t} \frac{1 - \exp(-\frac{m}{\tau})}{\frac{m}{\tau}} + \beta_{2t} \left(\frac{1 - \exp(-\frac{m}{\tau})}{\frac{m}{\tau}} - \exp\left(-\frac{m}{\tau}\right) \right) \quad (14)$$

Wenn nun Gleichung (14) genauer betrachtet wird fällt auf, dass diese Ähnlichkeiten zu der Faktorenanalyse mit $r = 3$ hat. Wenn man davon ausgeht, dass die β 's Faktoren und die Terme vor den Faktoren festgelegte Faktorladungen sind, können wir auch hier die Faktoren β_0 , β_1 und β_2 durch den multiplen Regressionsschätzer finden. Dabei nehmen wir wie in der Faktorenanalyse an, dass die β 's linear unabhängig voneinander sind und die Regressionsgleichung unkorrelierte Fehler mit Erwartungswert null hat. Ein entscheidender Unterschied zu der Faktorenanalyse ist, dass im Nelson-Siegel Modell ein zusätzlicher Parameter τ , durch einen optimierenden Algorithmus, gefunden werden muss. Diesem Problem widmen wir uns in Abschnitt 3.2.

Gleichung (14) hat nun den entscheidenden Vorteil gegenüber der Modellierung mittels Polynomial Splines, dass monotone Zinsstrukturen durch β_0 besser beschrieben werden, als auch Zinsen, welche vorher nicht beobachtet wurden, da die resultierende Zinsstrukturkurve nach Nelson und Siegel, nicht wie ein Polynom divergiert. Darüber hinaus ist Gleichung (14) ebenfalls stetig für jedes m definiert und die ermittelten Faktoren lassen Interpretationen zu.

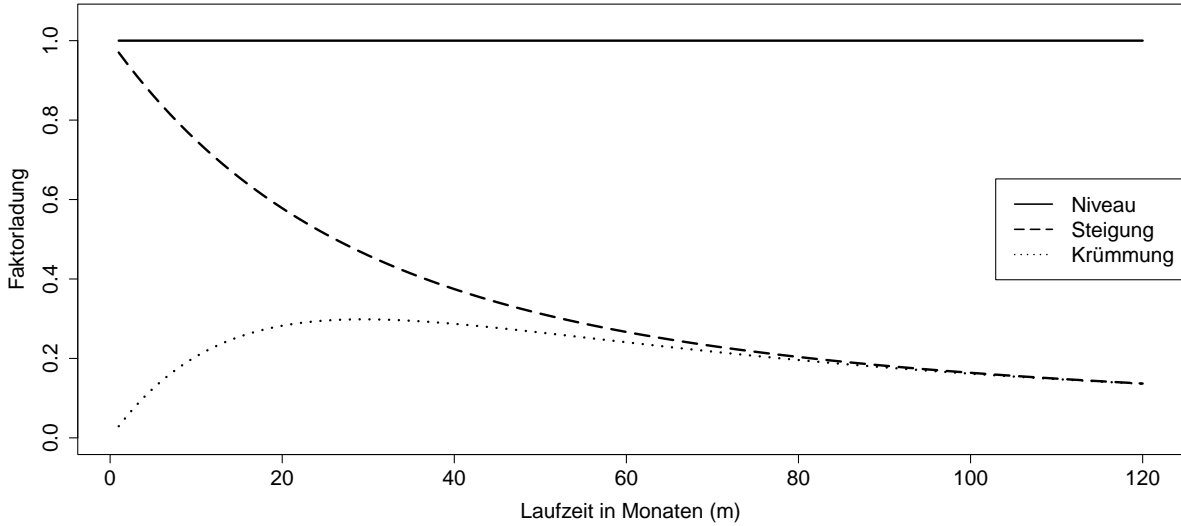
3.2 Das Nelson und Siegel Modell nach Diebold und Li

In der Arbeit von [Diebold und Li \(2006\)](#) wurde das Nelson-Siegel Modell wie folgt formuliert:

$$y_t(m) = \beta_{0t} + \beta_{1t} \frac{1 - \exp(-\lambda_t m)}{\lambda_t m} + \beta_{2t} \left(\frac{1 - \exp(-\lambda_t m)}{\lambda_t m} - \exp(-\lambda_t m) \right) \quad (15)$$

Wobei Gleichung (15) der Gleichung (14) entspricht, mit $\lambda_t = \frac{1}{\tau}$. Ein entscheidender Beitrag von Diebold und Li, ist die neue Interpretation der Faktoren, die Bestimmung von λ_t und die Dynamisierung der Nelson und Siegel Modellgleichung, was zu einem späteren Zeitpunkt behandelt wird. Betrachten wir dazu nochmals die jeweiligen Faktoren β_0 , β_1 und β_2 . Für β_0

Abbildung 3.2: Nelson–Siegel Faktorladungen



Anmerkung: Faktorladungen des Nelson-Siegel Modells nach Diebold und Li mit $\lambda = 0.0609$

wird deutlich, dass egal welche Laufzeit betrachtet wird, der Einfluss von β_0 auf die Zinsstrukturkurve immer der gleiche ist, da $y_t(\infty) = \beta_0$ ist. Diebold und Li definierten aus diesem Grund β_0 als das Niveau der Zinsstrukturkurve. Der Einfluss von β_1 ist besonders für kurzfristige Laufzeiten gegeben und nimmt exponentiell mit steigender Laufzeit ab. Dadurch wird die Zinsstrukturkurve stärker bei kürzeren und schwächer bei längeren Laufzeiten beeinflusst und verändert somit auch die Steigung der gesamten Zinskurve. Diebold und Li definierten die Steigung der Zinsstruktur als Differenz zwischen der größten und kleinsten betrachteten Laufzeit als $y_t(120) - y_t(3) = -0.78\beta_1 + 0.06\beta_2$. Man erkennt, dass der Einfluss von β_1 im Bezug auf die definierte Steigung von Diebold und Li am größten ist. Allgemein kann man dies auch als $y_t(\infty) - y_t(0)$ definieren, was genau $-\beta_1$ ergibt. Deswegen wird β_1 auch als Steigung der Zinsstrukturkurve interpretiert. Aus Nelson und Siegel wissen wir bereits, dass die dritte Faktorladung die mittleren Laufzeiten am stärksten beeinflusst und sehr klein wird, für besonders große oder kleine Laufzeiten. Daraus folgt, dass der Effekt von β_2 auf die Zinsstrukturkurve besonders für mittlere Laufzeiten ausgeprägt ist, was die Krümmung der Zinsstrukturkurve besonders beeinflusst. Diebold und Li definierten die Krümmung als zweimal einer mittleren Laufzeit, minus der größten und kleinsten Laufzeit. Es ergibt sich $2y_t(24) - y_t(3) - y_t(120) = 0.00053\beta_1 + 0.37\beta_2$. Auch hier wird deutlich, dass für die Definition der Krümmung, β_2 maßgeblich verantwortlich ist. Aus diesem Grund wird der dritte Faktor β_2 von Diebold und Li auch als Krümmungsfaktor der Zinsstrukturkurve verstanden. Alle drei Effekte sind auch sehr gut in [Abbildung 3.2](#) erkennbar.

Die Faktoren würden sich nun durch den kleinsten Quadrate-Schätzer bestimmen lassen, wenn für jede Periode t , ein λ_t bestimmt wäre. Diebold und Li entschieden sich deswegen, einen

festen Wert für λ_t zu bestimmen, um dadurch das Modell erheblich zu vereinfachen und die Schätzgenauigkeit der Faktoren zu erhöhen.

Um den Effekt von λ_t besser zu verstehen, betrachten wir zunächst Gleichung (15). Ein hohes λ_t sorgt für stark ausgeprägte β_1 und β_2 Faktorladungen und ein kleines λ_t sorgt für schwach ausgeprägte β_1 und β_2 Faktorladungen. Dadurch, dass λ_t bestimmt, wie steil die jeweiligen Faktorladungen verlaufen, kann dies auch dazu genutzt werden, das Maximum für die Krümmungsfaktorladung von β_2 zu bestimmen. Laut Diebold und Li, wird für solche Zwecke meist davon ausgegangen, dass das Maximum der dritten Faktorladung bei Laufzeiten zwischen zwei und drei Jahren liegt. Diebold und Li entschieden sich dann, den Mittelwert von 30 Monaten, zur Maximierung der Ladung zu wählen. Dies wird durch ein $\lambda_t = 0.0609$ erreicht. Dadurch ist der Effekt des dritten Faktors β_2 bei 30 Monaten am stärksten.

3.3 Die Erweiterung des Nelson und Siegel Modells nach Svensson

[Svensson \(1995\)](#) fiel damals auf, dass das Nelson-Siegel Modell bei komplexeren Zinsstrukturkurven schwächelte. Die Fehlergüte der berechneten Spotzinssätze der Nelson und Siegel Zinsstrukturkurve $i_{tm}(b)$, im Vergleich zu den echten Zinssätzen der Laufzeit m , waren insbesondere hoch, bei Zinsstrukturkurven die mehr als einen *Buckel* (oder U-Form) im Verlauf aufwiesen. Um dem Modell ein wenig mehr Flexibilität für komplexere Zinsstrukturkurven zu ermöglichen, erweiterte Svensson die Nelson-Siegel Gleichung um einen weiteren Faktor β_3 :

$$i_{tm}(b) = \beta_{0t} + \beta_{1t} \frac{1 - \exp\left(-\frac{m}{\tau_1}\right)}{\frac{m}{\tau_1}} + \beta_{2t} \left(\frac{1 - \exp\left(-\frac{m}{\tau_1}\right)}{\frac{m}{\tau_1}} - \exp\left(-\frac{m}{\tau_1}\right) \right) \quad (16)$$

$$+ \beta_{3t} \left(\frac{1 - \exp\left(-\frac{m}{\tau_2}\right)}{\frac{m}{\tau_2}} - \exp\left(-\frac{m}{\tau_2}\right) \right)$$

Wobei hier der Vektor an Modell-Parametern $b = (\beta_{0t}, \beta_{1t}, \beta_{2t}, \beta_{3t}, \tau_1, \tau_2)$ ist. Bei genauerer Betrachtung des Svensson-Modells wird deutlich, dass sich die neue vierte Faktorladung, von der dritten dadurch unterscheidet, dass dort ein zweites τ eingeführt wird. Also kann der neue Faktor β_3 , wie β_2 , als Krümmungsfaktor betrachtet werden, aufgrund der zwei verschiedenen Parameter τ_1 und τ_2 , mit unterschiedlicher Ausprägung und somit auch unterschiedlichem Maximum. Dadurch, dass nun eine zweite Krümmung berücksichtigt wird, kann das Svensson-Modell komplexere Strukturen besser erklären, als das alte Nelson-Siegel Modell, ohne bedeutend mehr Komplexität in die ursprüngliche Modellgleichung von Nelson und Siegel einzubringen. Jedoch ist die Vereinfachung von Diebold und Li, λ_t zu fixieren, in dem Svensson-Modell nicht wie in Diebold und Li umsetzbar, da nun zwei λ_t 's ($\lambda_{t1} = \frac{1}{\tau_1}, \lambda_{t2} = \frac{1}{\tau_2}$) bestimmt werden müssen². Aus diesem Grund empfiehlt Svensson, die Parameter in b , mittels des Maximum-Likelihood-Verfahrens zu lösen.

²Wobei grundsätzlich die Option besteht τ_1 und τ_2 zur Vereinfachung vorher festzulegen. Dadurch ist es anschließend möglich die Faktoren $\beta_0, \beta_1, \beta_2$ und β_3 durch eine multiple Regressionsanalyse zu schätzen und die Komplexität des Modells zu reduzieren.

4 Dynamische Vorhersage von Zinsstrukturkurven

Wir haben bisweilen drei Verfahren zur Zinsstrukturmodellierung herleiten können. Die Faktorenanalyse, das Nelson und Siegel Modell nach Diebold und Li und das Nelson und Siegel Modell nach Svensson. Alle drei Verfahren haben gemeinsam, dass die beschreibenden Parameter als Faktoren verstanden werden können und in dem Verfahren nach Diebold und Li oder Svensson, haben die Faktoren zusätzlich eine Interpretation erhalten.

An dieser Stelle wollen wir nun eine Dynamik in den vorher beschriebenen Modellen einbringen um eine Vorhersage dieser zu ermöglichen. Neben einer ökonomischen Interpretation zukünftiger Zinsstrukturen ist es zum Beispiel für Banken entscheidend zu wissen, wie sich die Zinsstrukturkurve in der Zukunft entwickeln wird um Derivate zu bepreisen, ein anständiges Portfoliomanagement zu planen oder Risikomanagement zu betreiben.

4.1 Dynamisierung unserer Modellgleichungen

Die Vorhersage erfolgt dabei, wie es auch in [Diebold und Li \(2006\)](#) erfolgt ist, in zwei Schritten. Der erste Schritt besteht aus der Vorhersage der Faktoren für den Zeitpunkt $t + 1$ aus den Faktoren des Zeitpunktes t , mittels eines Autoregressiven Modells erster Ordnung $AR(1)$, beziehungsweise Vektorautoregressiven Modells erster Ordnung $VAR(1)$. Erst im zweiten Schritt werden dann die Zinsen für den Zeitpunkt $t + 1$, durch die vorhergesagten Faktoren des Zeitpunktes $t + 1$ bestimmt. Die Dynamik der Faktoren wird dann wie folgt definiert:

1. Schritt: $AR(1)$ ohne Erweiterung

Faktorenanalyse: $(\hat{f}_{i,t+1} - \mu_i^f) = \theta_i^f (\hat{f}_{it} - \mu_i^f) + \eta_{i,t+1}^f$ für $i = 0, \dots, r$

Diebold und Li: $(\hat{\beta}_{i,t+1}^{DL} - \mu_i^{DL}) = \theta_i^{DL} (\hat{\beta}_{i,t+1}^{DL} - \mu_i^{DL}) + \eta_{i,t+1}^{DL}$ für $i = 0, 1, 2$

Svensson: $(\hat{\beta}_{i,t+1}^S - \mu_i^S) = \theta_i^S (\hat{\beta}_{i,t+1}^S - \mu_i^S) + \eta_{i,t+1}^S$ für $i = 0, 1, 2, 3$

$$(17)$$

1. Schritt: $VAR(1)$ ohne Erweiterung

Faktorenanalyse: $(\hat{f}_{t+1} - \mu^f) = \Theta^f (\hat{f}_t - \mu^f) + \eta_{t+1}^f$

Diebold und Li: $(\hat{\beta}_{t+1}^{DL} - \mu^{DL}) = \Theta^{DL} (\hat{\beta}_t^{DL} - \mu^{DL}) + \eta_{t+1}^{DL}$

Svensson: $(\hat{\beta}_{t+1}^S - \mu^S) = \Theta^S (\hat{\beta}_t^S - \mu^S) + \eta_{t+1}^S$

$$(18)$$

Dadurch, dass wir in Gleichung (17) und (18) die Faktoren jeweils um ihren Mittelwert reduziert haben, können wir das $AR(1)$ und das $VAR(1)$ Modell ohne eine Achsenabschnittskonstante schätzen. Der Mittelwert für μ_i^k ($k = (f, DL, S)$), ergibt sich dann aus den 1 bis t vorher berechneten Faktoren i . In dem $VAR(1)$ Modell sind die unabhängigen Variablen dann $\hat{f}_t = (\hat{f}_{1t}, \dots, \hat{f}_{rt})'$, $\hat{\beta}_t^{DL} = (\hat{\beta}_{0t}^{DL}, \hat{\beta}_{1t}^{DL}, \hat{\beta}_{2t}^{DL})'$ und $\hat{\beta}_t^S = (\hat{\beta}_{0t}^S, \hat{\beta}_{1t}^S, \hat{\beta}_{2t}^S, \hat{\beta}_{3t}^S)'$. Der Mittelwert ist nun

$\mu^f = (\mu_1^f, \dots, \mu_r^f)'$, $\mu^{DL} = (\mu_0^{DL}, \mu_1^{DL}, \mu_2^{DL})'$ und $\mu^S = (\mu_0^S, \mu_1^S, \mu_2^S, \mu_3^S)'$. Die Regressionsmatrix Θ ergibt sich zuletzt jeweils zu

$$\Theta^f = \begin{pmatrix} \theta_{11}^f & \dots & \theta_{1r}^f \\ \vdots & \ddots & \vdots \\ \theta_{r1}^f & \dots & \theta_{rr}^f \end{pmatrix}, \Theta^{DL} = \begin{pmatrix} \theta_{11}^{DL} & \theta_{12}^{DL} & \theta_{13}^{DL} \\ \theta_{21}^{DL} & \theta_{22}^{DL} & \theta_{23}^{DL} \\ \theta_{31}^{DL} & \theta_{32}^{DL} & \theta_{33}^{DL} \end{pmatrix} \text{ und } \Theta^S = \begin{pmatrix} \theta_{11}^S & \dots & \theta_{14}^S \\ \vdots & \ddots & \vdots \\ \theta_{41}^S & \dots & \theta_{44}^S \end{pmatrix}$$

wobei η anschließend den Dimensionen entsprechend angepasst wird. Wenn wir hier wieder die Annahmen eines einfachen beziehungsweise eines multiplen Regressionsmodells stellen und darüber hinaus eine stationäre Verteilung für die Faktoren annehmen, können die Parameter θ_i^k und Θ^k ($k = f, DL, S$) über eine Regression geschätzt werden.

Schauen wir uns dann die Gleichung (18) genauer an, erkennen wir, dass sich das VAR(1) Modell lediglich von dem AR(1) Modell dahingehend unterscheidet, dass für die Prognose des Faktors i , die Effekte der anderen Faktoren j , mit $j \neq i$, berücksichtigt werden.

Die Zinsen für den Zeitpunkt $t + 1$ für die Laufzeit m ergeben sich dann im 2. Schritt aus:

2. Schritt:

$$\hat{y}_{m,t+1} = \hat{\mu}_m + \hat{\lambda}_{m1} \hat{f}_{1,t+1}, \dots, \hat{\lambda}_{mr} \hat{f}_{r,t+1} + \varepsilon_{t+1}^f, \quad \hat{\mu}_m = \frac{1}{t} \sum_{d=1}^t y_{md}$$

$$\hat{y}_{t+1}(m) = \hat{\beta}_{0,t+1}^{DL} + \hat{\beta}_{1,t+1}^{DL} \frac{1 - \exp(-\lambda m)}{\lambda m} + \hat{\beta}_{2,t+1}^{DL} \left(\frac{1 - \exp(-\lambda m)}{\lambda m} - \exp(-\lambda m) \right) + \varepsilon_{t+1}^{DL} \quad (19)$$

$$\begin{aligned} \hat{i}_{t+1,m}(b) &= \hat{\beta}_{0,t+1}^S + \hat{\beta}_{1,t+1}^S \frac{1 - \exp\left(-\frac{m}{\tau_1}\right)}{\frac{m}{\tau_1}} + \hat{\beta}_{2,t+1}^S \left(\frac{1 - \exp\left(-\frac{m}{\tau_1}\right)}{\frac{m}{\tau_1}} - \exp\left(-\frac{m}{\tau_1}\right) \right) \\ &+ \hat{\beta}_{3,t+1}^S \left(\frac{1 - \exp\left(-\frac{m}{\tau_2}\right)}{\frac{m}{\tau_2}} - \exp\left(-\frac{m}{\tau_2}\right) \right) + \varepsilon_{t+1}^S \end{aligned}$$

4.2 Hinzunahme weiterer Variablen für die dynamischen Vorhersage

Im Gegensatz zu [Diebold und Li \(2006\)](#), wollen wir nun unsere Schätzung durch die Hinzunahme weiterer ökonomisch relevanter Variablen verbessern. Dabei werden wir bisher unbeobachtete Variablen, die einen Zusammenhang zur Zinsstrukturkurve haben berücksichtigen, um ausschließlich die jeweiligen Effekte³ der Faktoren auf die Zinsstruktur zu messen. Dadurch sollten die zu interessierenden Faktoren um *noise*, also unerwünschte Effekte, bereinigt werden und somit der Effekt der Faktoren auf die Zinsstrukturkurve deutlicher.

Wir erweitern die bisherigen dynamischen Gleichungen wie [Diebold, Rudebusch und Aruoba \(2006\)](#) ihre Modellgleichung erweitert haben⁴. Dabei werden wir drei makroökonomische Variablen hinzufügen, die allgemein als fundamentaler Bestandteil einer makroökonomischen Analyse verstanden werden. CU_t ist die Produktionsauslastung zum Zeitpunkt t , welche das Niveau der realen ökonomischen Aktivität relativ zum Potential der Volkswirtschaft widerspiegelt.

³Im Fall von Diebold und Li: Niveau, Steigung und Krümmung.

⁴Jedoch verzichten wir auf eine Darstellung der Dynamik durch eine Impulse Response Function und einer Schätzung mittels des Kalman-Filters.

FFR_t ist der Leitzins zum Zeitpunkt t , der durch die jeweilige Zentralbank der Volkswirtschaft gegeben wurde und $INFL_t$ ist die jeweilige Inflation der Volkswirtschaft zum Zeitpunkt t . Die dynamischen Gleichungen werden nun wie folgt angepasst:

1. Schritt: $AR(1)$ mit Erweiterung

Faktorenanalyse: $(\tilde{F}_{i,t+1} - M_i^f) = \theta_i^f (\tilde{F}_{it} - M_i^f) + \eta_{i,t+1}^f$ für $i = 0, \dots, r$

Diebold und Li: $(\tilde{B}_{i,t+1}^{DL} - M_i^{DL}) = \theta_i^{DL} (\tilde{B}_{it}^{DL} - M_i^{DL}) + \eta_{i,t+1}^{DL}$ für $i = 0, 1, 2$ (20)

Svensson: $(\tilde{B}_{i,t+1}^S - M_i^S) = \theta_i^S (\tilde{B}_{it}^S - M_i^S) + \eta_{i,t+1}^S$ für $i = 0, 1, 2, 3$

1. Schritt: $VAR(1)$ mit Erweiterung

Faktorenanalyse: $(\tilde{f}_{t+1} - \mu^f) = \Theta^f (\tilde{f}_t - \mu^f) + \eta_{t+1}^f$

Diebold und Li: $(\tilde{\beta}_{t+1}^{DL} - \mu^{DL}) = \Theta^{DL} (\tilde{\beta}_t^{DL} - \mu^{DL}) + \eta_{t+1}^{DL}$ (21)

Svensson: $(\tilde{\beta}_{t+1}^S - \mu^S) = \Theta^S (\tilde{\beta}_t^S - \mu^S) + \eta_{t+1}^S$

Wobei im $AR(1)$ Modell die unabhängigen Variablen nun durch die Folgenden Vektoren definiert werden: $\tilde{F}_{it} = (\tilde{f}_{it}, CU_t, FFR_t, INFL_t)'$, $\tilde{B}_{it}^{DL} = (\tilde{\beta}_{it}^{DL}, CU_t, FFR_t, INFL_t)'$ und $\tilde{B}_{it}^S = (\tilde{\beta}_{it}^S, CU_t, FFR_t, INFL_t)'$. Die jeweiligen zu schätzenden Regressoren im Erweiterten Fall sind dann $\theta_i^k = (\theta_{i1}^k, \theta_{i2}^k, \theta_{i3}^k, \theta_{i4}^k)$ und die Mittelwerte ergeben sich zu $M_i^k = (\mu_i^k, \mu^{CU}, \mu^{FFR}, \mu^{INFL})'$ für $k = f, DL, S$.

Für das $VAR(1)$ Modell werden die unabhängigen Variablen ähnlich wie für das $AR(1)$ Modell angepasst: $\hat{f}_t = (\hat{f}_{1t}, \dots, \hat{f}_{rt}, CU_t, FFR_t, INFL_t)'$, $\hat{\beta}_t^{DL} = (\hat{\beta}_{0t}^{DL}, \hat{\beta}_{1t}^{DL}, \hat{\beta}_{2t}^{DL}, CU_t, FFR_t, INFL_t)'$ und $\hat{\beta}_t^S = (\hat{\beta}_{0t}^S, \hat{\beta}_{1t}^S, \hat{\beta}_{2t}^S, \hat{\beta}_{3t}^S, CU_t, FFR_t, INFL_t)'$. Die Mittelwerte sind nun $\mu^f = (\mu_1^f, \dots, \mu_r^f, \mu^{CU}, \mu^{FFR}, \mu^{INFL})'$, $\mu^{DL} = (\mu_0^{DL}, \mu_1^{DL}, \mu_2^{DL}, \mu^{CU}, \mu^{FFR}, \mu^{INFL})'$ und $\mu^S = (\mu_0^S, \mu_1^S, \mu_2^S, \mu_3^S, \mu^{CU}, \mu^{FFR}, \mu^{INFL})'$. Die Regressoren ändern sich zu:

$$\Theta^f = \begin{pmatrix} \theta_{11}^f & \cdots & \theta_{1,r+3}^f \\ \vdots & \ddots & \vdots \\ \theta_{r+3,1}^f & \cdots & \theta_{r+3,r+3}^f \end{pmatrix}, \Theta^{DL} = \begin{pmatrix} \theta_{11}^{DL} & \cdots & \theta_{16}^{DL} \\ \vdots & \ddots & \vdots \\ \theta_{61}^{DL} & \cdots & \theta_{66}^{DL} \end{pmatrix} \text{ und } \Theta^S = \begin{pmatrix} \theta_{11}^S & \cdots & \theta_{17}^S \\ \vdots & \ddots & \vdots \\ \theta_{71}^S & \cdots & \theta_{77}^S \end{pmatrix}$$

Wie auch in dem Fall ohne makroökonomische Variablen, wird η den Dimensionen entsprechend angepasst. Der Unterschied zu den Regressionsmodellen ohne Erweiterungen ist, dass wir in der Berechnung von θ_i und Θ und somit auch für die Schätzung der Faktoren der Periode $t + 1$, den Einfluss der makroökonomischen Variablen berücksichtigen.

Die jeweiligen Zinsen der Laufzeit m für den Zeitpunkt $t + 1$ ergeben sich dann fast analog zu Gleichung (19), jedoch mit geänderten Faktoren:

2. Schritt:

$$\begin{aligned}\tilde{y}_{m,t+1} &= \tilde{\mu}_m + \tilde{\lambda}_{m1}\tilde{f}_{1,t+1}, \dots, \tilde{\lambda}_{mr}\tilde{f}_{r,t+1} + \varepsilon_{t+1}^f, \quad \tilde{\mu}_m = \frac{1}{t} \sum_{d=1}^t y_{md} \\ \tilde{y}_{t+1}(m) &= \tilde{\beta}_{0,t+1}^{DL} + \tilde{\beta}_{1,t+1}^{DL} \frac{1-\exp(-\lambda m)}{\lambda m} + \tilde{\beta}_{2,t+1}^{DL} \left(\frac{1-\exp(-\lambda m)}{\lambda m} - \exp(-\lambda m) \right) + \varepsilon_{t+1}^{DL} \\ \tilde{i}_{t+1,m}(b) &= \tilde{\beta}_{0,t+1}^S + \tilde{\beta}_{1,t+1}^S \frac{1-\exp\left(-\frac{m}{\tau_1}\right)}{\frac{m}{\tau_1}} + \tilde{\beta}_{2,t+1}^S \left(\frac{1-\exp\left(-\frac{m}{\tau_1}\right)}{\frac{m}{\tau_1}} - \exp\left(-\frac{m}{\tau_1}\right) \right) \\ &\quad + \tilde{\beta}_{3,t+1}^S \left(\frac{1-\exp\left(-\frac{m}{\tau_2}\right)}{\frac{m}{\tau_2}} - \exp\left(-\frac{m}{\tau_2}\right) \right) + \varepsilon_{t+1}^S\end{aligned}\tag{22}$$

In Gleichung (22) sehen wir dann, dass bei der direkten Berechnung der Zinsen, die makroökonomischen Variablen CU_t , FFR_t und $INFL_t$ nicht vorkommen. Diese finden nur bei der Berechnung der Faktoren Berücksichtigung. Somit unterscheidet sich Gleichung (22) zu Gleichung (19) dadurch, dass die in Gleichung (22) verwendeten Faktoren vorher um makroökonomische Einflüsse bereinigt wurden und sich dadurch die Schätzung der Zinsen verbessert haben sollte. Möchte man darüber hinaus mehr als eine Periode in die Zukunft schätzen, so lässt sich der vorher beschriebene Algorithmus auch iterativ erweitern. So kann eine h -monatige Vorhersage entstehen, indem zuerst der Faktor für die Periode $t+1$ geschätzt wird, um aus diesem die Schätzungen für die Perioden $t+2$ anzufertigen. Anschließend kann aus dem Faktor der Periode $t+2$ eine Prognose des Faktors für die Periode $t+3$ berechnet werden. Dies wird nun so oft wiederholt, bis der Faktor und daraus folgend der Zins der Periode $t+h$ ermittelt wurde.

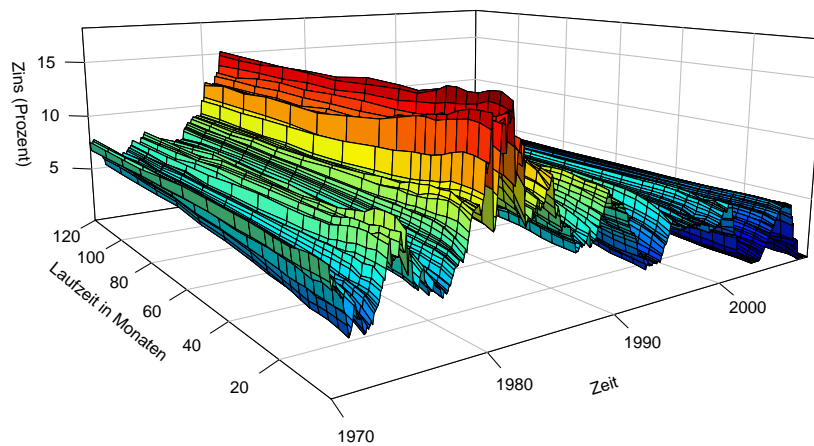
5 Anwendung

In diesem Abschnitt wollen wir uns mit der praktischen Umsetzung der vorher beschriebenen Theorie widmen. Wir werden dabei analysieren, wie die drei theoretischen Modelle Zinsstrukturkurven innerhalb der beobachteten Daten (*in-sample*) und außerhalb der beobachteten Daten (*out-of-sample*) abbilden können. Aber zunächst müssen wir schauen, welche Daten für die Umsetzung herangezogen werden. An dieser Stelle sei darauf verwiesen, dass sämtliche Programmierarbeit im Zusammenhang mit dieser Bachelorarbeit, auf meiner GitHub Seite einzusehen ist: <https://github.com/justinfranken/DMdZvS>.

5.1 Verwendete Daten

In der folgenden Anwendung werden wir zwei verschiedene Datensätze verwenden. Die Zinsdaten entstanden aus der Arbeit von [Jungbacker et al. \(2014\)](#), welcher monatliche U.S. amerikanische Staatsanleihen, sogenannte US Treasury Yields, zu Nullkuponpreisen betrachtete. Schauen wir zunächst auf Abbildung 5.1. In der Zeit vor 1985 und nach 2007 sind aufgrund von Öl und Finanzkrisen in Amerika, starke Bewegungen auf dem Anleihemarkt zu erkennen.

Abbildung 5.1: US-amerikanische Staatsanleihen Zinsen

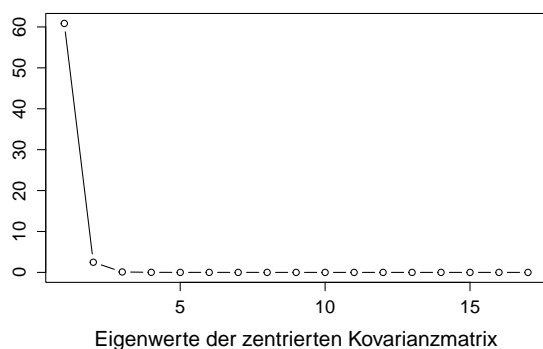


Anmerkung: Monatliche Zinssätze US-amerikanischer Staatsanleihen zu Nullkuponpreisen von 1970 bis 2009 mit 17 beobachteten Laufzeiten von drei bis 120 Monaten über 480 beobachtete Perioden. Daten wurden aus der Arbeit von [Jungbacker et al. \(2014\)](#) entnommen.

Um unserer Stationaritäts-Annahme gerecht zu werden, werden wir in der Anwendung den Zeitraum von 1985 bis 2007 betrachten, welcher $T = 265$ beobachtete Perioden enthält, mit $N = 17$ beobachteten Laufzeiten der folgenden Länge in Monaten: 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108 und 120. Der zweite Datensatz enthält die für die Vorhersage notwendigen makroökonomischen Variablen CU , FRR und $INFL$. Diese habe ich aus der FRED Datenbank der Federal Reserve Bank of St. Louis (<https://research.stlouisfed.org/econ/mccracken/fred-databases/>) entnommen. Da die Inflation $INFL$ in der Datenbank zunächst als Verbraucherpreisindex angegeben ist, habe ich $INFL$ durch den Quotienten $(VPI_{\text{aktuell}}/VPI_{\text{Vorjahr}} * 100) - 100$ berechnet. Auch CU , FRR und $INFL$ betrachten wir monatlich in dem Zeitraum von 1985 bis 2007, wodurch ebenfalls $T = 265$ beobachtete Perioden vorhanden sind.

5.1.1 Anzahl an verwendeten Faktoren in der Faktorenanalyse

Abbildung 5.1.1: Screeplot



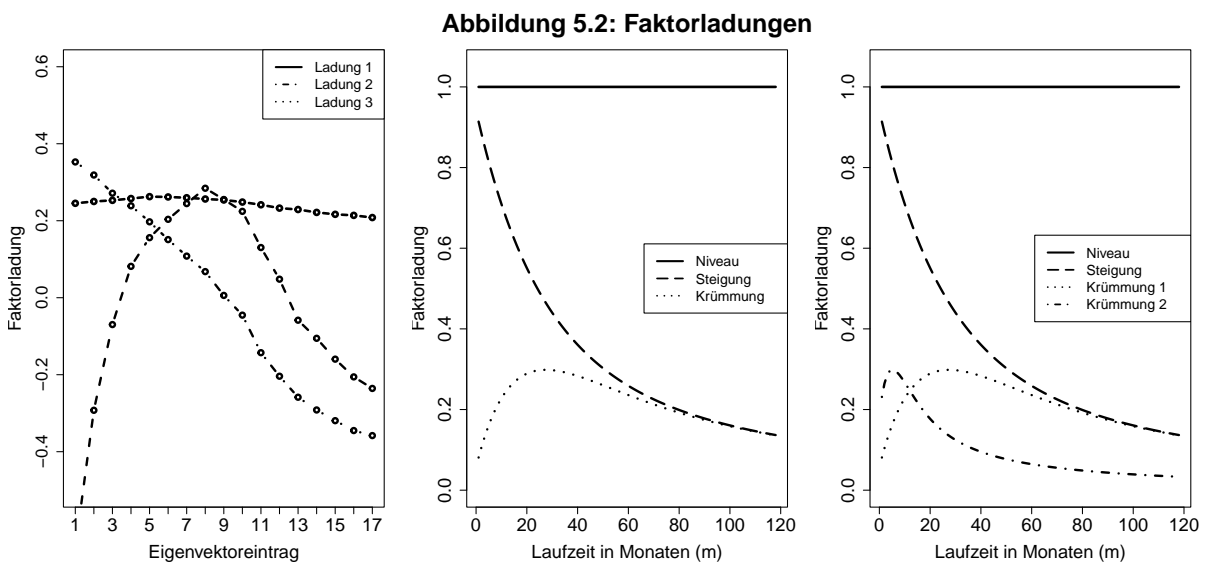
In dem Abschnitt 2.3 haben wir uns der Frage gewidmet, wie viele Faktoren r in der Faktorenanalyse genutzt werden sollten. Im folgenden bedienen wir uns insbesondere an den berechneten Eigenwerten der Kovarianzmatrix der zentrierten beobachteten Zinsdaten. Wie in [Abbildung 5.1.1](#) zu sehen ist, erklärt in diesem Screeplot der erste Eigenwert unglaublich viel Varianz aller Faktoren. Wir können nämlich

erkennen, dass durch die ersten drei Eigenwerte bereits 99,9% Varianz der Faktoren erklärt wird,

wobei die ersten beiden Faktoren 99,8% Varianz erklären und der erste Faktor 95,59% der Varianz erklärt. Normalerweise sollten wir nun $r = 1$ wählen, da wir im Screeplot nach einem Knick in der Grafik suchen und auch die erklärte Varianz mit 95,59% ausreichend hoch ist. Dennoch wollen wir im folgenden $r = 3$ verwenden, da wir immer noch die Dimensionen von 17 auf drei reduzieren konnten, wodurch wir ähnlich wie [Bliss \(1995\)](#) die Zinsstrukturkurve erklären können und anschließend in der Lage sind, resultierenden Faktorladungen und Faktoren mit den Nelson und Siegel Modellen zu vergleichen.

5.2 Modellierung von US-Staatsanleihen

Um nun Zinsstrukturkurven mittels der jeweiligen Verfahren zu Modellieren, müssen wir zunächst klären, welche Parameter der Svensson Methode wir in der Analyse verwenden werden. Betrachten wir dazu die Abbildung 5.2. Dort werden die jeweiligen Faktorladungen der Model-



Anmerkung: Faktorladungen der drei Modelle im Vergleich. Faktorladungen der Faktorenanalyse (Links) mit $r = 3$ Faktoren, Nelson-Siegel Ladungen nach Diebold und Li (Mitte) mit $\lambda = 0.0609$ und Nelson-Siegel Ladungen nach Svensson (Rechts) mit $\tau_1 = 1/0.0609$ und $\tau_2 = 4$.

le miteinander verglichen. Für das Svensson Modell sind die Parameter $\tau_1 \approx 16.42$ und $\tau_2 = 4$ fixiert, um die Komplexität, wie auch im Diebold und Li Modell, zu reduzieren. Dadurch, dass $\tau_1 = 1/\lambda$ ist, unterscheiden sich im Svensson Modell die ersten drei Ladungen nicht von dem Diebold und Li Modell. τ_2 wurde auf 4 fixiert, da so die vierte Ladung als zweiter Krümmungsfaktor mehr Komplexität für kurzfristige Zinsen in der Struktur berücksichtigen kann. Des Weiteren können wir in Abbildung 5.2 erkennen, dass die Eigenschaften der Ladungen aus der Faktorenanalyse, ähnliche Eigenschaften haben, wie die Ladungen im Diebold und Li Modell. Dadurch können wir die Faktoren der Faktorenanalyse, wie im Diebold und Li Modell als Niveau, Steigung und Krümmung interpretieren. Im [Anhang](#) ist für den interessierten Leser eine Abbildung der berechneten Faktoren aus den Modellen aufbereitet worden. Interessant an dieser Abbildung ist, dass die jeweiligen Faktoren, ähnliche Verläufe, trotz verschiedener Modelle aufweisen, was insbesondere durch die Interpretation der Faktoren erklärt werden kann. Für den

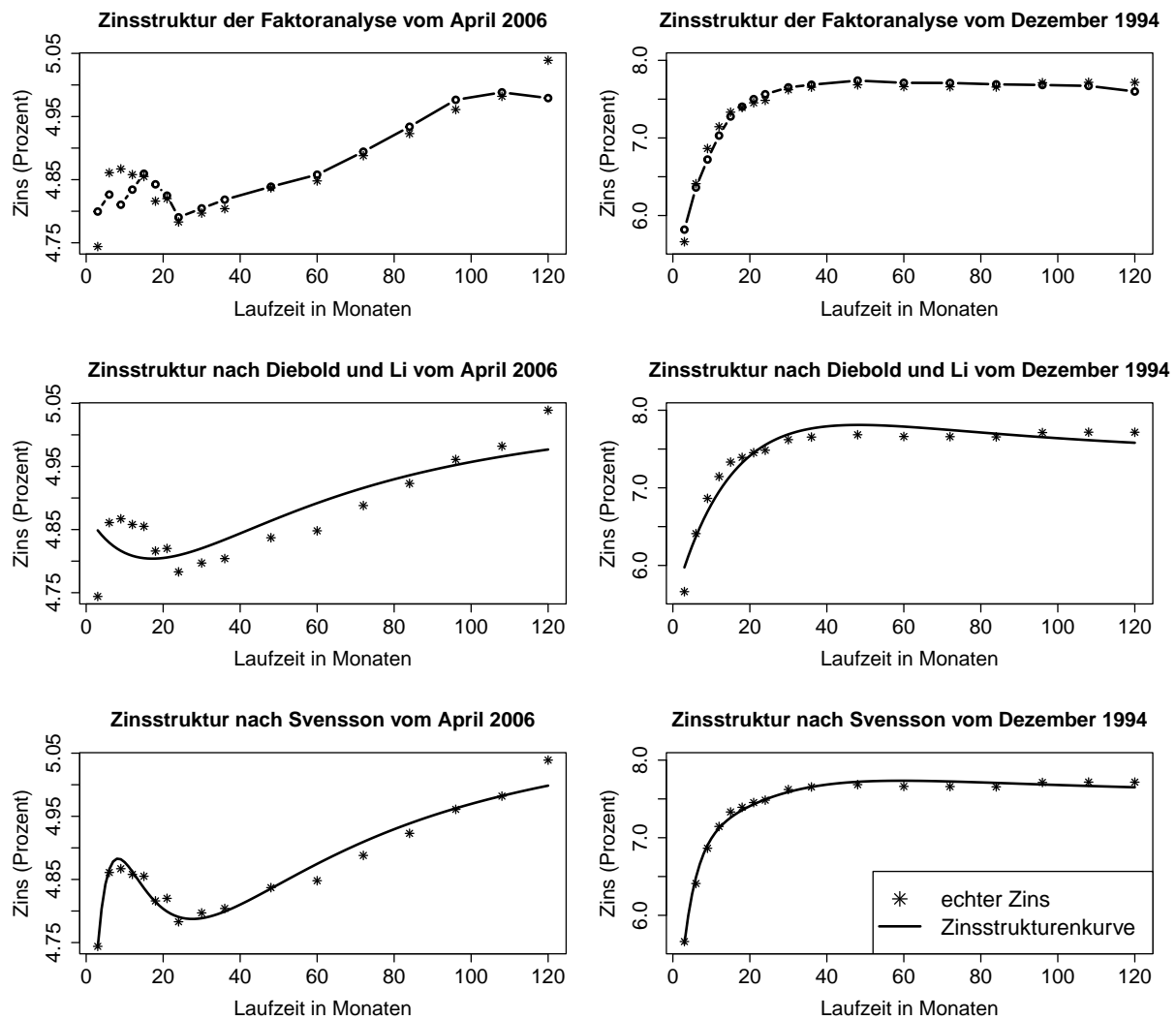


Abbildung 5.2.1: Vergleich der Zinsstrukturkurvenmodellierung mittels Faktorenanalyse, Diebold und Li und Svensson Verfahren, mit Ladungen aus Abb. 5.2. Links ist eine komplexere, Rechts eine monotone Zinsstruktur abgebildet. Sterne repräsentieren beobachtete US-Nullkupon-Staatsanleihen.

Vergleich der *in-sample* Modellierung einer Zinsstrukturkurven, werden wir die drei Verfahren jeweils dazu nutzen, um eine *monotonere* und eine *komplexere* Zinsstrukturkurve zu modellieren. Anschließend werden wir diese gemeinsam mit den echten beobachteten US-Nullkupon-Staatsanleihen grafisch darstellen, um die Genauigkeit der Modellierung zu vergleichen. Wir können dann in Abbildung 5.2.1 erkennen, dass die Zinsstrukturkurve der Modellierung mittels einer Faktorenanalyse, monotone als auch komplexe Zinsstrukturen relativ gut abbilden kann, jedoch Probleme bekommt, bei komplexeren Strukturen in der kurzen Frist. Des Weiteren sind im Gegensatz zu den anderen beiden Modellen, lediglich die 17 beobachteten Zinspunkte definiert. Die Diebold und Li Zinsstrukturkurve kann auch monotone Strukturen gut abbilden, hat jedoch Probleme bei komplexeren Strukturen, denn das Modell konnte aufgrund zusätzlicher Krümmungen, die Struktur schlecht erklären, da dies in den Ladungen nicht berücksichtigt wird. Das Svensson Modell ist in den ausgewählten Beispielen, der klare Sieger. Das Verfahren kann aufgrund der Erweiterung der Nelson-Siegel Kurve um eine zweite Krümmung, komplexe Zinsstrukturkurven besonders gut abbilden, trifft aber auch monotone Zinsstrukturen sehr

gut. Dies erklärt auch, warum viele Zentralbanken das Svensson Modell zur Modellierung der Zinsstrukturkurven nutzen.

5.3 Vorhersage von US-Staatsanleihen

Für die Analyse der *out-of-sample* Beobachtungen, werden wir die Zinsstruktur wie in Abschnitt 4 vorhersagen. Die Vorhersage erfolgt dabei mit und ohne makroökonomische Variablen, als auch durch einen Autoregressiven und einen Vektorautoregressiven Prozess erster Ordnung. Die Zinsstruktur wird zudem um eine Periode (1 Monat), sechs Perioden (6 Monate) und 12 Perioden (1 Jahr) vorhergesagt. Um zu überprüfen wie gut die Vorhersagen sind, teilen wir den Datensatz auf. Für den Fall einer 1 Monats Vorhersage, werden die ersten 100 Perioden dazu genutzt, um die Periode 101 vorherzusagen. Anschließend werden wir mit 101 Perioden, die Periode 102 vorhersagen. Dies wiederholen wir dann solange, bis die 265-te Periode vorhergesagt wurde. Analog werden wir 6 Monate und 12 Monate vorhersagen, wobei dann die erste vorhergesagte Periode die 106 beziehungsweise die 112 sein wird. Für jede vorhergesagte Periode berechnen wir dann die Fehler zwischen Vorhersage und echtem Zins. Der Fehler wird anschließend quadriert und aus allen quadrierten Fehlern wird dann der Mittelwert (MSE) berechnet, von dem wir noch die Wurzel ziehen um den *Root Mean Square Error* (RMSE) zu erhalten. Der RMSE kann dann als mittlere Abweichung, in Prozent, zwischen vorhergesagtem Zins und echtem Zins für die vorhergesagten Perioden zwischen April 1993 und Januar 2007 verstanden werden. In Tabelle 1 können wir dann sehen, dass die Faktorenanalyse bessere Vorhersagen treffen kann, wenn bei dieser makroökonomische Variablen berücksichtigt werden

Tabelle 1: RMSE der Vorhersagefehler

	AR(1)		VAR(1)	
	ohne Makrovar.	mit Makrovar.	ohne Makrovar.	mit Makrovar.
Faktorenanalyse				
1 m	0.2416	0.2381	0.2481	0.2940
6 m	0.7563	0.7035	0.7882	1.0199
12 m	1.1882	1.0482	1.2240	1.5370
Diebold und Li				
1 m	0.2538	0.2680	0.2535	0.9653
6 m	0.8251	0.8988	0.7997	3.2251
12 m	1.2650	1.3415	1.3035	3.9277
Svensson				
1 m	0.2594	0.2764	0.3169	1.2771
6 m	0.8506	0.9237	1.0544	3.5257
12 m	1.2799	1.3529	1.4989	4.1446

und die Vorhersage durch ein AR(1) Prozess erfolgt. Dieser Effekt tritt jedoch nicht bei den anderen Modellen auf, als auch nicht bei einer Vorhersage durch einen VAR(1) Prozess. In dem Fall von Diebold-Li und Svensson, wird die Vorhersage durch einen VAR(1) Prozess sogar besonders schlecht, wenn Makrovariablen berücksichtigt werden. Weiterhin erkennen wir, dass außer für die Diebold und Li 1 Monats-Vorhersage, der AR(1) Prozess besser funktioniert, als

der VAR(1) Prozess und wie zu erwarten war, sind die Vorhersagen besser, je weniger Perioden in die Zukunft geschätzt wird. Darüber hinaus wird deutlich, dass die Faktorenanalyse bessere Vorhersagen treffen kann als das Modell von Diebold und Li und das Modell von Diebold und Li, kann bessere Vorhersagen treffen als das Modell von Svensson. Weiterführende Informationen bezüglich deskriptiver Statistiken der verwendeten Datensätze als auch ausführliche Ergebnisse der Vorhersageberechnungen, sind im [Anhang](#) aufbereitet worden.

6 Schlussbemerkung

Diese Bachelorarbeit hat aufgezeigt, wie Zinsstrukturkurven durch eine Faktorenanalyse oder durch das Nelson- und Siegel-Modell beschrieben werden können. Wir haben uns dann Abwandlungen der ursprünglichen Nelson- und Siegel-Modellgleichung nach Diebold Li und Svensson angeschaut und darüber hinaus festgestellt, dass die Faktoren der Modelle als Niveau, Steigung und Krümmung (oder zusätzliche Krümmung) verstanden werden können. Für out-of-sample Beobachtungen widmeten wir uns anschließend einer dynamischen Beschreibung der Modellgleichungen. In der beschriebenen Dynamik haben wir zusätzlich fundamental wichtige makroökonomische Variablen berücksichtigt, um die Vorhersagen zu verbessern. Durch die Anwendung haben wir zum Schluss aufklären können, dass das Svensson Modell in-sample Beobachtungen sehr gut beschreiben kann, aber die beiden anderen Modelle kleinere Probleme bei komplexeren Zinsstrukturen aufweisen. Bei der out-of-sample Vorhersage von Zinsstrukturkurven war es zwar mit allen Modellen möglich, gute Prognosen zu treffen, jedoch erzielte die Faktorenanalyse die besten Ergebnisse. Die Vorhersagen konnten durch den AR(1) Prozess am besten geschätzt werden und außer für die Prognose von Faktoren der Faktorenanalyse durch ein AR(1) Prozess, ließ sich auch keine Verbesserung des RMSE durch Makrovariablen erkennen.

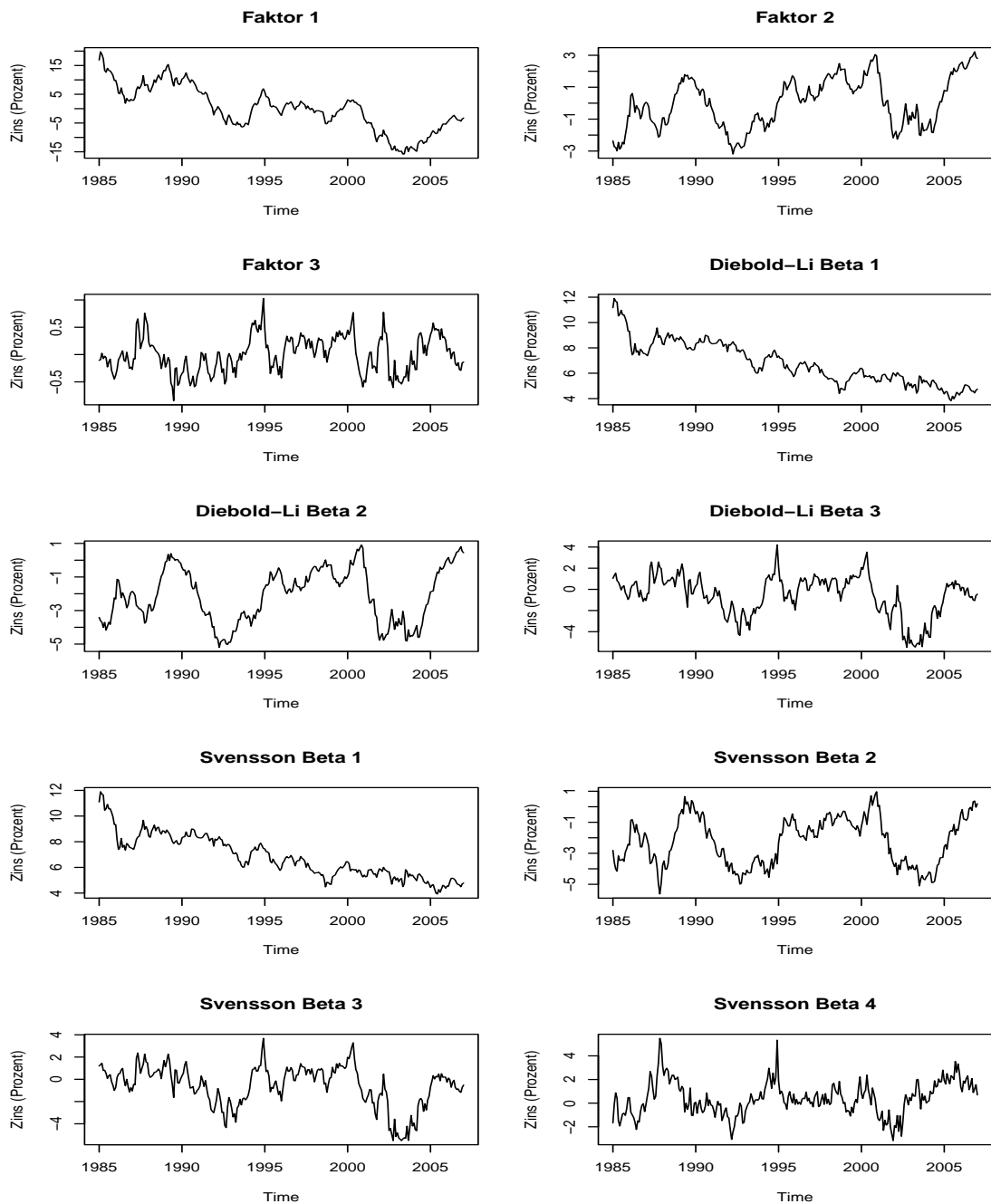
Über die Bachelorarbeit hinausgehend wäre eine fortführende Analyse der Vorhersagekraft der dynamischen Modelle für AR- und VAR-Prozesse höherer Ordnungen von Interesse, sowie die Lockerung einer Stationaritäts-Annahme in der Prognose. Zudem gibt es noch weitere aufschlussreiche Forschungsarbeiten zur Zinsstrukturanalyse. Beispielsweise haben [Diebold, Rudebusch und Aruoba \(2006\)](#) untersucht, wie die makroökonomischen Variablen im Zusammenhang zu den Modellfaktoren Niveau, Steigung und Krümmung stehen. [Mönch \(2008\)](#) ging noch einen Schritt weiter und hat die Schätzung von Zinsstrukturen verbessert, indem er dafür ausschließlich makroökonomische Variablen betrachtet, ohne vorher induzierte Interpretationen der Faktoren in seinen VAR-Modellen vorzunehmen. Eine noch tiefgreifendere Analyse der Zinsstrukturkurve haben [Otto und Salish \(2022\)](#) geleistet. Otto und Salish beschreiben die Zinsstrukturkurve, indem die beobachteten Zinsen zunächst durch Splines stetig definiert werden, um anschließend durch eine funktionale Faktorenanalyse die Struktur zu erforschen und eine Dynamik zu entwickeln. Darüber hinaus formulieren Otto und Salish Informationskriterien zur Identifizierung der optimalen Anzahl an berücksichtigten Faktoren, sowie der Größe der Ordnungen in den AR- und VAR-Prozessen.

7 Literaturverzeichnis

- Ahn, S. C. und Horenstein, A. R. (2013). Eigenvalue ratio test for the number of factors. *Econometrica*, 81: 1203 - 1227.
- Bai, J. und Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70: 191 - 221.
- Bliss, R. R. (1997). Movements in the term structure of interest rates. *Federal Reserve Bank of Atlanta*, S. 16 - 33.
- Breitung, J. und Choi, I. (2013). Factor models. In *Handbook of Research Methods and Applications in Empirical Macroeconomics*, S. 249 - 265. Edward Elgar Publishing.
- Diebold, F. X. und Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics*, 130: 337 - 364.
- Diebold, F. X., Rudebusch, G. D. und Aruoba, S. B. (2006). The macroeconomy and the yield curve: A dynamic latent factor approach. *Journal of Econometrics*, 131: 309 - 338.
- Jungbacker, B., Koopman, S. J. und Van der Wel, M. (2014). Smooth dynamic factor analysis with application to the US term structure of interest rates. *Journal of Applied Econometrics*, 29: 65 - 90.
- McCulloch, J. H. (1971). Measuring the term structure of interest rates. *The Journal of Business*, 44: 19 - 31.
- McCulloch, J. H. (1975). The tax-adjusted yield curve. *The Journal of Finance*, 30: 811 - 830.
- Mönch, E. (2008). Forecasting the yield curve in a data-rich environment: A no-arbitrage factor-augmented VAR approach. *Journal of Econometrics*, 146: 26 - 43.
- Nelson, C. R. und Siegel, A. F. (1987). Parsimonious modeling of yield curves. *The Journal of Business*, 60: 473 - 489.
- Otto, S und Salish, N (2022), Dynamic factor model for functional time series: Identification, estimation, and prediction. Working Paper.
- Stock, J. H. und Watson M. W. (2015). *Introduction to econometrics*. Pearson.
- Svensson, L. E. O. (1995). Estimating forward interest rates with the extended Nelson & Siegel method. *Quarterly review / Central Bank of Sweden*, S. 13 - 26.

8 Anhang

Abbildung 8.1: Faktoren der Modelle von 1985 bis 2007



Anmerkung: Faktoren der jeweiligen Modelle von 1985 bis 2007, mit den Faktorladungen aus Abb. 5.2 und berechnet aus den Zinsdaten US-amerikanischer Nullkuponanleihen, welche aus der Arbeit von [Jungbacker et al. \(2014\)](#) entnommen wurden. Man beachte, dass für die Berechnung der Faktoren 1,2 und 3, die Zinsdaten vor der Analyse um den Mittelwert reduziert wurden.

Tabelle 2: Deskriptive Statistik der Zins- und Makrodaten

	Mittelwert	Standardabweichung	Min	1st Qu.	Median	3rd Qu.	Max
3	4.799	2.005	0.876	3.390	5.076	5.839	9.130
6	4.941	2.017	0.958	3.534	5.174	6.192	9.324

9	5.052	2.020	1.012	3.537	5.275	6.269	9.343
12	5.172	2.046	1.039	3.741	5.374	6.457	9.559
15	5.308	2.073	1.065	3.878	5.516	6.635	9.954
18	5.390	2.057	1.143	4.041	5.570	6.756	10.145
21	5.460	2.034	1.217	4.155	5.579	6.834	10.230
24	5.504	2.006	1.297	4.184	5.599	6.831	10.371
30	5.640	1.983	1.435	4.340	5.721	7.036	10.661
36	5.742	1.942	1.599	4.435	5.772	7.127	10.715
48	5.941	1.898	1.955	4.633	5.907	7.379	11.126
60	6.056	1.845	2.294	4.682	5.978	7.478	11.196
72	6.188	1.834	2.562	4.727	6.070	7.654	11.450
84	6.271	1.79	2.856	4.815	6.154	7.656	11.571
96	6.352	1.764	3.074	4.910	6.185	7.712	11.425
108	6.406	1.757	3.237	4.954	6.238	7.763	11.562
120	6.404	1.723	3.347	5.021	6.236	7.716	11.531
CU	79.72	3.184	71.63	78.25	80.12	82.28	85.62
FFR	5.103	2.22	0.980	3.300	5.310	6.530	9.850
INFL	3.058	1.083	1.069	2.264	2.982	3.664	6.380

Anmerkung: Deskriptive Statistiken der Verwendeten Zinsdaten aus der Arbeit von [Jungbacker et al. \(2014\)](#) und den Makroökonomischen Daten aus der FRED Datenbank: CU (Produktionsauslastung der USA), FFR (Leitzins der FED) und INFL (berechnete Inflation des Verbraucherpreisindex der USA).

Tabelle 3: Ausführliche Ergebnisse der Vorhersagefehler

Faktorenanalyse			RMSE	MSE	Max. MSE	Min. MSE	SD MSE
1m	AR(1)	ohne M.	0.2416	0.0584	0.4862	5.6197e-06	0.0845
		mit M.	0.2381	0.0567	0.4298	3.1070e-07	0.0802
	VAR(1)	ohne M.	0.2481	0.0616	0.5336	7.1139e-06	0.0884
		mit M.	0.2940	0.0864	0.7077	5.9755e-06	0.1249
6m	AR(1)	ohne M.	0.7563	0.5720	4.1489	4.6747e-08	0.7489
		mit M.	0.7035	0.4950	3.0611	9.2277e-06	0.6103
	VAR(1)	ohne M.	0.7882	0.6213	4.9746	6.4311e-06	0.8477
		mit M.	1.0199	1.0403	6.5787	4.6205e-06	1.2703
12m	AR(1)	ohne M.	1.1881	1.4118	10.4193	0.0008	1.8247
		mit M.	1.0482	1.0987	10.1990	8.1760e-05	1.5443
	VAR(1)	ohne M.	1.2239	1.4981	10.6419	0.0003	2.0176
		mit M.	1.5370	2.3624	11.2537	0.0006	2.8259
Diebold und Li			RMSE	MSE	Max. MSE	Min. MSE	SD MSE
1m	AR(1)	ohne M.	0.2538	0.0644	0.5513	1.4907e-05	0.0922
		mit M.	0.2680	0.0718	0.5949	3.4587e-06	0.1045

	VAR(1)	ohne M. mit M.	0.2535 0.9653	0.0643 0.9318	0.5663 11.5726	4.2160e-07 2.4241e-05	0.0970 1.3496
6m	AR(1)	ohne M. mit M.	0.8251 0.8988	0.6808 0.8079	4.3612 5.4732	6.6664e-05 1.6110e-05	0.8512 1.0493
		VAR(1)	ohne M. mit M.	0.7997 3.2251	0.6396 10.4014	4.8334 156.0724	3.6966e-06 0.0018
	AR(1)		ohne M. mit M.	1.2650 1.3415	1.6002 1.7995	12.2871 10.1218	9.2405e-07 6.9325e-05
		VAR(1)	ohne M. mit M.	1.3035 3.9277	1.6991 15.4271	10.7326 260.5435	0.0002 3.2032e-05
Svensson			RMSE	MSE	Max. MSE	Min. MSE	SD MSE
1m	AR(1)	ohne M. mit M.	0.2594 0.2764	0.0673 0.0764	0.6135 0.6742	7.7946e-06 2.8459e-05	0.0979 0.1135
		VAR(1)	ohne M. mit M.	0.3169 1.2771	0.1004 1.6310	0.7910 12.6553	2.1119e-05 0.0015
	AR(1)		ohne M. mit M.	0.8506 0.9237	0.7235 0.8532	5.0844 6.0330	0.0003 9.9479e-07
		VAR(1)	ohne M. mit M.	1.0544 3.5257	1.1118 12.4309	6.1250 258.3894	1.1343e-07 0.0038
12m	AR(1)		ohne M. mit M.	1.2799 1.3529	1.6382 1.8304	10.1642 9.5949	3.1683e-05 1.0516e-05
		VAR(1)	ohne M. mit M.	1.4989 4.1446	2.2467 17.1781	11.8660 560.4187	0.0001 6.5696e-05

Anmerkung: Ergebnisse der Vorhersagefehler in ausführlicher Form. **RMSE** oder **Root Mean Squared Error**, ist die Wurzel des **MSE** (Mean Squared Error). Der **Mean Squared Error** ist der Mittelwert aller quadrierten Fehler. **Max. MSE** ist der größte Wert aller quadrierten Fehler, **Min. MSE** ist der kleinste Wert aller quadrierten Fehler. **SD MSE** ist die Standardabweichung der quadrierten Fehler.

9 Schriftliche Versicherung

Ich versichere hiermit, dass ich die vorstehende Bachelorarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, dass die vorgelegte Arbeit noch an keiner anderen Hochschule zur Prüfung vorgelegt wurde und dass sie weder ganz noch in Teilen bereits veröffentlicht wurde. Wörtliche Zitate und Stellen, die anderen Werken dem Sinn nach entnommen sind, habe ich in jedem einzelnen Fall kenntlich gemacht.

Datum und Unterschrift: