



Berner Fachhochschule

Soziale Arbeit

Data Mining mit Administrativdaten der Sozialen Sicherheit

Schlussbericht, 22. November 2013

**Tobias Fritschi, Oliver Hümbelin, Christoph Schaller, Robert Fluder, Bernhard Anrig,
Urs Sauter, Kilian Koch, Livia Bannwart, Luca Bösch**

Projektleitung und Umsetzung:

Tobias Fritschi

Oliver Hümbelin

Christoph Schaller

Robert Fluder

Bernhard Anrig

Urs Sauter

Kilian Koch

Livia Bannwart

Luca Bösch

Kontakt:

Tobias Fritschi, Dozent, lic. rer. pol. (tobias.fritschi@bfh.ch / T: +41 31 848 36 84)

www.soziale-arbeit.bfh.ch/de/forschung/publikationen/soziale_sicherheit

Inhaltsverzeichnis

Inhaltsverzeichnis	3
Zusammenfassung	5
Tabellenverzeichnis	10
Abbildungsverzeichnis	11
1 Einleitung	12
1.1 Ausgangslage	12
1.2 Zielsetzung und Fragestellung	13
2 Verwendete Datensätze	14
3 Vorgehensweise	16
3.1 Arbeitsschritt 1 – Daten aufbereiten	17
3.2 Arbeitsschritt 2 – Analyse von Verlaufsmustern	17
3.3 Arbeitsschritt 3 – Erstellen von Risikoprofilen	18
3.4 Arbeitsschritt 4 – Modell zur Prognose der Risikostruktur	18
3.5 Arbeitsschritt 5 – Indikatoren für die Systeme der Sozialen Sicherheit	19
3.6 Arbeitsschritt 6 – Risikomonitoring SECO	19
3.7 Datenaufbereitung	20
3.7.1 Grundsätzliche Bemerkungen	20
3.7.2 Festlegung von Zuständen	20
3.7.3 Personenmerkmale und Ereignisse	21
4 Verlaufs- und Risikomusteranalyse	23
4.1 Konzeptionelle Eingrenzungen	23
4.1.1 Eingrenzung der Grundgesamtheit	23
4.1.2 Umgang mit der zeitlichen Dimension	24
4.2 Verlaufstypologien	24
4.2.1 Verlaufsindikatoren	24
4.2.2 Methodische Überlegungen zur verwendeten TwoStep-Clusteranalyse	25
4.2.3 Verlaufstypologie für das Gesamtsystem	26
4.2.4 Verlaufstypologie differenziert nach Eintrittssystemen	28
4.3 Risikoprofile	32
4.3.1 Datengestützte Risikoprofile als Ergebnis einer Regressionsanalyse	32
4.3.2 Risikoprofile für Eintritte über ALV	34
4.3.3 Risikoprofile für Eintritte über SH	38
4.3.4 Risikoprofile für Eintritte über IV	43
4.3.5 Risikoprofile für Eintritte ins Gesamtsystem	49
4.4 Modell zur Prognose der Risikostruktur	55
4.4.1 Prognose der Risikostruktur für die Eintrittskohorten 2007 bis 2010	55
4.4.2 Beurteilung der Modellgüte und der Schätzparameter	56
4.5 Indikatoren für die Systeme der Sozialen Sicherheit	58
4.5.1 Verlaufsindikatoren	58
4.5.2 Mobilitätsindikatoren und Distanz zur Erwerbstätigkeit	59
4.6 Weiterführende Analysen	61
5 Entwicklung eines Monitoring-Tools	62
5.1 Datengrundlagen und Variablen	62
5.2 Auswertungsmöglichkeiten	65
5.2.1 Erstellen von Pivot-Tabellen	66
5.2.2 Berichte mit Grafiken aus SQL-Server Reporting Services	68

6	Literaturverzeichnis	70
7	Anhang	71
7.1	Rechnerische Grundlagen des Erwerbsintegrationsindex	71
7.2	Auswertung der Clusterergebnisse anhand der für die Clusteranalyse verwendeten Merkmale	74
7.3	Variablen aus einzelnen Systemen	77
7.4	Ergebnis Parameterschätzung Risikomodell	78
7.5	Anleitung zur Einrichtung des CUBE-Zugriffs	80

Zusammenfassung

Einleitung

Die Schweiz verfügt über ein umfassendes Netz der Sozialen Sicherheit. Im Leistungsvollzug entstehen dabei Administrativdatensätze, die über die Sozialversicherungsnummer zeitlich verknüpfbar sind und weite Teile des Systems der Sozialen Sicherheit abbilden. 2009 hat der Fachbereich Soziale Arbeit im Auftrag des Bundesamtes für Sozialversicherungen erstmals eine integrale Analyse solcher Administrativdaten vorgenommen und die Übergänge zwischen den Leistungssystemen Arbeitslosenentschädigung (ALV), Sozialhilfe (SH), IV-Rente (IV) und IV-Taggelder (IVT) quantifiziert (Fluder et al. 2009). Damit wurde die Grundlage für die Expertise zur Analyse von Sozialleistungsdaten am Fachbereich Soziale Arbeit der Berner Fachhochschule gelegt. Mehrere Projekte im Auftrag von Bundesämtern folgten, zu Verläufen im Bereich der Sozialhilfe (Salzgeber et al. 2010), im Bereich der ALV (Fluder et al. 2012) und im Bereich der IV (Bolliger et al. 2012, Fluder et al. 2013).

Fragestellung

Das interdepartementale Projekt „Data Mining mit Administrativdaten der Sozialen Sicherheit“ hat zum Ziel, ausgehend von Administrativdaten komplexe quantitative Analysemethoden für das Verständnis des Systems der Sozialen Sicherheit fruchtbar zu machen und diese in einem Monitoring-Tool zu konsolidieren. Die Umsetzung des Projektes ist dabei durch eine enge Zusammenarbeit des Fachbereichs Soziale Arbeit sowie des Departements Technik und Informatik gekennzeichnet. Damit können aus den Disziplinen Ökonomie, Soziologie, Statistik und Informatik Kompetenzen in das Projekt miteingebracht werden. Der vorliegende Bericht enthält sowohl die Ergebnisse, die zuhanden des Praxispartners SECO erarbeitet wurden, als auch die Ergebnisse des BFH-Forschungsprojektes.

Grundlage der Analysen bilden Einzelverläufe von Personen, die in den Jahren 2005 bis 2010 in der Schweiz Leistungen der sozialen Sicherheit (ALV, IV und Sozialhilfe) beansprucht haben. Dabei werden die Sozialversicherungsdaten sowie die Daten der Sozialhilfestatistik mit Informationen zur Erwerbstätigkeit aus den individuellen Konten der AHV verknüpft, so dass auch Erwerbsverläufe vor, während und nach dem Bezug von Sozialleistungen rekonstruiert werden können.

Folgende Fragestellungen sollen im Gesamtprojekt beantwortet werden:

- Welche Methoden und welches Vorgehen aus dem Bereich des Data Mining, insbesondere aus den strukturerkennenden Verfahren wie z.B. verschiedene Formen der Clusteranalyse lassen sich erfolgreich zur Analyse eines grossen Administrativdatensatzes anwenden?
- Lassen sich mittels dieser komplexen Methoden interpretierbare Klassen (Verlaufsmuster) identifizieren?
- Lassen sich Faktoren finden, mit welchen sich die Zugehörigkeit zu solchen typischen Klassen prognostizieren lässt?
- Lassen sich diese Methoden im Bereich der Qualitätssicherung der Daten anwenden, um die Qualität der Daten zu evaluieren und gegebenenfalls zu verbessern?
- Lassen sich die als relevant festgestellten Methoden des Data Mining in einem Analyse-Instrument für die Langzeitbeobachtung konsolidieren, welches die Wiederholung der Analysen aus den Blickwinkeln der vier Systeme der Sozialen Sicherheit ermöglicht und welches für jährliche Erweiterungen des Datenmaterials geeignet ist? Kann dieses Instrument auch auf andere Administrativdatensätze angewendet werden?

Vier Grundtypen des Leistungs- und Erwerbsverlaufs

Durch eine Verschränkung von Cluster- und Sequenzanalyse konnten die Vielzahl der beobachteten Verläufe zu vier Clustern gruppiert werden, welche die Grundlage für eine Verlaufstypologie darstellen. In der gesamtsystemischen Betrachtung konnten vier Gruppen unterschieden werden (vgl. Tabelle 1).

- Der grösste Cluster 4 zeichnet sich durch kurze Unterstützungsphase (ALV) und einem *schnellen Eintritt in den Arbeitsmarkt* (39 Prozent) aus.
- Der nächstgrössere Cluster 3 fasst *Verläufe am Rande des Arbeitsmarktes* zusammen, die längere und/oder mehrere Perioden ALV umfassen sowie Phasen ohne Leistungsbezug und ohne Erwerbstätigkeit (38 Prozent).
- Weitere Cluster entstanden um die *Sozialhilfe* (12 Prozent)
- und die *Invalidenrente* (10 Prozent).

Eine Clustertypologie, die eine Differenzierung nach Eintrittssystemen beinhaltet, enthält darüber hinaus eine feinere Einteilung in zwölf Gruppen.

Tabelle 1: Verlaufstypologie - Eintritt über alle Systeme

Cluster	Benennung	N	%
1	Mit IV-Rente	11'110	10.1%
2	Existenzsicherung primär durch Sozialhilfe	13'558	12.3%
3	Am Rand des Arbeitsmarktes (ALV, Erwerb und Phasen ohne Erwerb und Leistungen)	42'320	38.3%
4	Kurze Unterstützungsphase zu Beginn, danach Erwerb	43'520	39.4%
Gesamt		110'508	100%

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010

Bemerkung: durchschnittlicher Silhouettenwert=0.3¹

Risikoprofile von Neueintritten für problematische Leistungsverläufe

Weitere Auswertungen fassten die Analyse von „Risiko“ ins Auge, wobei hier eine konzeptionelle Operationalisierung von Risiko über die Nähe bzw. Entfernung von Verläufen am Arbeitsmarkt erfolgte. Diese Analysen wurden sowohl für Verläufe ausgehend von Eintritten in das Gesamtsystem sowie ausgehend von Eintritten über die Einzelsysteme ALV, SH und IV durchgeführt.

In der Zusammenfassung werden die Auswertungen auf Eintritte über ALV gezeigt. Zentraler Indikator zur Beurteilung des Risikos ist dabei die absolvierte Ausbildung. Die Risikowahrscheinlichkeit verändert sich markant im Vergleich von Personen mit und ohne nachobligatorischen Bildungsabschluss. Weiter zeigten sich

¹ Der Silhouettenwert spricht für eine mittelmässige Qualität des Clusterergebnisses. Das Mass ist ein Durchschnitt aller Datensätze $(B-A) / \max(A,B)$, wobei A der Abstand des Datensatzes zu seinem Clusterzentrum und B der Abstand des Datensatzes zu dem am nächsten liegenden, nicht zugehörigen Clusterzentrum ist. Ein Umrisskoeffizient von 1 würde bedeuten, dass alle Fälle direkt in ihren Clusterzentren liegen. Ein Wert von -1 würde bedeuten, dass alle Fälle in den Clusterzentren anderer Cluster liegen. Ein Wert von 0 bedeutet, dass die Fälle im Durchschnitt gleich weit entfernt von ihrem eigenen Clusterzentrum und dem nächsten benachbarten Cluster liegen.

unterschiedliche Risiken bei einer Differenzierung nach Nationalität, aber auch nach Kanton. Bei Konstellationen mit simultanem Auftreten mehrerer Risikofaktoren kumuliert sich das Risiko entsprechend. So weisen geschiedene Personen ohne Schulabschluss, aus einem EU8-Land, der Altersgruppe 46 – 65 entsprechend eine mehr als dreifach Wahrscheinlichkeit auf, komplexe Verläufe ohne Wiedereingliederung in den Arbeitsmarkt zu erfahren, wenn diese mit durchschnittlichen Bezüger/innen verglichen werden, die über einen ALV-Bezug in den Leistungsbezug aus dem System der Sozialen Sicherheit eintreten.

Tabelle 2 zeigt vorhergesagte Wahrscheinlichkeiten für die drei ermittelten Verlaufstypen von Eintritten über ALV, für ausgewählte Personengruppen. Dabei wird ersichtlich, dass die Kumulation mehrerer Risikofaktoren wie geringer Bildung, Herkunft aus einem EU8-Land, der Verlust des Lebenspartners und fortgeschrittenes Alter zu einer deutlichen Veränderung der vorhergesagten Risikowahrscheinlichkeit führen. Als Gruppe die im Vergleich zum Durchschnitt günstigere Wahrscheinlichkeiten der Zugehörigkeiten zu unproblematischen Verläufen aufweist, können junge Erwachsene (unter 23 Jahre) mit einem Abschluss auf Sekundarstufe II ausgemacht werden. Diese Gruppe hat ein sehr geringes Risiko riskante Verläufe zu erleben und gleichzeitig ein etwas erhöhte Chance für Verläufe mit Aufnahme einer Erwerbstätigkeit nach kurzer Zeit. Wie sich allerdings den Werten in Tabelle 2 entnehmen lässt, sind die Abweichungen zu Durchschnittsbeziehenden eher gering.

Tabelle 2: Vorhergesagte Wahrscheinlichkeiten für ausgewählte Gruppen

	Typ 1: Komplex	Typ 2: Wechselhaft	Typ 3: Erwerb
Gruppe			
kein obligatorischer Schulabschluss aus EU8, geschieden/verwitwet/getrennt, 46 - 65 Jahre	24.7%	62.4%	12.9%
Abschluss auf Sekundarstufe I, aus EU8, zwischen 18 und 25 Jahren	13.4%	56.9%	29.6%
zwischen 45 und 65 Jahre, Schweizer-/in, ledig	8.6%	50.6%	40.8%
<i>Profil von Durchschnittsbeziehenden</i>			
lediger Mann, zwischen 26 und 45 Jahren, Schweizer, Gastgewerbe, mit Sekundarstufen II Abschluss.	7.6%	37.6%	54.7%
Abschluss auf Sekundarstufe II, unter 23 Jahre	4.1%	38.1%	57.8%

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010

Anmerkungen: Wahrscheinlichkeiten mit erhöhter Zugehörigkeit zu einem bestimmten Verlaufstyp sind **fett** markiert.

Monitoring der Risikostruktur von Leistungsbeziehenden

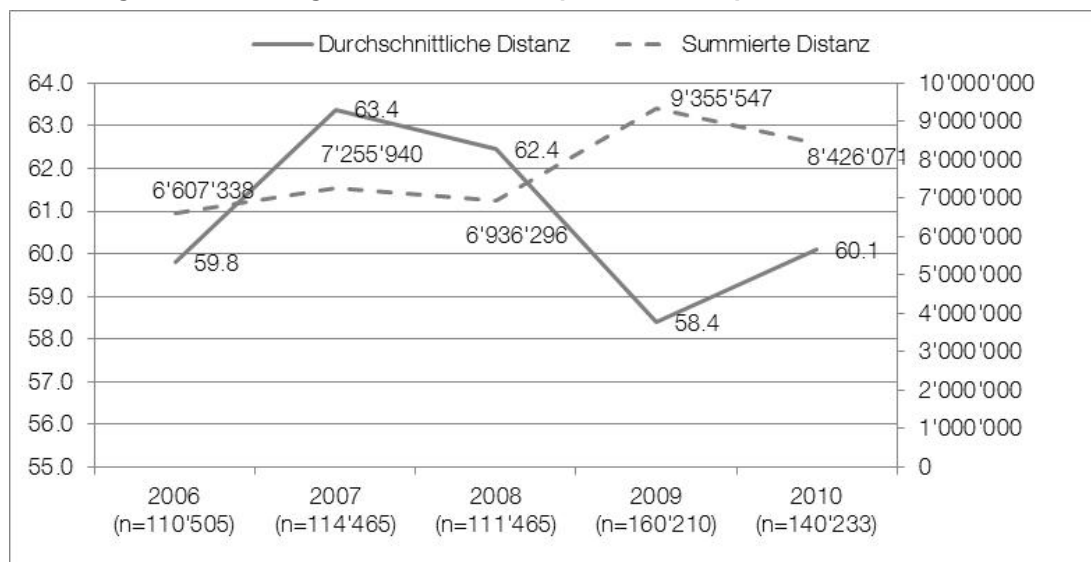
Auf der Grundlage der gemachten Auswertungen wurden weiter Überlegungen gemacht, wie sich daraus Indikatoren zur Beobachtung von strukturellen Verschiebungen ableiten lassen. Als Vorschlag konnte ein Risikomodell zur Prognostizierung der Risikostruktur entwickelt werden und auf Verlaufsbasis berechnete Systemindikatoren. Das Prognosemodell stellt Merkmale von Eintrittskohorten in Bezug zur erwarteten Distanz zur Erwerbstätigkeit und erlaubt somit eine kohortenweise Beurteilung der Veränderung der gesamtsystemischen Belastung der Sozialen Sicherheit. Im Rahmen der bisherigen Arbeiten wurde das Prognosemodell ausgehend von Eintritten über die ALV definiert. Dies aufgrund der spezifischen Bedürfnisse des Praxispartners SECO. Im Rahmen des Teilauftrags für das SECO wurden die Prognosewerte des Modells als Bestandteil des Tools zum Risikomonitoring integriert.

Als Systemindikatoren werden drei Arten unterschieden.

- (1) *Verlaufsindikatoren*. Diese umfassen die mittlere Dauer von Bezügen und Erwerbstätigkeit sowie alle relevanten Kombinationen von Parallel-Zuständen.
- (2) *Mobilitätsindikatoren* beschreiben die Komplexität von Verläufen über die Zahl der Zustandswechsel und die Zahl unterschiedlicher Zustände.
- (3) *Erwerbsintegrationsindex*. Dieser misst die Distanz zur Erwerbstätigkeit, wobei Phasen mit Leistungsbezügen über die empirisch beobachtete Wahrscheinlichkeit des Übertritts in die Erwerbstätigkeit gewichtet wurden (ALV ist „näher“ am Arbeitsmarkt als IV), vgl. nachfolgendes Anwendungsbeispiel.

Aus den in Abbildung 1 abgebildeten Werten lässt sich jahresweise ablesen, wie sich die Risikostruktur innerhalb von vier Jahren verändert hat. Höhere Werte entsprechen einer weiteren prognostizierten Distanz zur Erwerbstätigkeit. Damit ist eine Abschätzung der zukünftigen Gesamtbelastung des Systems der sozialen Sicherheit möglich. Für eine erste globale Betrachtung eignet sich die Interpretation der gestrichelten Linie (► Summierte prognostizierte Distanz). Dabei ist deutlich zu erkennen, wie sich die Zunahme der Fälle im Jahr 2009 in einer markant höheren summierten prognostizierten Distanz zur Erwerbstätigkeit widerspiegelt. Veränderungen ergeben sich jedoch nicht nur aus der reinen Zahl von Neubeziehenden. Vielmehr ist zweitens die veränderte Zusammensetzung von Neubeziehenden zu berücksichtigen. Dies wird aus der Betrachtung der durchgezogenen Linie (► Durchschnittliche prognostizierte Distanz) besser ersichtlicher. So ist etwa die Zunahme der Zahl der Neubeziehenden im Jahr 2009 vor allem auf eine Zunahme bei der ALV zurückzuführen, bei welchen Verläufe erwartungsgemäss näher am Arbeitsmarkt sind. Dies äussert sich in einer tieferen durchschnittlichen prognostizierten Distanz.

Abbildung 1: Veränderung der Risikostruktur (2006 bis 2010)

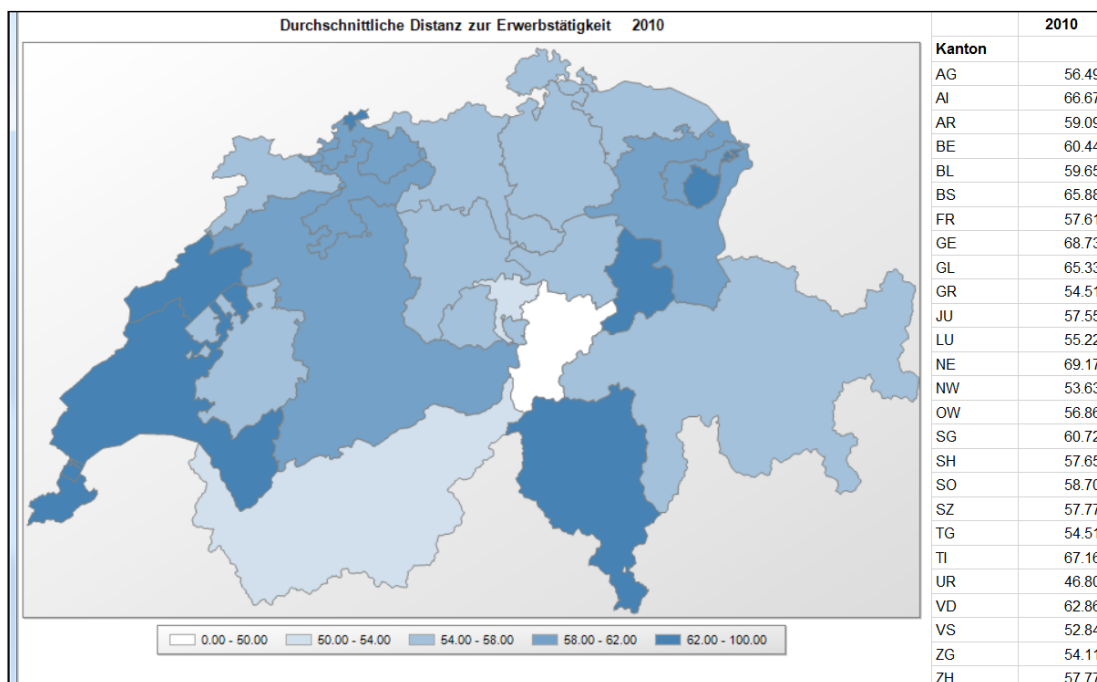


Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010, Berechnungen BFH
 Bemerkungen: Es ist die erwartete Distanz zur Erwerbstätigkeit abgebildet. Die Werte für das Jahr 2006 wurden empirisch beobachtet. Die Jahre 2007-2010 sind prognostiziert.

Entwicklung eines Auswertungstools mit Excel-Abfragemöglichkeiten

Als Teilauftrag für den Praxispartner SECO wurden die Auswertungen ausgehend von ALV-Eintrittskohorten konsolidiert in einem Monitoring-Tool. Dieses basiert auf einem SQL Server Analysis Services Cube. Der Zugriff kann über alle Programme erfolgen, die Analysis Services als Datenquelle unterstützen. Für Endbenutzer eignet sich dafür Microsoft Office Excel in der Version 2007 oder höher gut. Abfragen sind über Pivot-Tabellen möglich mit welchen Indikatoren in Bezug zur Kalenderzeit und weiteren Merkmalen (Region, Demografie) gesetzt werden können. Erweiterte Darstellungsmöglichkeiten erlaubt der SQL Server Reporting Service. Damit können Ergebnisse als Diagramme und kartographisch aufbereitet dargestellt werden. Abbildung 2 zeigt ein analoges Beispiel für einen Report mit einer Karte, die den Durchschnitt der Distanz zur Erwerbstätigkeit pro Kanton darstellt.

Abbildung 2: Risikoindikator Distanz zur Erwerbstätigkeit im Kantonsvergleich



Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010

Tabellenverzeichnis

Tabelle 1: Verlaufstypologie - Eintritt über alle Systeme	6
Tabelle 2: Vorhergesagte Wahrscheinlichkeiten für ausgewählte Gruppen	7
Tabelle 3: Gliederung der Auswertungsschritte	16
Tabelle 4: Codierung Zustände pro Quartal	21
Tabelle 5: Harmonisierte Variablen in SHIVALV	22
Tabelle 6: Codierung Zustände und Ereignisse	22
Tabelle 7: Filter und Fallzahlen der Grundgesamtheit für die Eintrittskohorte 2006	23
Tabelle 8: Übersicht zu den für die Clusteranalyse verwendeten Indikatoren	25
Tabelle 9: Verlaufstypologie - Eintritt über alle Systeme	27
Tabelle 10: Clustertypologie – differenziert nach Eintrittssystemen	28
Tabelle 11: Güte der Modelanpassung für alle Modelle	33
Tabelle 12: Typ 1 "Komplexe Verläufe mit ALV, Erwerb- und Sozialhilfe"	35
Tabelle 13: Typ 2 "ALV und Erwerb abwechselnd, mit Austritten aus Erwerb"	36
Tabelle 14: Vorhergesagte Wahrscheinlichkeiten für ausgewählte Gruppen	38
Tabelle 15: Typ 1 "Ohne Kontakt zum Arbeitsmarkt"	39
Tabelle 16: Typ 2 "Sozialhilfe und IV-Rente"	40
Tabelle 17: Typ 3 "Working Poor (Erwerb und SH)"	41
Tabelle 18: Vorhergesagte Wahrscheinlichkeiten für ausgewählte Gruppen	43
Tabelle 19: Typ 2 "Komplexe Verläufe mit offenem Ausgang"	44
Tabelle 20: Typ 3 "Sehr komplexe Verläufe mit langen Kombinationsbezügen"	45
Tabelle 21: Typ 4 "Einfache Verläufe, die schnell in Phasen mit IV-Rente/Erwerb münden" ..	46
Tabelle 22: Vorhergesagte Wahrscheinlichkeiten für ausgewählte Gruppen	48
Tabelle 23: Typ 1 "Mit IV-Rente"	50
Tabelle 24: Typ 2 "Existenzsicherung primär durch Sozialhilfe"	51
Tabelle 25: Typ 3 "Am Rand des Arbeitsmarktes (ALV, Erwerb, Phasen o. Erw./Leistungen)" ..	52
Tabelle 26: Vorhergesagte Wahrscheinlichkeiten für bestimmte Gruppen	54
Tabelle 27: Verlaufsindikatoren für Gesamtsystem und nach Eintrittssystemen	59
Tabelle 28: Mobilitätsindikatoren und Distanz zur Erwerbstätigkeit	60
Tabelle 29: Übertritts-Wahrscheinlichkeit, Zielzustand; Erwerbstätig	72
Tabelle 30: Transitionsmatrix	73
Tabelle 31: Risikomodell alle Eintrittssysteme (Zielvariable Distanz zur Erwerbstätigkeit) ..	78
Tabelle 32: Tabelle mit Wertelabels von ausgewählten Variablen	85

Abbildungsverzeichnis

Abbildung 1: Veränderung der Risikostruktur (2006 bis 2010)	8
Abbildung 2: Risikoindikator Distanz zur Erwerbstätigkeit im Kantonsvergleich	9
Abbildung 3: Zustandsverteilung von Verläufen im Gesamtsystem (n=110'508)	26
Abbildung 4: Zustandsverteilung von Verläufen mit Eintritt über ALV (n=85'262)	29
Abbildung 5: Zustandsverteilung von Verläufen mit Eintritt über Sozialhilfe (n=17'764)	30
Abbildung 6: Zustandsverteilung von Verläufen mit Eintritt über IV (n=7'024)	31
Abbildung 7: Bedeutsamkeit der Prädiktoren	34
Abbildung 8: Wahrscheinlichkeit der Zugehörigkeit zu Typ 1 - 3 für Bildungsgruppen	37
Abbildung 9: Zugehörigkeit zu Typ 1 - 4 für unterschiedliche Haushaltstypen	42
Abbildung 10: Zugehörigkeit zu Typ 1 - 4 für unterschiedliche IV-Rententeile	47
Abbildung 11: Zugehörigkeit zu Typ 1 - 4 für unterschiedliche Altersgruppen	53
Abbildung 12: Veränderung der Risikostruktur (2006 bis 2010)	56
Abbildung 13: Effekt des Alters auf die vorhergesagte Distanz zur Erwerbstätigkeit	57
Abbildung 14: Zahl der Zustandswechsel nach Eintritt in der zeitlichen Entwicklung	60
Abbildung 15: Übertritte von ALV zu SH nach Geschlecht und Nationalität	66
Abbildung 16: Anzahl ALV-Beziehende und Länge des ALV-Bezugs im Zeitverlauf	67
Abbildung 17: Länge des ALV-Bezugs im Kantonsvergleich	68
Abbildung 18: Risikoindikator Distanz zur Erwerbstätigkeit im Kantonsvergleich	69
Abbildung 19: Risikoindikator Distanz zur Erwerbstätigkeit als Zeitreihe	69

1 Einleitung

1.1 Ausgangslage

Der Fachbereich Soziale Arbeit der Berner Fachhochschule hat 2009 im Auftrag des Bundesamtes für Sozialversicherungen BSV die Untersuchung zur Quantifizierung der Übergänge zwischen den Leistungssystemen IV-Rente, IV-Taggelder, Arbeitslosenentschädigung und Sozialhilfe durchgeführt. Dies ermöglicht eine integrale Sicht auf die drei Leistungssysteme der sozialen Sicherheit. Dazu wurden die Administrativdatensätze der IV, der ALV sowie die Sozialhilfestatistik der Jahre 2004 - 2006 nach einem einheitlichen Modus aufbereitet, harmonisiert und ausgewertet (Fluder et al. 2009). Daran anschliessend stellt das BSV nun jährlich einen Gesamtdatensatz der 3 Leistungssysteme zusammen unter der Bezeichnung SHIVALV.

Auf Basis von zeitlich verknüpften Administrativdaten wurden durch den Fachbereich Soziale Arbeit der BFH weitere Projekte im Auftrag des SECO durchgeführt, so eine Pilotstudie zu Verläufen von ALE-Neubeziehenden und die darauf basierende Hauptstudie (Fluder et al. 2012). Im Auftrag des Bundesamts für Statistik BFS wurde eine erste Verlaufsstudie zu Sozialhilfebeziehenden durchgeführt (Salzgeber et al. 2010), die Neuauflage für einen aktualisierten Zeitraum ist derzeit in Bearbeitung. Im Bereich Invalidenversicherung wurde für das BSV die 5. IV-Revision evaluiert (Bolliger et al 2012). Zu den Verläufen in die IV wird derzeit ein Mandat für das BSV bearbeitet, in welchem die Verlaufsmuster in den Systemen der Sozialen Sicherheit vor dem ersten Rentenbezug für Neurentner/innen des Jahres 2010 untersucht werden (Fluder et al. 2013).

Aufbauend auf diesen Projekten sollen im vorliegenden Projekt mittels Methoden des Data Minings Risikoprofile erstellt werden bezüglich verschiedener Verlaufstypen innerhalb des Systems der Sozialen Sicherheit. Dabei sollen Merkmale des Leistungsbezugs der Systeme der Sozialen Sicherheit untersucht werden wie die Bezugsdauer, die Wiedereintritte und die Systemübergänge. Leistungskombinationen und deren Dynamik lassen sich nun über eine Periode von sechs Jahren (2005 bis 2010) analysieren. Die Ergebnisse aus den deskriptiven und wirkungsorientierten Auswertungen in bereits durchgeführten Projekten werden mit diesem Projekt durch Erkenntnisse ergänzt, die mittels komplexer iterativer Analysemethoden gewonnen werden.

Das Projekt Data Mining mit Administrativdaten der Sozialen Sicherheit ist ein interdepartementales Forschungsprojekt der Berner Fachhochschule, welches das Potenzial der Methoden des Data Mining im Bereich Soziale Sicherheit auslotet und daran anschliessende Vermarktungsmöglichkeiten sucht. Finanziert wird dieses hauptsächlich durch die BFH. Beteiligt daran ist neben dem Fachbereich Soziale Arbeit auch der Fachbereich Technik und Informatik TI. Die TI ist verantwortlich für die Mitkonzeption und Programmierung des Auswertungstools zum Langzeitmonitoring der Systeme der Sozialen Sicherheit in einer Gesamtsicht.

Das SECO unterstützt das Projekt mit einem Cashbeitrag und ermöglicht den Zugang zu den relevanten Datenbeständen. Im Gegenzug erhält das SECO erstens einen Teilbericht, der die mit dem Projektverantwortlichen beim SECO, Daniel Keller, abgestimmten Fragestellungen beantwortet. Zweitens erhält das SECO die Möglichkeit, unter Rücksprache mit der BFH den Prototyp des Auswertungstools für ein Langzeitmonitoring zu verwenden. Dessen Handhabung wird im Teilbericht beschrieben. Der Prototyp zeigt die Möglichkeiten für die spätere Entwicklung eines massgeschneiderten SECO-internen Instrumentes auf.

1.2 Zielsetzung und Fragestellung

Anhand der IV, ALV-, Sozialhilfe- und der IK-Daten der Jahre 2005 bis 2010 werden aussagekräftige Verlaufsmuster gefunden, die eine Grundlage für ein Risikomonitoring im Gesamtsystem der Sozialen Sicherheit bilden können. Ziel des Projekts ist es, komplexe Analysemethoden des Data Mining auf grosse Administrativdatensätze der Sozialen Sicherheit anzuwenden. Damit kann das Potential an enthaltenen Informationen in diesen Daten extrahiert werden. Die Analysemethoden sollen in einem Auswertungstool zusammengeführt werden, in welchem der Auftraggeber Auswertungen selber durchführen kann.

Folgende Fragestellungen sollen im Gesamtprojekt beantwortet werden:

1. Welche Methoden und welches Vorgehen aus dem Bereich des Data Mining, insbesondere aus den strukturerkennenden Verfahren wie z.B. verschiedene Formen der Clusteranalyse lassen sich erfolgreich zur Analyse eines grossen Administrativdatensatzes anwenden?
2. Lassen sich mittels dieser komplexen Methoden interpretierbare Klassen (Verlaufsmuster) identifizieren?
3. Lassen sich Faktoren finden, mit welchen sich die Zugehörigkeit zu solchen typischen Klassen prognostizieren lässt?
4. Lassen sich diese Methoden im Bereich der Qualitätssicherung der Daten anwenden, um die Qualität der Daten zu evaluieren und gegebenenfalls zu verbessern?
5. Lassen sich die als relevant festgestellten Methoden des Data Mining in einem Analyse-Instrument für die Langzeitbeobachtung konsolidieren, welches die Wiederholung der Analysen aus den Blickwinkeln der vier Systeme der Sozialen Sicherheit ermöglicht und welches für jährliche Erweiterungen des Datenmaterials geeignet ist? Kann dieses Instrument auch auf andere Administrativdatensätze angewendet werden?

Für den Teilbericht zuhanden des SECO waren insbesondere die Fragestellungen 2, 3 und 5 relevant. Da zu einem grossen Teil auf bereits standardisierte und harmonisierte Datensätze des BSV und des SECO zurückgegriffen wird, war die Fragestellung 4 zur Qualitätssicherung weniger relevant (vgl. Abschnitt 3.6). Fragestellung 1 bezieht sich auf die BFH-interne Entwicklung eines Instrumentariums an Methoden des Data Mining.

In Kapitel 2 wird die verwendete Datengrundlage beschrieben, Kapitel 3 stellt kurz die Vorgehensweise und Methodik in den einzelnen Arbeitsschritten vor. In Kapitel 4 geht auf die Resultate der Data Mining-Analysen ein, strukturiert nach Arbeitsschritten. Kapitel 5 schliesslich stellt die Verwendungsmöglichkeiten des Monitoring-tools dar.

2 Verwendete Datensätze

Die Daten bestehen aus Bezugsrecords der Sozialversicherungen (ALV und IV) und der Sozialhilfe, die durch das BSV im Gesamtdatensatz SHIVALV zusammengeführt und harmonisiert worden sind. Auf der Ebene von Individuen (bzw. Fallträgern in der Sozialhilfe) ist monatlich ersichtlich, welche Personen im Zeitraum von 2005 bis 2010 eines oder mehrere Systeme der sozialen Sicherheit beansprucht haben. Die SHIVALV-Daten können mit Angaben der individuellen Konten der Alters- und Hinterlassenenversicherung (AHV IK-Daten 2000-2010) ergänzt werden. Daraus lässt sich der Erwerbsstatus ableiten sowie der Bezug von IV-Taggeldern. Es ist allerdings zu beachten, dass IV-Taggeldbezüge nur dann betrachtet werden können, wenn in den Jahren 2005 bis 2010 einmal ein Bezug aus einem anderen System der Sozialen Sicherheit stattgefunden hat.²

Diese Datengrundlage erlaubt es, Bezugsmuster hinsichtlich der Nutzung des Gesamtsystems der Sozialen Sicherheit und der Erwerbstätigkeit im Längsschnitt zu untersuchen. Ergänzt werden können die Daten mit diversen zu administrativen Zwecken erhobenen Merkmalen der Beziehenden aus den verschiedenen Systemen der Sozialen Sicherheit, die für eine vertiefende Analyse zur Erklärung allfälliger Muster herbeigezogen werden können.

In allen Systemen stehen folgende Informationen für die Auswertung zur Verfügung:

- *Leistungsangaben (IV-Rente, IV-Taggeld, ALV, SH):*
Leistungstyp, Beginn und Ende eines Bezugs, Beträge
- *Demografische Angaben zu den Personen in allen Leistungssystemen:*
Wohnort, Alter, Geschlecht, Zivilstand, Nationalität
- *Spezifisch pro Leistungssystem* stehen noch weitere Informationen zur Verfügung: *IV:* Gebrechenscode, Teilrente, Unterstützungspflicht (Kinder) etc.; *ALV:* Ausbildung, Qualifikation, Branche etc.; *SHS:* Angaben zur *Unterstützungseinheit* (Fall bzw. Dossier): Einkommenskomponenten, Bruttobedarf, Anzahl Personen pro Fall, Haushaltstyp, involvierte Stellen etc. sowie Angaben zu den *unterstützten Personen:* Ausbildung, Qualifikation, Grund für Erwerbslosigkeit etc.

Für die Analyse wurden folgende Datenbestände verwendet:

- SHIVALV-Daten 2005 bis 2010: Für die einzelnen Jahre wurde die folgende Anzahl Einzeldaten geliefert:

2005	651'267
2006	638'592
2007	605'353
2008	584'152
2009	640'561
2010	663'583

² Diese Einschränkung führt dazu, dass die IV-Taggelder als Eintrittssystem nur eine untergeordnete Rolle spielen können. Aussagen für Verläufe von Personen mit Eintritt über IV-Taggelder sind im Rahmen dieses Projekts nicht möglich, da Verläufe von Personen, die danach keine weiteren Bezüge aus anderen Systemen der Sozialen Sicherheit aufweisen, fehlen. Bei einer späteren Wiederholung könnte die Abgrenzung der IK-Daten entsprechend angepasst werden, so dass diese Verläufe auch abgebildet würden.

- AHV-IK Daten 2000 bis 2010: Für die einzelnen Jahre wurden die folgenden Anzahl Einzeldaten geliefert, für Personen, die in den oben aufgeführten Datensätzen des SHIVALV mindestens einmal vorkommen:

2000	886'872
2001	924'674
2002	966'331
2003	1'003'861
2004	1'039'392
2005	1'077'790
2006	1'091'274
2007	1'091'511
2008	1'098'542
2009	1'066'475
2010	1'005'577

- AVAM / ASAL Daten des SECO 205 bis 2010: Für die einzelnen Jahre wurden die folgenden Anzahl Einzeldaten geliefert (monatliche Einträge, pro Person mehrere pro Jahr möglich):

	AVAM/PLASTA	ASAL/SIPAC
2005	2'577'060	1'818'518
2006	2'055'451	1'629'048
2007	2'409'723	1'346'181
2008	1'879'539	1'227'630
2009	2'327'353	1'674'554
2010	2'392'000	1'826'696

Die Datenbestände wurden vom BSV mit einer verknüpfbaren anonymen ID versehen, so dass die Personenrecords sowohl über die Jahre als auch für verschiedene Systeme innerhalb eines Jahres miteinander verknüpft analysiert werden konnten.

3 Vorgehensweise

Auf der Grundlage der bestehenden Datenbasis und hinsichtlich oben formulierter Ziele werden nachfolgend sechs Auswertungsschritte definiert. Im ersten Arbeitsschritt werden die Datengrundlagen aufbereitet. In einem zweiten Arbeitsschritt wird untersucht, welche Verlaufsmuster häufig auftreten und ob sich ähnliche bzw. unterschiedliche Verlaufsmuster identifizieren lassen. Ein dritter Arbeitsschritt fokussiert auf Zusammenhänge zwischen Bezugsmuster und Merkmalen der Beziehenden mit dem Ziel, Risikoprofile zu erstellen. Im vierten Arbeitsschritt wird ein Risikomodell erstellt, das zum Ziel hat, anhand von Eintrittsmerkmalen zukünftige Verschiebungen hinsichtlich der Distanz zur Erwerbstätigkeit zu prognostizieren. In einem fünften Schritt wurden deskriptive Auswertungen auf Systemebene vorgenommen, um die ermittelten Verlaufsmuster auch aggregiert betrachten zu können. Der sechste Auswertungsschritt befasst sich mit der Überführung der vorangehenden Auswertungen in ein Monitoring-tool.

Tabelle 1 zeigt die Gliederung der Auswertungsschritte und die dabei verwendeten Methoden und Softwares sowie deren zentrale Ergebnisse.

Tabelle 3: Gliederung der Auswertungsschritte

Arbeitsschritte	Software	Statistische Methoden	Zentrales Ergebnis
Daten aufbereiten (Arbeitsschritt 1)	SQL Server Integration Services, SQL Server	Datenbank	Originaldaten sind in Datenbank integriert und zu Arbeitsdatensätzen aufbereitet.
Bilden von Verlaufsmustern (Arbeitsschritt 2)	Stata	Sequenzdaten- und Clusteranalyse	Verlaufstypologie, Gruppierung von Verläufen
Bilden von Risikoprofilen / Schätzen der Zugehörigkeit zu Verlaufsmustern (Arbeitsschritt 3)	SPSS	Multinomiale Logit-Regression	Zusammenhangsmodelle, Risiko- und Resilienzfaktoren für Zugehörigkeit zu Verlaufstypen
Modell für Prognose der Risikostruktur (Arbeitsschritt 4)	SPSS	Lineare Regression	Ausarbeiten eines Risikomodelles zur Prognose der zukünftigen Risiken im System der Sozialen Sicherheit
Indikatoren für die Systeme der Sozialen Sicherheit (Arbeitsschritt 5)	SPSS/Stata	deskriptive Methoden der Sequenzdaten-analyse	Verweildauer- und Mobilitätsindikatoren
Risikomonitoring (Arbeitsschritt 6)	SQL Server Analysis Services Cube	Kreuztabelle, Datenbank	Auswertungs-tool mit hinterlegten Informationen aus Arbeitsschritten 1 bis 5. Räumliche, zeitliche und weitere Gruppierungen möglich.

3.1 Arbeitsschritt 1 – Daten aufbereiten

Als Grundlage für die nachfolgenden Arbeitsschritte wurden zu Beginn des Projekts zunächst die Originaldaten aus den verschiedenen Datenquellen in eine Datenbank integriert und zu Arbeitsdatensätzen aufbereitet. Die Originaldaten wurden in Form einzelner SPSS, SAS und CSV Dateien angeliefert, wobei pro Jahr und Datenquelle typischerweise eine separate Datei vorlag. Diese einzelnen Dateien wurden extrahiert und in einer relationalen SQL Server Datenbank gespeichert. Daten aus der gleichen Quelle wurden dabei jeweils in einer eigenen Tabelle abgelegt. So wurden z.B. die Dateien mit den SHIVALV Daten von 2005 bis 2010 in einer einzelnen Datenbanktabelle zusammengefasst. Dies ermöglichte die einfache Erzeugung von Datensätzen über mehrere Jahre für die Analyse.

Im Verlaufe des Projekts wurden dann die Arbeitsdatensätze für die einzelnen Analyseschritte basierend auf den jeweiligen Vorgaben und Anforderungen der verwendeten Methode aufbereitet. Bei der Aufbereitung wurden unter anderem Datensätze ausgefiltert, Daten verschiedener Quellen miteinander verknüpft, Attribute recodiert, neue Attribute errechnet sowie Aggregationen und Transformationen mittels SQL vorgenommen (Siehe auch Abschnitt 3.7 zur Datenaufbereitung). Die resultierenden Arbeitsdatensätze wurden entweder in neuen Datenbanktabellen gespeichert oder durch Views dynamisch aus den bestehenden Tabellen abgeleitet. Bei den eigentlichen Analyseschritten wurden diese Datensätze dann direkt von den Data Mining Programmen aus der Datenbank eingelesen.

3.2 Arbeitsschritt 2 – Analyse von Verlaufsmustern

Eine erste Möglichkeit zur Bestimmung von Verlaufsmustern besteht darin, die Komplexität der Bezugsmuster anhand konzeptioneller, theoretischer Eingrenzungen zu verdichten, wie es in den vergangenen BFH-Projekten mit den SHIVALV-Daten gemacht wurde. Hierbei wurden ausgehend von Eintritten in ein bestimmtes System der weitere Verlauf bzw. der vorangehende Verlauf im Rahmen der Sozialen Sicherheit untersucht. Typische Verlaufstypen dieser Vorgehensweise sind „ALE kurz“ oder „Pendler“, wobei ersterer in einem Bezug von Arbeitslosenentschädigung während weniger als 12 Monaten besteht und letzterer einen Verlaufstyp darstellt, der mindestens zwei Mal von einem Leistungssystem der Sozialen Sicherheit in ein anderes wechselt.

Sequenzdatenanalyse

Im Rahmen der methodischen Abklärungen zum Data Mining Prozess sind wir auf die im europäischen Raum bisher wenig verbreitete Sequenzdatenanalyse gestossen. Dieses Verfahren erlaubt im Gegensatz zur oben beschriebenen theoriegeleiteten Analyse die datengetriebene Beschreibung und Analyse von zeitlichen Sequenzen. Darunter fallen etwa Lebensverläufe, aber auch Berufs- oder Erwerbsverläufe (Scherer und Brüderl 2010). Typische Fragestellungen lauten: Welche Typen von Verläufen gibt es? Wie homogen sind die Verläufe von bestimmten Gruppen? Was sind Ursachen von bestimmten Verlaufsmustern? Die von Scherer und Brüderl vorgeschlagene Vorgehensweise scheint insbesondere für sozialwissenschaftliche Fragestellungen geeignet und zeichnet sich durch eine inhaltlich stringente Vorgehensweise aus, die klare Ergebnisse verspricht, ohne dass eine starke Eingrenzung der Ursprungsdaten notwendig ist.

Die Sequenzdatenanalyse stellt spezifische Anforderungen an die Datenstruktur. So müssen die Daten aus einer zeitlichen Abfolge von klar definierten Zuständen bestehen. Zur sequenziellen Darstellung von Verläufen müssen die Zeitachse und der Zustandsraum definiert werden:

- *Zeitachse:* Es muss ein standardisiertes Beobachtungsfenster definiert werden, wodurch die Sequenzen auf dieselbe Länge gebracht werden. Es gilt also Entscheidungen bezüglich möglicher Zensierungen rechts bzw. links zu fällen.
- *Zustandsraum:* Für jeden Monat muss jeder Beobachtung ein möglicher Zustand zugewiesen

werden. Dabei stellt sich im vorliegenden Fall die Frage, wie mit Doppelbezügen bzw. Doppelzuständen umgegangen wird. Möglich ist, dass ein Zustand einen anderen überschreibt oder dass Doppelzustände einen eigenen Zustand abbilden.

Nach Scherer und Brüderl (2010) beinhaltet eine Sequenzdatenanalyse:

(a) Beschreibende Kennzahlen für die individuellen Sequenzen

- Verweildauer in einzelnen Zuständen (beispielsweise: kumulierte oder mittlere Dauer der Arbeitslosigkeit)
- Anzahl Zustandswechsel (Ereignisse), Mobilitätskennzahlen (Beispielsweise Anzahl Wechsel zwischen den Systemen)

(b) Ähnlichkeit bzw. Distanzbestimmung zwischen Sequenzen (Distanzmass)

- Vergleich zu einer Referenzsequenz (beispielsweise durchgängig erwerbstätig)
- Paarweiser Vergleich von Sequenzen

(c) Typenbildung von Verläufen mittels Clusteranalyse. Dafür wird zuerst eine Distanzmatrix berechnet (vgl. Schritt (b)), auf deren Basis die Verläufe anschliessend klassifiziert werden.

Clusteranalyse

In der Clusteranalyse werden Gruppen so gebildet, dass sich die Beobachtungen innerhalb der Gruppe möglichst ähnlich und die Unterschiede zwischen den Gruppen möglichst gross sind. Es wird also eine Gruppierung der Daten angestrebt, die vorerst losgelöst von inhaltlichen Kriterien erfolgt. Ob diese Gruppierung nützlich bzw. sinnvoll ist, muss im Zuge der Interpretation der Ergebnisse entschieden werden. Grundlage der Clusteranalyse ist die Distanzmatrix, die Abstände zwischen den Verläufen auf metrischem Skalenniveau wiedergibt.

3.3 Arbeitsschritt 3 – Erstellen von Risikoprofilen

Ergebnis des ersten Arbeitsschrittes ist eine Typologie. Mit dieser Typologie als Zielvariable lassen sich im dritten Arbeitsschritt Zusammenhangsmodelle berechnen (Multinomial logistische Regression). Damit werden Faktoren eruiert, die in Zusammenhang mit der Zugehörigkeit zu einem bestimmten Verlaufstyp stehen. Analysiert werden dabei persönliche Merkmale der Personen (Geschlecht, Nationalität, Alter, Beruf, Bildungsstand etc.). Ziel dieses Arbeitsschrittes ist es Risikoindikatoren und Risikogruppen zu identifizieren.

Die statistischen Analysen zielen dabei auf eine adäquate Modellierung der Zusammenhänge ab, die als deskriptive Modelle mit optimaler Anpassung an die Daten verstanden werden sollen.³ Dabei gilt es zu bedenken, dass die gefundenen Zusammenhänge nicht kausal zu interpretieren sind. Wenn etwa ein bestimmtes Merkmal (bspw. Nationalität) auf ein erhöhtes Risiko hinweist, dann ist damit nicht gesagt, dass eine Veränderung (Einbürgerung) das ursprünglich entdeckte Risiko senkt. Vielmehr scheint es sinnvoll, gefundene Zusammenhänge als Hinweise auf Gruppenmerkmale zu verstehen, die auf spezifische Verläufe verweisen.

3.4 Arbeitsschritt 4 – Modell zur Prognose der Risikostruktur

Im vorangehenden Arbeitsschritt wurden Erklärungsfaktoren für die Zugehörigkeit zu bestimmten Verlaufstypen ermittelt. Dabei können gewisse Faktoren bzw. Ausprägungen als eigentliche Risikofaktoren für

³ “[...] regression can be justified without reference to causality, as it can be considered nothing more than a method for obtaining a best-fitting descriptive model” (Morgan & Winship 2007:123).

die Zugehörigkeit zu einem problematischen Verlaufstyp verstanden werden. Problematische Verlaufstypen sind solche, die stark von einem durchgehenden Erwerbsverlauf abweichen. In Arbeitsschritt 2 wurde eine solche Abweichung mit dem Distanzmass gemessen.

In Arbeitsschritt 4 soll nun die Höhe der Distanz erklärt werden mittels Faktoren des Erstzustandes von persönlichen Merkmalen sowie danach eintretenden persönlichen Ereignissen (Veränderungen gegenüber dem Erstzustand). Dadurch wird es möglich, Erklärungsfaktoren für problematische Verläufe innerhalb des Systems der Sozialen Sicherheit zu gewichten, als Methode dient dabei die lineare Regression. Für jedes Eintrittssystem sowie für alle Systeme gleichzeitig werden unterschiedliche Modelle geschätzt.

Auf Basis des Regressionsmodells kann schliesslich eine Prognose für zukünftige Entwicklungen der Risikostruktur erstellt werden. Diese Prognose kann entsprechend unterschiedlichen Eingrenzungen der Neueintritte nach Systemen auch für Einzelsysteme gemacht werden. Z.B. kann für Neueintritte in der ALV gemessen werden, ob sich die Risikostruktur für Verläufe (Distanz zu durchgehender Erwerbsarbeit) in Zukunft verbessern oder verschlechtern wird.

3.5 Arbeitsschritt 5 – Indikatoren für die Systeme der Sozialen Sicherheit

Der Blick der Analysen soll hier auf die Systeme der sozialen Sicherung gerichtet werden. In unserer Analyse wird die Struktur der Daten nach Personen beibehalten. Als Ereignisse werden im 5. Arbeitsschritt vorerst Wechsel des Zustands bezüglich Bezug aus dem System der Sozialen Sicherheit sowie der Erwerbstätigkeit betrachtet.

Beispiele relevanter Ereignisse aus der Perspektive der ALV sind der Eintritt in die Erwerbstätigkeit, Eintritt in die Sozialhilfe, Eintritt in die IV. Um Aussagen auf der Systemebene ableiten zu können, werden die Wechsel und Verweildauern für einzelne Zustände gemessen und für bestimmte Zeiträume aggregiert. Mit solchen Mobilitätsindikatoren können einfache Indikatoren zu den Handlungsspielräumen innerhalb der unterschiedlichen Systeme ausgewiesen werden.

Zudem können die aufbereiteten Daten nach einzelnen Eintrittssystemen gefiltert werden, so dass Verläufe in den Systemen und der Erwerbstätigkeit ausgehend von Neueintritten in ein bestimmtes System der Sozialen Sicherheit betrachtet werden können.

3.6 Arbeitsschritt 6 – Risikomonitoring SECO

Ein Risikomonitoring im Sinne des SECO setzt auf einer höheren Aggregationsebene an und soll die Veränderung von Risikoindikatoren hinsichtlich persönlicher Merkmale (Geschlecht, Alter etc.) oder hinsichtlich regionaler Unterschiede aufzeigen. Zu diesem Zweck soll ein Auswertungstool zuhanden des SECO programmiert werden, in dem die relevanten Informationen aus den vorangehenden 5 Arbeitsschritten enthalten sind, so dass durch Abfrage von Quer- und Längsschnitten global und für Untergruppen die interessierenden Risikoabklärungen durch das SECO selber durchgeführt werden können.

Mögliche Risikoindikatoren: Veränderung der Anzahl unerwünschter Ereignisse (Zunahme Wiederanmeldungen bei der Sozialhilfe, Zunahme Aussteuerungen) oder Zunahme von unerwünschten Verläufen, die sich stärker vom Ideal der durchgehenden Erwerbsarbeit entfernt haben oder voraussichtlich entfernen werden.

Daraus könnte ein nationales Frühwarnsystem entwickelt werden, welches bruchhafte Veränderungen oder

unerwünschte Trends sichtbar macht.

3.7 Datenaufbereitung

3.7.1 Grundsätzliche Bemerkungen

Bei der Harmonisierung und Standardisierung stützen wir uns weitgehend auf die bereits durch die zuständigen Ämter vorgenommenen Normierungen und Harmonisierungen, da wir als Basis mit dem bereits durch das BSV aufbereiteten Datensatz SHIVALV arbeiten (vgl. Abschnitt 3.6.3). Zusätzliche Normierungen und Aggregationen einzelner Variablen (vgl. z.B. Abschnitt 3.6.2) wurden aufgrund der Erfahrungen aus vorangehenden Projekten umgesetzt.

3.7.2 Festlegung von Zuständen

In SHIVALV liegt für jede Person jahrweise die Information zu Bezügen der Sozialen Sicherheit vor (festgehalten in 12er-Strings). Die Aufbereitung beinhaltet

- a) die monatsweise Prüfung aller möglichen Leistungskombinationen, die zu Zuständen verdichtet werden
- b) eine Aggregation auf Quartale. Testauswertungen und Diskussionen mit unterschiedlichen Expertinnen und Experten haben zum Schluss geführt, dass die Aufbereitung in Quartalen die beste Grundlage zur Bildung einer Verlaufstypologie darstellt.⁴

⁴ -SH: Genauigkeit ½ Jahr (erst nach ½ Jahr kein Bezug Ablösung)

-ALV: Genauigkeit 2 Monate (jeder Monat mit einem Bezugstag wird gezählt)

-IK: Genauigkeit 3 Monate (Arbeitgeber geben oft Auszahlungsmonat an, Erwerbsunterbrüche innerhalb des Jahres bei nachträglichen Buchungen nicht abgebildet)

Tabelle 4: Codierung Zustände pro Quartal

Einfache Codierung (1. Version)					
Zustandskombination: 1: zutreffend 2: nicht zutreffend Egal: 1 oder 2					
Code 1	ALV	IVR	SH	IVT	ER
A0	1	2	2	2	2
A1	1	2	2	2	1
B0	1	1	2	egal	2
B1	1	1	2	egal	1
D0	1	2	2	1	2
D1	1	2	2	1	1
E0	1	1	1	egal	2
E1	1	1	1	egal	1
I0	2	1	2	egal	2
I1	2	1	2	egal	1
C0	1	2	1	2	2
C1	1	2	1	2	1
J0	2	1	1	egal	2
J1	2	1	1	egal	1
U0	2	2	1	1	2
U1	2	2	1	1	1
S1	2	2	1	2	1
S0	2	2	1	2	2
T0	2	2	2	1	2
T1	2	2	2	1	1
X1	2	2	2	2	1
X0	2	2	2	2	2
F0	1	2	1	1	2
F1	1	2	1	1	1

Einfache Codierung (2. Version)	
Zustandskombination	Code 2
ALV	A0
ALV und Erwerb	A1
ALV SH	C0
ALV SH mit Erwerb	C1
IV-Taggeld	T0
IV-Taggeld mit Erwerb	T1
SH kombi IVR und/oder IVT, ev. ALV	U0 J0 E0 F0 > kombi1
SH kombi IVR und/oder IVT, ev. ALV mit Erwerb	U1 J1 E1 F1 > kombi1_er
Sozialhilfe	S0
Sozialhilfe mit Erwerb	S1
IV-Rente, ev. IVT	I0
IV-Rente mit Erwerb, ev. mit IVT	I1
Erwerb	X1
Kein Einkommen und keine Leistungen	X0
IV-Rente und/oder IVT mit ALV	B0 D0 > kombi2
IV-Rente und/oder IVT mit ALV mit Erwerb	B1 D1 > kombi2_er

Quelle: Darstellung BFH

Ausgangspunkt für die Festlegung der quartalsweisen Zustände ist die Kombinationsmatrix auf der linken Seite in Tabelle 2. Für die Codierung eines Bezugs bzw. einer Erwerbstätigkeit in einem Quartal muss dieser in mindestens 2 der 3 Monate des Quartals vorliegen. Auf der rechten Seite von Tabelle 2 wird zudem festgehalten, welche Codes der einfachen Codierung in einem zweiten Schritt verdichtet wurden, so dass insgesamt pro Quartal 16 Zustandskombinationen möglich sind.

3.7.3 Personenmerkmale und Ereignisse

Es wurde ein Personendatensatz mit allen im Zeitraum 2005 bis 2010 erfassten Personen gebildet. Die Beobachtungseinheit (Datenzeile) entspricht einer Person und beinhaltet alle relevanten Informationen, die für den Arbeitsschritt „exploratives Data Mining zu Risikoprofilen“ notwendig sind. Die Informationen stammen aus drei Quellen:

- SHIVALV 2005 bis 2010
- IK-Daten 2005 bis 2010
- Separate AVAM/ASAL bzw. PLASTA/SIPAC-Daten 2005 bis 2010

In Tabelle 5 werden die harmonisierten Variablen aufgeführt, die für die Systeme IV, ALV und SH in gleicher Form vorhanden sind.

Tabelle 5: Harmonisierte Variablen in SHIVALV

Harmonisierte gemeinsame Variablen		
gesl	GESCHLECHT	1=M, 2=F (mit Versichertenregister abgeglichen)
kanton	KANTON	Kantonsnummer BFS (1-26)
zivil	ZIVILSTAND	1=Ledig, 2=Verheiratet, 3=Verwitwet, 4=Geschieden, 5=Getrennt (Bei nur ALV-Bezügern => Verheiratet/Getrennt = 2)
alterj	VOLLENDETES ALTERSJAHR (per 31. Dez.)	(mit Versichertenregister abgeglichen)
nation	NATIONALITÄT CH/AUSLAND	1=CH, 2=AUSLÄNDER
nation_code_BFS	STAAT	Staatencodes BFS (8xxx, Beilage)
Harmonisierung bei Bezug aus mehreren Systemen:		
(IV/ALV = Kanton+Nation aus ALV, Zivilstand aus IV)		
(IV/SH = Kanton , Nation , Zivilstand alle aus SH)		
(ALV/SH = Kanton+Nation aus ALV, Zivilstand aus SH)		

Quelle: SHIVALV

Für einzelne Systeme liegen weitere Variablen vor, vgl. Anhang.

Die Personendaten liegen jeweils in jährlich Auszügen vor. Dabei muss in der Aufbereitung die vorhandene Zensierung abgebildet werden. Solange Personen nicht mit den Systemen der sozialen Sicherheit in Kontakt stehen, liegen keine Informationen vor. Durch laufende Ein- und Austritte können so individuell unterschiedliche Beobachtungszeiträume entstehen. Diese können teilweise durch die Fokussierung auf Eintrittskohorten aufgefangen werden. Verkürzte Erhebungszeiträume für die unterschiedliche Teilerhebungen (ALV, IV, SH) können nach wie vor auftreten. Möglich sind sowohl Links- wie Rechtszensierungen. Diesem Umstand wird insofern Rechnung getragen, dass die Aufbereitung jeweils Variablen zu vorhanden (Teil-)Beobachtungsjahren vorsieht und daraus Analysen zur Interpretierbarkeit der zeitlichen Veränderungen möglich sind.

Durchgeführt wird eine Aufbereitung, die den Zustand zu Beginn des Systemeintrittes und am Ende abbildet. Veränderungen werden in Ereignis-Variablen festgehalten, die ausgehend vom Jahr des Systemeintrittes Informationen zu Ereignissen von vier Folgejahren abbilden. Veränderungen vier Jahre nach Systemeintritt können ignoriert werden. Damit steht der Beobachtungsfokus im Einklang mit den Analysen der Verläufe. Als zusammenfassende Variable wird bei *nominalen Merkmalen* die Anzahl Wechsel notiert, bei *ordinalen Merkmalen* werden die Wechsel zusätzlich durch vier Variablen bewertet. Ausserdem wird der letzte beobachtete Zustand in einer Zustandsvariablen und einer Variable mit der relativen Zeit bis zum Wechsel festgehalten. Falls es aus inhaltlichen Gründen nicht passender erscheint, beinhaltet die Aufbereitung demnach zwei Varianten.

Tabelle 6: Codierung Zustände und Ereignisse

<i>Nominale Merkmale</i>	<i>Ordinale Merkmale</i>
<ul style="list-style-type: none"> - 1. Zustand - 4. Ereignisvariablen pro Jahr mit 1=Wechsel, 0=kein Wechsel, -99=nicht beobachtet - 1 zusammenfassende Variable mit Anzahl Zustandswechsel - letzter Zustand - Jahr letzter Zustand (relative Zeit) 	<ul style="list-style-type: none"> - alle Variablen für nominale Merkmale -4 Variablen mit ergänzenden Informationen zum Wechsel, die allenfalls bereits eine Bewertung des Wechsels beinhalten. Nachfolgen sind Konkretisierungen vorgenommen.

Quelle: Darstellung BFH

4 Verlaufs- und Risikomusteranalyse

4.1 Konzeptionelle Eingrenzungen

4.1.1 Eingrenzung der Grundgesamtheit

Zur besseren Interpretierbarkeit der Ergebnisse erfolgen die Analysen mit der Eingrenzung auf Personen, die im Jahr 2006 zum ersten Mal mit dem System der Sozialen Sicherung in Kontakt gekommen sind. Diese Gruppe bezeichnen wir als Eintrittskohorte 2006. Die Kriterien zur Bestimmung dieser Gruppe sind nachfolgend beschrieben.

A) Identifizierung von Neueintritten

Anhand der SHIVALV-Daten wird für jede Person der erste Bezugsmonat eruiert. Da die SHIVALV-Daten erst ab Januar 2005 zur Verfügung stehen, lassen sich für das Jahr 2005 Neueintritte nicht vollumfänglich von Durchbeziehenden unterscheiden. Deswegen begrenzen wir die Analyse auf Eintritte im Jahr 2006. Damit steht für die Mustererkennung der Verlaufsanalysen für jede Person ein Beobachtungszeitraum von mindestens vier Jahren zur Verfügung. Testauswertungen haben gezeigt, dass Verläufe innerhalb der vier Jahre eine gewisse Konsolidierung erfahren – wenige Wechsel sind zu späteren Zeitpunkten zu beobachten. Das Kriterium „Neueintritte“ wird präzisiert, indem wir Fälle ausschliessen, die im Jahr 2004 bereits ALE oder IV-Taggelder bezogen haben, wie es sich anhand der AHV-IK-Daten rekonstruieren lässt

B) Ausschliessen von Personen mit fehlerhaften Einkommensdaten

Erwerbsverläufe sind über die AHV-IK Daten in der Regel monatsweise abgebildet. Teilweise ist die monatsweise Verortung nicht gegeben. Solche Fälle werden ausgeschlossen.

C) Ausschliessen Personen mit Rentenbezüge innerhalb des Beobachtungszeitraumes

Ebenfalls nicht berücksichtigt werden Personen, die innerhalb des Beobachtungszeitraums das Rentenalter erreichen. Damit wird die Grundgesamtheit auf Personen im Erwerbsalter eingegrenzt.

Die Auswirkungen der Filter auf die Fallzahlen der Grundgesamtheit werden in Tabelle 5 dargestellt. 141'091 Personen erfüllen das Eintrittskriterium. Mit der Anwendung der Filterkriterien verbleiben 110'508 Personen, deren Verläufe analysiert werden.

Tabelle 7: Filter und Fallzahlen der Grundgesamtheit für die Eintrittskohorte 2006

Eintrittskriterium - Erstbezug von ALE, SH, IV oder IV-T im Jahr 2006(ausgehend von 2005)
141'091 Personen

Filter 1 - ALE-IVT Bezug 2004
(ohne Berücksichtigung anderer Filter)
- 15,210

Filter 2 - Fehlerhafte Einkommensdaten
(ohne Berücksichtigung anderer Filter)
- 10'715

Filter 3 - Rentenbezüge innerhalb des Beobachtungszeitraumes
(ohne Berücksichtigung anderer Filter)
- 6'765

Grundgesamtheit mit Eintritts- und Filterkriterien
110'508

Quelle: Darstellung BFH

4.1.2 Umgang mit der zeitlichen Dimension

Das *Zeitfenster* für die Verlaufsanalysen wird für alle Personen auf vier Jahre standardisiert, d.h. für jede Person wird ein Zeitraum über vier Jahre betrachtet, auch wenn möglicherweise Angaben über längere Zeiträume vorliegen. Damit werden Verlaufsmuster besser vergleichbar. Wie bereits erwähnt, haben Testauswertungen zudem ergeben, dass weitere Veränderungen hinsichtlich der Bezugsmuster nach vier Jahren deutlich seltener vorkommen. Die Testauswertungen haben weiter ergeben, dass eine monatsweise *Aufbereitung der Bezugsmuster* mit einer gewissen Unschärfe verbunden ist. Damit das Auftreten von Parallelbezügen nicht überschätzt wird, haben wir uns daher entschlossen, mit einer quartalsweisen Betrachtung zu arbeiten. Diese beinhaltet, dass Bezüge bzw. Zustände nur als solche gewertet werden, wenn in zwei von drei betrachteten Monaten ein Bezug stattgefunden hat.

4.2 Verlaufstypologien

Damit Muster in der Vielzahl der beobachteten Verläufe erkannt werden können, kombinieren wir an dieser Stelle Methoden der Cluster- und der Sequenzdatenanalyse. Wie eingangs beschrieben, ermöglicht eine Clusteranalyse Beobachtungen so zu gruppieren, dass die Beobachtungen innerhalb eines Clusters möglichst ähnlich und die Unterschiede zwischen den Cluster möglichst gross sind. Damit werden Muster in den Daten freigelegt, die weitgehend im Informationsgehalt der Daten zu begründen sind. Auf konzeptionelle Festlegungen wird weitgehend verzichtet. Festgelegt wird die Auswahl der für die Analysen verwendeten Indikatoren, bei deren Berechnung wir auf Methoden der Sequenzdatenanalyse zurückgreifen. Eine weitere Festlegung erfolgt bei der Konkretisierung der methodischen Schritte (Wahl des Verfahrens der Clusterbildung, Distanzmass, Bestimmung der Clusterzahl).

4.2.1 Verlaufsindikatoren

Eingang in die Analysen finden Verlaufsmerkmale, die auf individueller Ebene Auskunft zur kumulierten Dauer von Bezügen, Phasen mit Erwerb sowie Phasen ohne Erwerb und Bezügen über vier Jahre geben. Ebenso berücksichtigt wird ein entwickelter Indikator zur Bestimmung der Distanz eines Verlaufes zum Erwerb, den wir als Erwerbsintegrationsindex⁵ bezeichnen (vgl. Anhang 7 zur Berechnung). Der Erwerbsintegrationsindex beziffert für jeden Verlauf dessen Distanz zu einem idealtypischen Verlauf mit durchgängiger Erwerbstätigkeit. Damit ist eine numerische Aussage zur Erwerbsintegration über den gesamten Beobachtungszeitraum von vier Jahren möglich. Ergänzt werden diese metrischen Merkmale mit nominalen Indikatoren, die den Zustand zu bestimmten Zeitpunkten ausgehend vom Eintritt wiedergeben (Zustand zu Beginn, nach einem Jahr und nach vier Jahren). Mit letzteren wird die chronologische Abfolge eines Verlaufes abgebildet. Ebenfalls berücksichtigt wird die Komplexität von Verläufen, gemessen als Anzahl unterschiedlicher Zustände und Anzahl Zustandswechsel. Insgesamt basiert die Clusteranalyse somit auf 22 Indikatoren (Tabelle 6).

⁵ Grundlage der Berechnung ist die Darstellung eines Verlaufes als Sequenz von quartalsweiser Abfolge von Zuständen (bsp. ALV,ALV,ALV,Erwerb etc.). Alle empirisch beobachteten Verläufe werden dabei mit einem idealtypischen Verlauf mit durchgängiger Erwerbstätigkeit verglichen. Rechnerisch geschieht der Vergleich über die *Levensthein-Distanz*. Dabei wird gezählt, wie viele Transformationsschritte benötigt werden, bis eine Ausgangssequenz in eine Sequenz von durchgängiger Erwerbstätigkeit überführt ist. Die empirisch beobachtete Möglichkeit von einem Zustand mit Unterstützungsleistungen in den Zustand „Erwerbstätig“ zu gelangen, wird bei der Berechnung mitberücksichtigt. Damit tragen wir dem Umstand Rechnung, dass sich Bezüge bei unterschiedlichen Systemen unterschiedlich „nahe“ am Arbeitsmarkt befinden (ALV ist etwa „näher“ als IV).

Tabelle 8: Übersicht zu den für die Clusteranalyse verwendeten Indikatoren

<i>Verlaufsindikatoren – metrisch</i>	<i>Verlaufsindikatoren - nominal</i>	<i>Komplexität des Verlaufes</i>
Erwerbsintegrationsindex	Zustand zu Beginn	Anzahl Zustände
Quartale keine Sozialleistungen, nicht erwerbstätig	Zustand nach einem Jahr	Anzahl Zustandswechsel
Quartale Erwerbstätig	Zustand nach vier Jahren	
Quartale Erwerbstätig + alle möglichen Kombinationen		
Quartale Erwerbstätig, ALV + weitere Leistungskombi		
Quartale ALV, Sozialhilfe und Erwerb		
Quartale IV-Rente und Erwerb		
Quartale Sozialhilfe und Erwerb		
Quartale IV-Taggelder und Erwerb		
Quartale ALV und Erwerb		
Quartale ALV		
Quartale ALV und Sozialhilfe		
Quartale IV-Taggelder		
Quartale ALV + weitere Leistungskombi		
Quartale Sozialhilfe		
Quartale Leistungskombi (ohne ALV)		
Quartale IV-Rente		

Quelle: Darstellung BFH

4.2.2 Methodische Überlegungen zur verwendeten TwoStep-Clusteranalyse

Es existieren zahlreiche Clusterverfahren mit unterschiedlichen Vor- und Nachteilen. Wir haben uns für die TwoStep-Clusteranalyse entschieden, da es sich um ein Verfahren handelt, das sich explizit für die Anwendung mit grossen Datensätzen eignet, wenn die Zahl der Cluster unbekannt ist bzw. offen gelassen wird. Häufig verwendete Verfahren zur Bestimmung der Clusterzahl, wie die visuelle Analyse mittels Dendrogrammen⁶, sind bei grossen Datenmengen schlicht nicht mehr anwendbar. Die TwoStep-Clusteranalyse gewinnt durch ein vorgeschaltetes Präclusterverfahren an Leistungsfähigkeit. In einer zweiten Stufe wird aus den Präclustern der ersten Stufe in einem hierarchisch-agglomerativen Verfahren ein hierarchisches Mengensystem von Präclustern gebildet, das auch eine Hierarchie der Ausgangsstichprobe ist. Wegen dieser rechnerisch optimierenden Vorgehensweise wird die TwoStep-Clusteranalyse in der Literatur explizit als Verfahren des Data-Mining mit grossen Datensätzen erwähnt (vgl. Wiedenbeck und Züll 2010:534). Die optimale Zahl der Cluster wird dabei als Teil der Analyse mitgeschätzt⁷. Da die Eingangsvariablen sowohl metrische wie nominale Merkmale beinhalten, wurden die Analysen mit dem Log-Likelihood-Distanzmass durchgeführt. Auch wurden alle Merkmale standardisiert, damit die Ergebnisse nicht von den unterschiedlichen Skalen beeinflusst werden. Ferner zeigte sich im Laufe der Analysen, dass diese ohne die Bereinigung von Ausreissern durchgeführt werden können (vgl. Abschnitt 3.7).

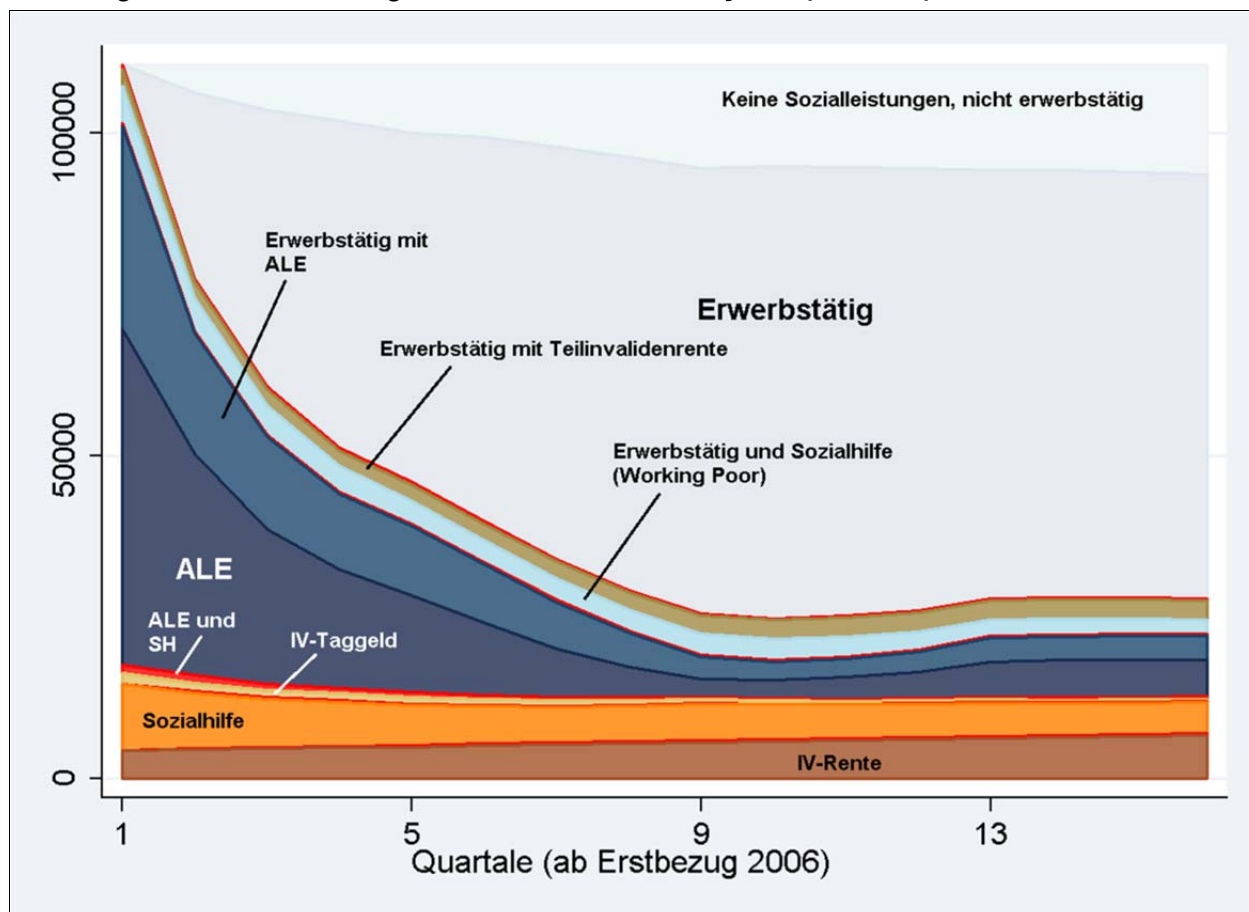
⁶ Ein Dendrogramm ist Teil einer hierarchischen Clusteranalyse. Die Gruppierung der Beobachtungen ist dabei schrittweise entlang der zunehmenden Distanzen abgebildet. Auf diese Weise lässt sich die optimale Zahl der Cluster visuell erkennen.

⁷ Bei den nachfolgend präsentierten Ergebnissen, wurde jeweils die vom Algorithmus favorisierte Zahl der Cluster verwendet. Ausnahme bildet die Clusterung für Personen, die über Sozialhilfe eingetreten sind. Hier wurde die Zahl der Cluster von drei auf vier erhöht, weil das Ergebnis mit vier Clustern inhaltlich besser zu interpretieren ist.

4.2.3 Verlaufstypologie für das Gesamtsystem

Abbildung 3 gibt einen ersten Eindruck zu den beobachteten Verläufen. Gezeigt wird für jedes beobachtete Quartal, wie viele Personen jeweils in einem der festgelegten Zustände sind. Daraus ergibt sich eine aggregierte Zustandsverteilung der Eintrittskohorte 2006 über den Zeitraum von vier Jahren. In der Abbildung lässt sich erkennen, dass die Mehrheit der Eintritte in das Gesamtsystem mit einem ALE-Bezug starten. Der Anteil an Personen mit ALE nimmt mit fortschreitender Zeit rapide ab, während die Zahl der Erwerbstätigen zunimmt. Der Bestand der Personen mit Sozialhilfe bleibt relativ konstant. Bei IV-Renten ist eine leichte Zunahme zu beobachten.

Abbildung 3: Zustandsverteilung von Verläufen im Gesamtsystem (n=110'508)



Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010

Bemerkung: Rote Flächen sind Personenbestände von selten vorkommenden Kombinationszuständen

Aus der Clusteranalyse mit allen Verläufen im Gesamtsystem ergeben sich vier Gruppen, die in Tabelle 7 aufgelistet sind. Die Benennung der Cluster ist aus dem Vergleich der Ausprägungen der Ausgangsvariablen entwickelt, die jeweils für die identifizierten Cluster berechnet wurden. Dabei wurde das Augenmerk auf Verlaufsmerkmale gerichtet, die deutlich zur Unterscheidung der Cluster beigetragen haben (vgl. Anhang 7.2).

Tabelle 9: Verlaufstypologie - Eintritt über alle Systeme

Cluster	Benennung	N	%
1	Mit IV-Rente	11'110	10.1%
2	Existenzsicherung primär durch Sozialhilfe	13'558	12.3%
3	Am Rand des Arbeitsmarktes (ALV, Erwerb und Phasen ohne Erwerb und Leistungen)	42'320	38.3%
4	Kurze Unterstützungsphase zu Beginn, danach Erwerb	43'520	39.4%
Gesamt		110'508	100%

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010

Bemerkung: durchschnittlicher Silhouettenwert=0.3⁸

Cluster 1 – Hauptcharakteristikum dieses Clusters ist der Kontakt mit der IV. Auffällig ist zudem, wie sich die Komplexität der Verläufe stark unterscheidet. Eine Vielzahl (63%) von Verläufen beinhaltet keinen bis einen Wechsel. Gleichzeitig finden sich Verläufe mit der höchsten beobachteten Komplexität (vier bis maximal acht Wechsel). Immerhin 8% der Personen in Cluster 1 sind Verläufe mit erhöhter Komplexität. Die hohe Zahl der Wechsel resultiert dabei aus kurzen Phasen mit abwechselnden Kombinationsbezügen. Über alle Verläufe hinweg betrachtet sind überdurchschnittlich lange IV-Bezüge oder IV-Bezüge mit Erwerb zu beobachten, während Phasen mit Erwerb, Sozialhilfe oder ALV sehr selten vorkommen. Von allen Clustern weist Cluster 4 die höchste mittlere Distanz zur Erwerbstätigkeit auf (~128).

Cluster 2 – Ebenso wie bei Cluster 3 befinden sich in diesem Cluster Personen mit eher komplexen Verläufen mit überdurchschnittlich vielen Wechsel (~4). Hier finden die Wechsel allerdings primär zwischen Sozialhilfe und Erwerbstätigkeit statt, während ALV-Bezüge selten bleiben. Es resultieren lange Bezüge mit Sozialhilfe (~15 Monate), teilweise kombiniert mit Erwerbstätigkeit (~6 Monate), während die mittlere Dauer von Erwerbstätigkeit ohne weitere Leistungsbezüge kurz ausfällt (~3 Monate). Verläufe dieses Clusters beginnen bereits meist in der Sozialhilfe (84%). Nach vier Jahren Beobachtungszeit enden 47% der Verläufe mit Sozialhilfebezug. Unterdurchschnittlich wenige (28%) enden mit einer Phase der Erwerbstätigkeit. Die mittlere Distanz zur Erwerbstätigkeit fällt bei diesem Cluster hoch aus (~118).

Cluster 3 – Verläufe in diesem Cluster zeichnen sich durch überdurchschnittlich viele Wechsel (~4) aus. Diese Wechsel finden überwiegend zwischen ALV und Erwerbstätigkeit statt. Auffällig lange sind Bezüge der ALV (~9 Monate). Vereinzelt können sehr kurze Phasen mit Sozialhilfe vorkommen. Überproportional häufig sind ferner Austritte aus dem Erwerbsleben (kein Erwerb, keine Bezüge). Verläufe starten meist bei der ALV (63%). Nach vier Jahren sind 17% wieder bei der ALV gemeldet, während 33% aus dem Erwerbsleben ausgetreten und immerhin 45% Erwerbstätig sind. Dieses Cluster weist eine mittelhohe Distanz zur Erwerbstätigkeit auf (~64).

⁸ Der Silhouettenwert spricht für eine mittelmässige Qualität des Clusterergebnisses. Das Mass ist ein Durchschnitt aller Datensätze $(B-A) / \max(A,B)$, wobei A der Abstand des Datensatzes zu seinem Clusterzentrum und B der Abstand des Datensatzes zu dem am nächsten liegenden, nicht zugehörigen Clusterzentrum ist. Ein Umrissskoeffizient von 1 würde bedeuten, dass alle Fälle direkt in ihren Clusterzentren liegen. Ein Wert von -1 würde bedeuten, dass alle Fälle in den Clusterzentren anderer Cluster liegen. Ein Wert von 0 bedeutet, dass die Fälle im Durchschnitt gleich weit entfernt von ihrem eigenen Clusterzentrum und dem nächsten benachbarten Cluster liegen.

Cluster 4 - Dieses Cluster zeichnet sich durch Verläufe mit wenigen Wechsel (~2) aus, die meist mit einem kurzen (~3 Monate⁹) ALV-Bezug beginnen und anschliessend in Erwerbstätigkeit übergehen. Nach einem Jahr sind bereits 100% der Personen in diesem Cluster erwerbstätig. Der Verbleib in der Erwerbstätigkeit ist relativ konstant. Personen dieses Clusters sind im Mittel 42 Monate erwerbstätig (von 48 Beobachtungsmonaten). Nach vier Jahren Beobachtungszeit sind 98% erwerbstätig. Über vier Jahre betrachtet resultiert eine geringe mittlere Distanz zur Erwerbstätigkeit (~16).

4.2.4 Verlaufstypologie differenziert nach Eintrittssystemen

Da die Kontaktaufnahme mit einem bestimmten Zweig der Sozialen Sicherung bereits viele Informationen über die Situation der betroffenen Person beinhaltet, ist - je nach Eintrittssystem - mit gänzlich unterschiedlichen Verläufen zu rechnen. Daher wird eine weitere Typologie präsentiert, die aus der Eingrenzung nach Eintrittssystemen¹⁰ resultiert und zwölf Verlaufstypen umfasst (vgl. Tabelle 10).

Tabelle 10: Clustertypologie – differenziert nach Eintrittssystemen

Eintrittssystem	% von allen Verläufen	Cluster	Beschreibung	N	%
ALV	77.2%	1	Komplexe Verläufe mit ALV, Erwerb und Sozialhilfe	5'264	6.2%
		2	ALV und Erwerb wechseln, zwischenzeitliche Austritte aus Erwerb	40'397	47.4%
		3	Kurz ALV mit anschliessender Erwerbsphase	39'601	46.4%
			<i>Gesamt ALV</i>	<i>85'262</i>	<i>100.0%</i>
Sozialhilfe	14.6%	4	Ohne Kontakt zum Arbeitsmarkt (Lange Phasen mit SH und ohne ALV-Bezüge/Erwerb)	5'659	35.1%
		5	Sozialhilfe und IV-Rente	1'756	10.9%
		6	Working Poor (Erwerb und SH)	4'677	29.0%
		7	Kurze Phase mit SH, später erwerbstätig	4'031	25.0%
			<i>Gesamt Sozialhilfe</i>	<i>16'123</i>	<i>100.0%</i>
Invalidenrente	6.2%	8	Einfache Verläufe mit Vollrente	3'246	47.1%
		9	Komplexe Verläufe mit offenem Ausgang (IV-Rente, IV und Erwerb sowie keine Leistung/kein Erwerb)	2'036	29.5%
		10	Sehr komplexe Verläufe mit langen Kombinationsbezügen (IV, IVT und ALV)	1'341	19.5%
		11	Einfache Verläufe, die schnell in Phasen mit IV-Rente und Erwerbstätigkeit münden	271	3.9%
			<i>Gesamt Invalidenrente</i>	<i>6'894</i>	<i>100.0%</i>
IV-Taggeld	2.0%	12	<i>Gesamt IV-Taggeld</i>	<i>2'229</i>	<i>32.3%</i>

⁹ Weil uns die Präsentation der Zeiträume in Monaten verständlicher schien als in Quartalen, haben wir die auf Quartals-Basis berechneten Werte in Monate umgerechnet. Damit sind logischerweise lediglich Ergebnisse in 3-Monatsschritten möglich.

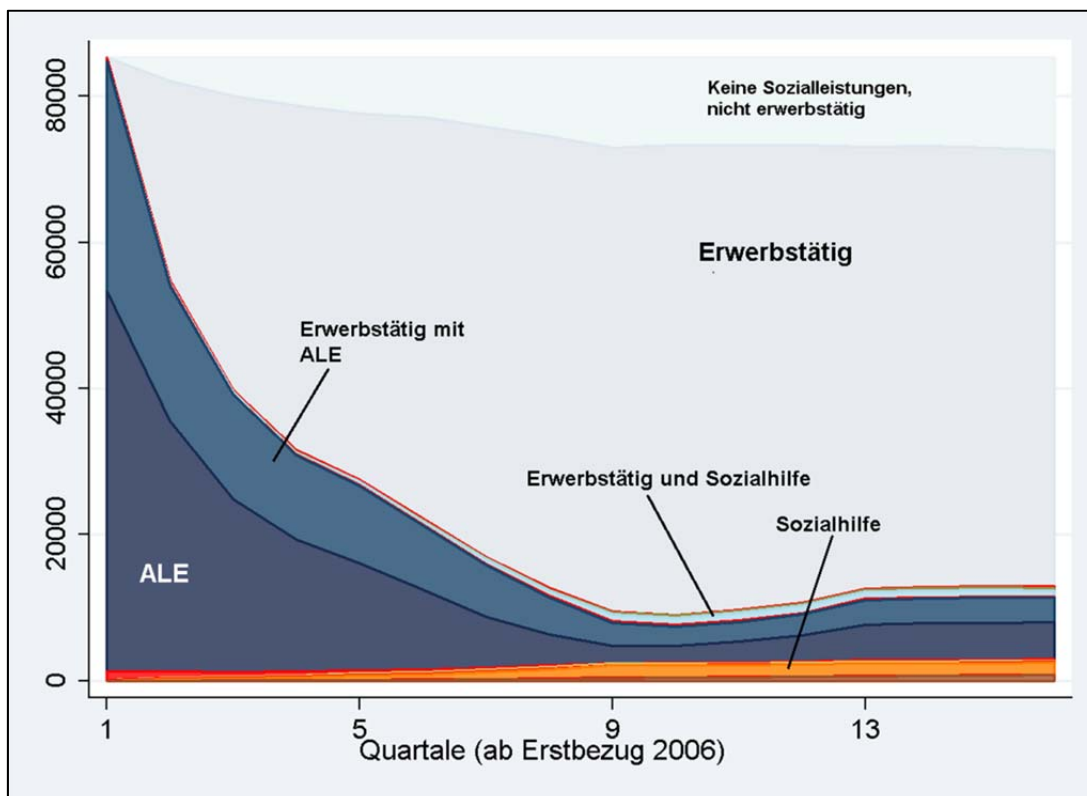
¹⁰ Die fallweise Festlegung des Eintrittssystems ist ausschliessend und in der Reihenfolge ALV, Sozialhilfe, Invalidenrente und IV-Taggeld geschehen. Diese Regel dient dem Umgang mit Kombinationsbezügen bei Eintritt. Konkret bedeutet dies, dass ein Verlauf, der bereits Teil der ALV-Typologie ist, nicht für nachfolgende Clusteranalysen verwendet wird.

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010

Verläufe mit Eintritten über ALE münden meist in die Erwerbstätigkeit, wie dem mit fortschreitender Zeit steigendem Anteil an erwerbstätigen Personen in Abbildung 4 zu entnehmen ist. Markant steigt auch der Anteil an Personen, die weder Leistungen beziehen noch erwerbstätig sind. In wesentlich geringerem Ausmass ist eine Zunahme des Bestandes an Personen in der Sozialhilfe bzw. mit IV Rente auszumachen.

Verläufe, die mit einem Eintritt über ALV beginnen, lassen sich zu drei Clustern gruppieren. Beinahe die Hälfte gehört zu *Cluster 3*, das sich mit Verläufen mit wenig Wechseln (~1), kurzen ALV-Bezügen und nachhaltiger Beschäftigung im Arbeitsmarkt (~42 Monate) charakterisieren lässt. Verläufe in *Cluster 2* sind wesentlich komplexer (~ 3 Wechsel) und mit erschwelter Integration in den Arbeitsmarkt verbunden. Es resultieren überdurchschnittlich lange ALE Bezüge (~ 9 Monate) und ALE-Mehrfachbezüge, die mit unterdurchschnittlicher Erwerbsdauer (~ 24 Monate) einhergehen. Zwischenzeitlich scheinen sich diese Personen nicht am Arbeitsmarkt zu orientieren, verfügen demnach zeitweise über die Möglichkeit ihren Lebensunterhalt ohne Unterstützungsleistungen zu bestreiten. Es resultieren Episoden ohne Erwerb und ohne Leistungsbezüge. Mit dem Ende der vierjährigen Beobachtungsdauer sind 31% der Personen in diesem Cluster im Zustand „Nicht erwerbstätig, kein Bezug“. *Cluster 1* ist vom Arbeitsmarkt am weitesten entfernt (mittlere Distanz=110) und beinhaltet komplexe Verläufe mit vielen Wechseln (~ 4), die auch den Bezug von Sozialhilfe beinhalten. Viele Kombinationsbezüge treten auf. Phasen mit Erwerb kommen vor, bleiben aber selten und sind eher kurz (~ 6 Monate).

Abbildung 4: Zustandsverteilung von Verläufen mit Eintritt über ALV (n=85'262)



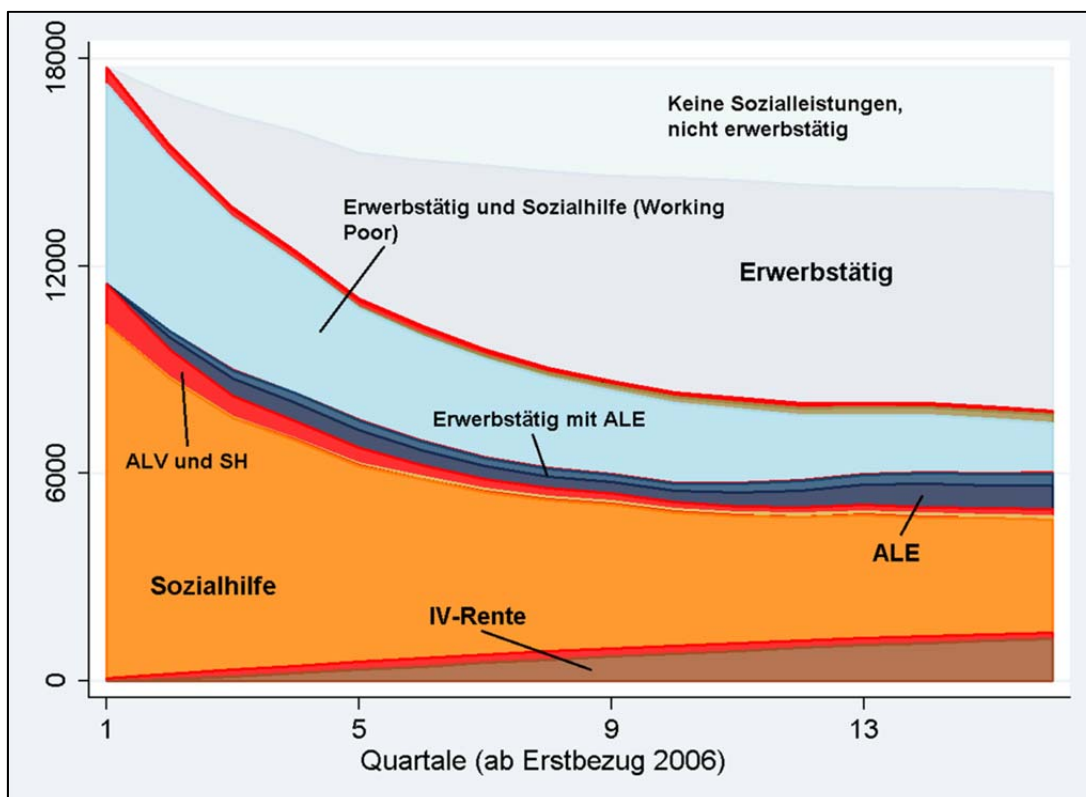
Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010

Bemerkung: Rote Flächen sind Personenbestände von selten vorkommenden Kombinationszuständen

Eintritte über Sozialhilfe erfolgen häufig nur mit Sozialhilfe als finanzielle Grundlage der Lebensbewältigung (vgl. Abbildung 5). Ebenfalls häufig sind Eintritte von Personen, die geläufig als Working Poor bezeichnet werden, d.h. diese Personen verfügen neben der Sozialhilfe über ein Erwerbseinkommen. Mit fortschreitender Zeit nimmt der Bestand an Personen, die auf Sozialhilfe angewiesen sind ab, während die Zahl der Erwerbstätigen, aber auch die Zahl der Personen ohne Leistungen und Erwerb sowie die Zahl der Personen mit IV-Rente zunimmt.

In Abgrenzung zu Verläufen mit Eintritten über andere Systeme sind Verläufe mit Eintritt über die Sozialhilfe im Mittel mit mehr Wechseln (~3) verbunden, was auf eine generell erhöhte Komplexität von Verläufen mit Eintritt über Sozialhilfe hindeutet. Zusammenfassen lassen sich die Verläufe zu vier Clustern. Verläufe in *Cluster 7* befinden sich im Mittel deutlich näher an der Erwerbstätigkeit (~32). Eine Unterstützungsphase mit Sozialhilfe, die teilweise mit Erwerbstätigkeit einhergeht, wird meist nach kurzer Zeit von einer langen Periode mit Erwerbstätigkeit abgelöst (~36 Monate). Auch Verläufe in *Cluster 6* beinhalten Erwerbsphasen. Allerdings sind reine Erwerbsphasen seltener bzw. kürzer (~ 9 Monate), während die Kombination von Erwerb und Sozialhilfe häufiger vorkommt (~ 18 Monate). Deswegen wird dieses Cluster als Working-Poor-Verlauf bezeichnet. Die verbleibenden beiden Cluster zeichnen sich durch eine maximal mögliche Distanz zur Erwerbstätigkeit aus, wobei sich die Art der finanziellen Existenzsicherung deutlich unterscheidet. Verläufe in *Cluster 5* beinhalten eine mittlere Sozialhilfebezugsdauer (~12 Monate), die gegen Ende der Beobachtungszeit meist in eine IV-Rente mündet (81%). Bei Verläufen von *Cluster 4* dauert die Existenzsicherung mittels Sozialhilfe am längsten (~ 24 Monate), überdurchschnittlich lange sind zudem Phasen ohne Erwerb und Bezüge (~ 12 Monate). Wie die Existenzsicherung in diesen Perioden erfolgt, ist unbekannt.

Abbildung 5: Zustandsverteilung von Verläufen mit Eintritt über Sozialhilfe (n=17'764)



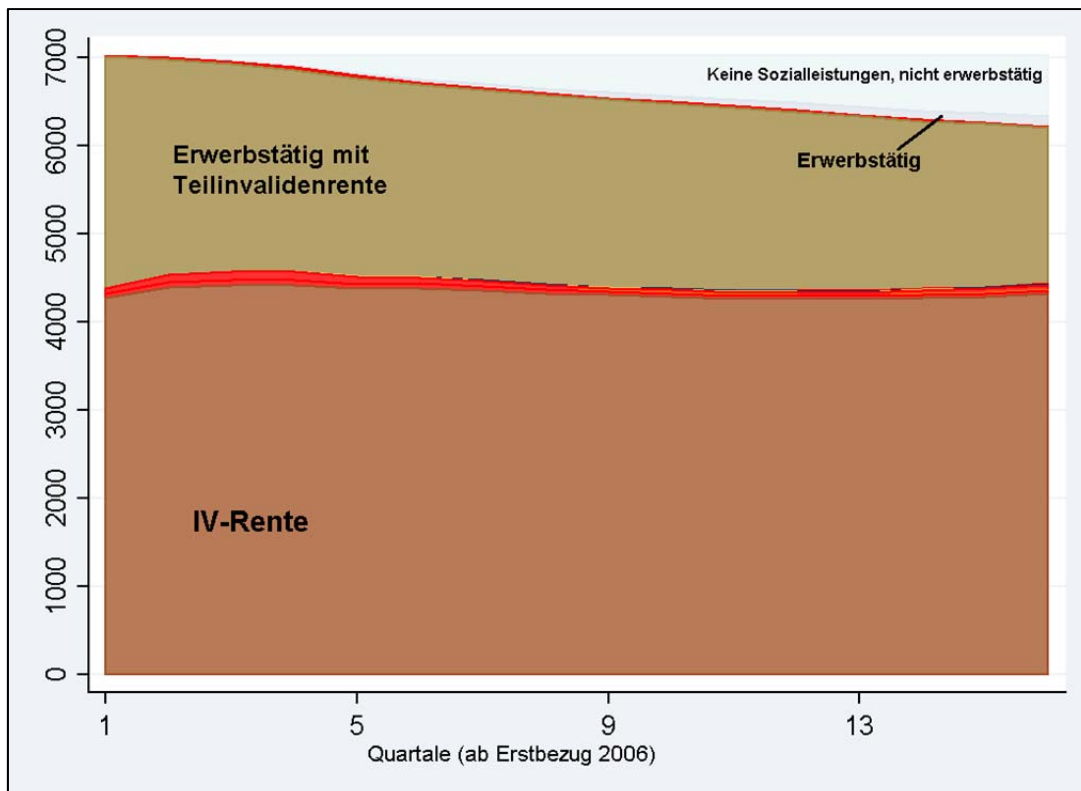
Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010

Bemerkung: Rote Flächen sind Personenbestände von selten vorkommenden Kombinationszuständen

Verläufe mit *Eintritten über IV-Rente* sind über die Zeit gesehen mit deutlich geringeren Zahl an Zustandsverschiebungen verbunden (Abbildung 6). Der Anteil an Personen mit IV-Rente ist relativ konstant, während der Anteil an Personen mit Teilrente und Erwerbstätigkeit in geringem Masse abnimmt. Die Zahl der Personen, die ihr Lebensunterhalt alleine über ein Erwerbseinkommen bestreitet, nimmt allerdings nur geringfügig zu. Dafür steigt der Anteil der Personen die keine Sozialleistungen beziehen und nicht erwerbstätig sind.

Laut Silhouetten-Kohäsionswert erlaubt die Clusterung mit Eingrenzung auf Eintritte über IV-Rente eine deutliche Unterteilung in vier Gruppen. Dabei spielt hier die Nähe zur Erwerbstätigkeit eine untergeordnete Rolle. Vielmehr sind es Merkmale zur Komplexität des Verlaufes und zur Dauer von Renten- sowie Kombinationsbezügen, die für die Differenzierung ausschlaggebend sind. Verläufe in *Cluster 8* lassen sich als einfache Verläufe (ohne Wechsel) beschreiben, die mit einer IV-Rente beginnen und enden (100%). Ebenfalls als einfach (ohne Wechsel) bezeichnet werden können Verläufe in *Cluster 11*. Diese beginnen und enden mit einer IV-Teilrente, die mit einem Erwerbseinkommen ergänzt wird. Mehr Veränderung ist bei Verläufen zu beobachten (~ 1 Wechsel), die in *Cluster 9* zusammengefasst werden. Dabei sind am Ende des Verlaufes unterschiedliche Zustände zu beobachten; am häufigsten sind IV-Vollrenten (44%), am nächst häufigsten sind Zustände ohne Erwerb und ohne Leistung/Rente (33%). Bei diesen Personen ist eine Rente also durch eine unbekannte Quelle der Existenzsicherung ersetzt worden (möglicherweise Partnerin oder Partner). Auch möglich sind IV-Renten kombiniert mit Erwerbseinkommen (16%). Manchmal steht am Ende des Verlaufes eine Periode mit Erwerb ohne weitere Leistungsbezüge (5%). Bei Verläufen in *Cluster 10* werden schliesslich am meisten Wechsel beobachtet (~ 2). Diese finden meist von einer IV-Rente zu einer IV-Rente mit Erwerb statt (oder umgekehrt). Hinzu kommen Phasen mit Leistungskombinationsbezügen (IV-Rente plus SH und/oder IV-Taggeld).

Abbildung 6: Zustandsverteilung von Verläufen mit Eintritt über IV (n=7'024)



Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010

Bemerkung: Rote Flächen sind Personenbestände von selten vorkommenden Kombinationszuständen

Dem letzten *Cluster 12* zugeordnet werden Verläufe, die mit einem *Eintritt mit IV-Taggeldern* beginnen. Da Verläufe von Eintritten mit IV-Taggeldern ohne spätere Bezüge von SH, ALV oder IV fehlen, bilden Auswertungen zu dieser Teilgruppe die Realität nicht adäquat ab (vgl. Kapitel 2). Auf eine weitere Differenzierung wird daher verzichtet.

4.3 Risikoprofile

Inhaltlich sind Risikoprofile eine Verknüpfung der Verlaufstypologie und der Merkmale der Personen bzw. Merkmale zu deren Umfeld (Kanton, Gemeinde, Haushalt). Es soll daraus ersichtlich werden, ob Konstellationen in der Situation der Personen zu beobachten sind, die bestimmte Verläufe begünstigen. Methodisch entscheiden wir uns hier für einen Regressionsansatz, über den Zusammenhänge eindeutig modelliert werden können (Festlegen von Zielgrösse und erklärenden Merkmalen). Gleichzeitig kann auf diese Weise der Einfluss einzelner Merkmale unter Konstanthaltung aller im Modell berücksichtigten Variablen berechnet werden, was Aussagen zur prädiktiven Stärke einzelner Merkmale zulässt. Dabei werden diese Voraussagen nicht als kausale Zusammenhänge interpretiert, sondern als Indikatoren auf Risiken bestimmter Verlaufstypen innerhalb des Systems der Sozialen Sicherheit. Ebenfalls ist keine Aussage über die Konstellationen von Merkmalen möglich, die zum Eintritt in das System der Sozialen Sicherheit führen.

4.3.1 Datengestützte Risikoprofile als Ergebnis einer Regressionsanalyse

Das Vorliegen von Personenmerkmalen ist auch davon abhängig, mit welchem System der sozialen Sicherung jemand in Kontakt gekommen ist, da bei Personen, die AVL, IV oder Sozialhilfe beziehen, je unterschiedliche Merkmale erhoben werden. Es bietet sich daher an, die Analysen nach Eintrittssystemen zu differenzieren. Zielvariable der Regressionsschätzung bilden jeweils die bereits vorgestellten Verlaufstypologien. Da es sich dabei um eine kategoriale Variable handelt, drängt es sich auf, Zusammenhänge über eine multinomial logistische Regression zu modellieren.

Die Risikoprofile werden anhand der geschätzten Zusammenhangsmodelle erstellt. Risikoprofile werden schrittweise vorgestellt, indem zuerst die *Risiko-* und *Resilienzmerkmale* identifiziert und darauf aufbauend Risikogruppen bestimmt werden.

- Im ersten Schritt werden Merkmale isoliert betrachtet, d.h. es werden Merkmalsausprägungen hervorgehoben, bei welchen sich signifikante Unterschiede bei der Zuteilung zu den Verlaufstypen gezeigt haben. Unter *Risikofaktoren* werden Ausprägungen geführt, die die Wahrscheinlichkeit zum Zieltyp zu gehören erhöhen, entsprechend sind unter *Resilienzmerkmale* jene Ausprägungen geführt, die die Wahrscheinlichkeit erhöhen, nicht zum Zieltyp zu gehören (sondern zum Referenztyp). Dieser Logik folgend ist die Festlegung von Ziel- und Referenztyp entscheidend. Darauf wird später eingegangen.
Zur Beurteilung der Bedeutsamkeit eines bestimmten Merkmals werden Faktorwerte ausgewiesen. Bei den dargestellten Risiko- bzw. Resilienzfaktorwerten handelt es sich um exponierte Beta-Koeffizienten aus dem Schätzmodell, auch bekannt als Odds-Ratio. Interpretiert werden kann der Wert als Faktor, um den sich das Verhältnis von Wahrscheinlichkeit zu Gegenwahrscheinlichkeit verändert, zu einem entsprechenden Verlaufstyp (Ziel bzw. Referenztyp) zu gehören, wenn die ausgewiesene Kategorie mit einer festgelegten Referenzkategorie verglichen wird. Als Referenzkategorie wird jeweils die Kategorie mit mittlerer Effektstärke gewählt. Damit kann die ausgewiesene Effektstärke als Veränderung zum Durchschnitt betrachtet werden.
- Im zweiten Schritt werden Merkmalskombinationen betrachtet, um damit *Risikogruppen* bestimmen zu können. Ausgehend von der Durchschnittsgruppe (Personen mit durchschnittlicher Ausprägung auf relevanten Merkmalen) werden Risiko- und Resilienzgruppen bestimmt, die entsprechend höhere bzw. tiefere Risiken haben, ungünstige Verläufe zu erfahren. Die Darstellung erfolgt hierbei über die gruppenweise vorhergesagte Wahrscheinlichkeit der Zugehörigkeit zu einem bestimmten Verlaufstyp.

Um einen Überblick über die Aussagekraft der berechneten Modelle zu erhalten, wurden in Tabelle 11 verschiedene Gütemasse der Modellanpassung für die einzelnen Systeme (ALV, SH und IV) wie auch für das Gesamtsystem berechnet. Dabei zeigt sich, dass Eintritte über die IV am besten vorausgesagt werden können. Die verschiedenen Eintritte über die ALV und über die SH werden von den verwendeten Modellen nahezu gleich gut geschätzt. Einzig im McFadden R2 zeigt sich eine etwas schlechtere Schätzung bei der SH als bei der ALV. Die Werte für das Gesamtsystem sind schliesslich leicht höher als für die ALV und die SH, liegen aber weit unter derjenigen für die IV. Insgesamt betrachtet sind die Gütemasse ausser bei der IV auf einem relativ tiefen Niveau.

Dabei gilt es jedoch zu berücksichtigen, dass je nach Eintrittssystem aufgrund der erhobenen Registerdaten unterschiedliche Merkmale zur Situation bei Eintritt zur Verfügung stehen. So ist z.B. der Ausbildungsstand der Person nur für das Eintrittssystem ALV mit genügender Qualität vorhanden. Die Ausprägungen der verwendeten Variablen werden in den Tabelle 5 und Tabelle 32 beschrieben.

Tabelle 11: Güte der Modellanpassung für alle Modelle

	ALV	SH	IV	Gesamt
Cox und Snell	0.11	0.11	0.28	0.15
Nagelkerke	0.13	0.12	0.31	0.17
McFadden	0.07	0.04	0.14	0.07

Quelle: Modellschätzungen anhand SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010

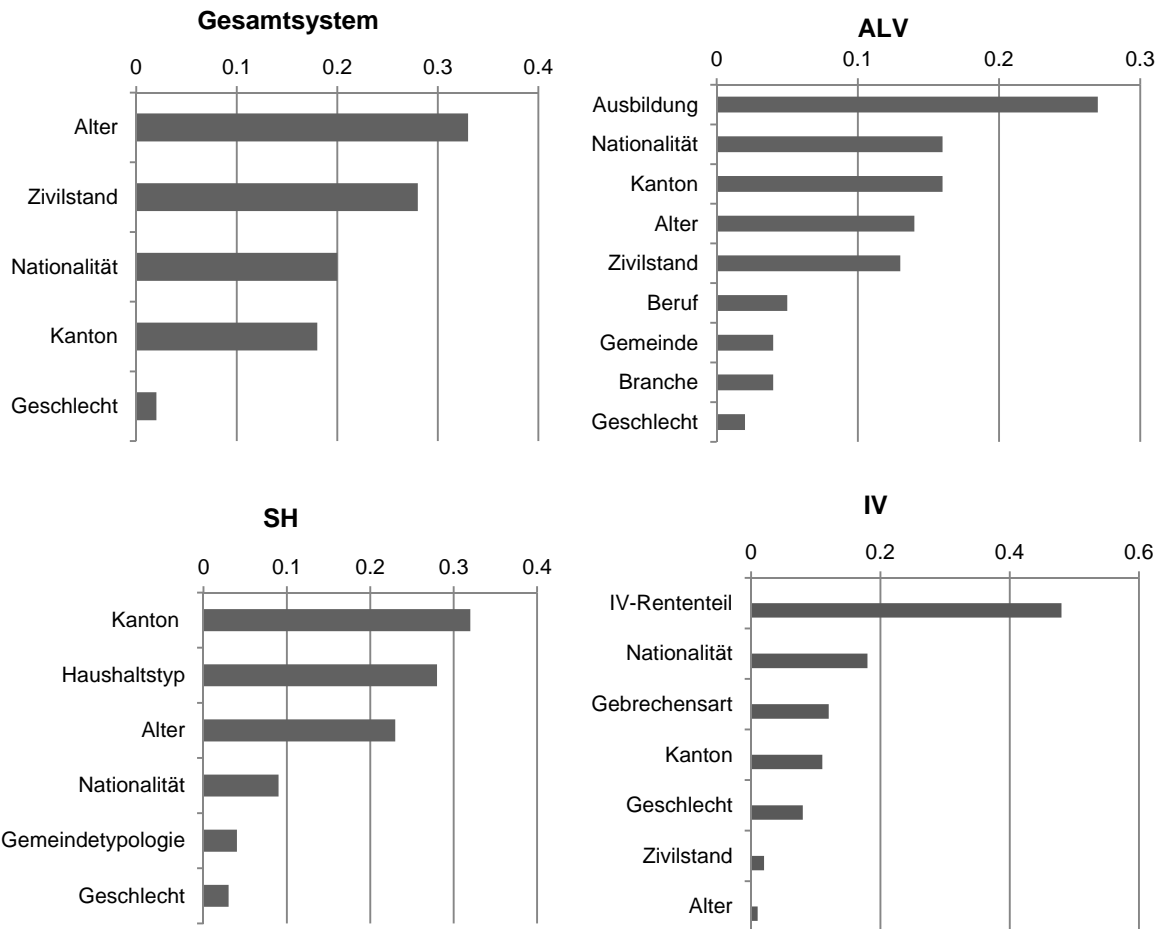
In einem nächsten Schritt wurde für die vier Modelle berechnet, wie gross der Beitrag der einzelnen Variablen zu Erklärungsleistung des Gesamtmodells ist. In Abbildung 7 sind die Bedeutsamkeiten der einzelnen Variablen für alle Modelle in hierarchischer Reihenfolge aufgelistet. Für das Gesamtmodell zeigt sich, dass vor allem das Alter und der Zivilstand entscheidende Faktoren sind, die bestimmen, wie Personen mit den Systemen der sozialen Sicherheit in Kontakt kommen.

Für die Zuordnung zur ALV Typologie spielt hingegen vor allem der Bildungsstand eine entscheidende Rolle.

Die zentralste Variable für die Vorhersage des Verlaufes in der Sozialhilfe bildet interessanterweise die regionale Verortung in einem Kanton, dabei stellt sich die Frage, ob dieser Zusammenhang auf besonderen Zusammensetzungen der Bevölkerung oder auf der Behördenpraxis beruht. Diese Frage kann jedoch im Rahmen dieses Berichtes nicht genauer beantwortet werden und muss Gegenstand von späteren Untersuchungen sein. Daneben spielen auch der Haushaltstyp und das Alter eine wichtige Rolle beim Modell der Sozialhilfe.

Schliesslich bildet für die IV-Typologie der IV-Rententeil mit Abstand den wichtigsten Faktor im Modell. Insgesamt kann gesagt werden, dass in allen Modellen äusserst verschiedene Grössen eine zentrale Rolle spielen. Allerdings ist nochmals zu bemerken, dass nicht bei allen Eintrittssystemen dieselben Merkmale in die Analyse mit einbezogen werden konnten.

Abbildung 7: Bedeutsamkeit der Prädikatoren



Quelle: Modellschätzungen anhand SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010

4.3.2 Risikoprofile für Eintritte über ALV

Im Folgenden wird nun genauer auf die einzelnen Modelle eingegangen. Dabei werden verschiedene Auswertungen bezüglich der Risikoprofile in den Systemen der sozialen Sicherheit dargestellt. Dabei soll in einem ersten Schritt genauer auf Verläufe in der ALV eingegangen werden.

Zielvariable ist die ALV-Typologie:

- Typ 1: „Komplexe Verläufe mit ALV, Erwerb und Sozialhilfe“
- Typ 2: „Phasen mit ALV und Erwerb wechseln sich ab“
- Typ 3: „Kurzer ALV-Bezug mit anschliessender Erwerbsphase“

Die ersten beiden Verlaufstypen beinhalten eher prekäre Verläufe, während Typ 3 jene Verläufe zusammenfasst, die nach absehbarer Zeit in eine beständige Phase von Erwerbstätigkeit münden. Da Typ 3 demnach eher erwünschte Verlaufsmuster umfasst, wird dieser Typ als Referenzkategorie verwendet.

Tabelle 12 und Tabelle 13 beinhalten die *Risiko- und Resilienzmerkmale* mit Fokus auf komplexe Verläufe (Typ 1) und Verläufe mit erschwertem Wiedereintritt in den Arbeitsmarkt (Typ 2). Die Merkmale mit prädiktiven Eigenschaften sind dabei in absteigender Reihenfolge entsprechend der Höhe des Risiko- bzw. Resilienzfaktors sortiert. Diese Werte widerspiegeln gleichzeitig die Prädiktorstärke der entsprechenden Merkmalausprägungen, was Vergleiche ermöglicht. Angegeben ist die Effektstärke auf das Odds Ratio (Exp(B)). Nicht signifikante Effekte werden mit ‚-‘ angegeben.

Ein erhöhtes Risiko zur Gruppe der Personen mit komplexen Verläufen zu gehören weisen Personen mit geringer Ausbildung auf (kein nachobligatorischer Abschluss). Auch die Nationalität spielt eine Rolle; Ausländerinnen und Ausländer der EU-8 Länder gehören zur Risikogruppe. Im kantonalen Vergleich zeigt sich, dass insbesondere Personen in Basel-Stadt bzw. in Waadt mit komplexeren Verläufen behaftet sind. Auch sind Personen, die in urbanem Gebiet leben, eher von komplexen Verläufen betroffen. Demographische Faktoren wie Zivilstand, Alter und Geschlecht gehen ebenfalls mit erhöhtem Risiko einher. So neigen Personen, die ihren Partner bzw. ihre Partnerin verloren haben, zu komplexeren Verläufen. Dies lässt sich nicht über das Alter erklären, da dies im Modell kontrolliert wird. Es zeigt sich aber, dass Personen im letzten Erwerbsaltersabschnitt ebenfalls einem erhöhten Risiko zu ungünstigen Verläufen ausgesetzt sind. Keine auffälligen Muster liessen sich bezüglich Beruf und Branche erkennen. Allerdings wurde hier mit verdichteten Kategorisierungen gearbeitet. Allenfalls ergäbe eine Auswertung der detaillierten Codes Hinweise auf bestimmte riskante Berufsgruppen.

Tabelle 12: Typ 1 "Komplexe Verläufe mit ALV, Erwerb- und Sozialhilfe"

Risikomerkmale	Risikofaktor	Resilienzmerkmale	Resilienzfaktor
Ausbildung: weniger als obligatorischer Schulabschl.	2.9	Kanton: Nidwalden	3.6
Ausbildung: Sek 1	2.8	Ausbildung: Tertiär	2.8
Nationalität: EU 8	2.3	Nationalität: EU 17 Nord	2.3
Kanton: Basel-Stadt	2.1	Nationalität: EU 17 Süd	2.2
Kanton: Waadt	2.0	Beruf: Gastgewerbe	1.8
Zivilstand: : Verwitwet, geschieden, getrennt	2.0	Geschlecht: Frau	1.4
Gemeindetypologie: Zentren	1.7	Zivilstand: verheiratet	1.4
Alter: 45 bis 65	1.6	Alter: Unter 23 Jahre	1.3
Geschlecht: Mann	1.4	Gemeindetypologie: Agrarische Gemeinden	1.2
Beruf	-	Branche	-
Branche	-		

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010

Bemerkungen: Referenzkategorien der Risiko- bzw. Resilienzmerkmale sind:

- Ausbildung: Sekundarstufe 2
- Kanton: Jura
- Nationalität: Schweiz
- Zivilstand: ledig
- Gemeindetypologie: Periurbane Gemeinden
- Alter: 26 bis 45 Jahre
- Beruf: Sonst Dienstleistungen

Es können weitere Faktoren gefunden werden, die eher vor riskanten Verläufen „schützen“, sogenannte Resilienzfaktoren. Auffällig ist der kantonale Effekt von Nidwalden, der sich anhand der bestehenden Daten nur schwer erklären lässt. Erstaunlich ist, dass Personen aus EU 17 Nord und Süd-Ländern noch seltener komplexe Verläufe aufweisen als Schweizerinnen und Schweizer. Zusätzlich verringert wird das Risiko, wenn diese Person in einer agrarischen Gemeinde wohnt. Ausserdem sind verheiratete Personen mit guter Bildung weniger betroffen.

Anstelle der Sozialhilfe (Typ 1) stehen bei Verläufen von Typ 2 eher Episoden ohne Erwerb und Leistungen im Vordergrund. Zwischenzeitlich scheiden diese Personen aus dem Arbeitsmarkt aus. Alles in allem sind diese Verläufe etwas weniger kritisch zu beurteilen und verlaufen etwas näher am Arbeitsmarkt als Typ 1. Entsprechend sind Risiko- und Resilienzfaktoren weniger ausgeprägt. Erstaunlicherweise sind kantonale Unterschiede am grössten, während die Ausbildung auch hier eine relevante Rolle spielt. Gleichzeitig erleben eher Jugendliche und junge Erwachsene solche Verläufe. Häufiger sind es Frauen, allerdings ist die Zuordnung zu Verläufen über das Geschlecht verzerrt, weil Frauen im nachfolgenden Sozialhilfebezug häufig nicht in den Daten erfasst sind, da der Fallträger in einem Paarhaushalt oft der Mann ist.

Tabelle 13: Typ 2 "ALV und Erwerb abwechselnd, mit Austritten aus Erwerb"

Risikomerkmale	Risikofaktor	Resilienzmerkmale	Resilienzfaktor
Kanton: Genf	1.9	Alter: 45 bis 65 Jahre	1.7
Kanton: Tessin	1.8	Nationalität: Schweiz	1.6
Ausbildung: Sek I1	1.8	Kanton: Obwalden	1.5
Ausbildung: weniger als obligatorische Ausbildung	1.6	Gemeindetypologie: Agrarische Gemeinden	1.2
Nationalität: EU 8	1.3	Zivilstand: ledig	1.2
Alter: 18 bis 25 Jahre	1.3	Ausbildung: Tertiär	1.1
Gemeindetypologie: Zentren	1.1	Geschlecht: Mann	1.1
Geschlecht: Frau	1.1	Beruf	-
Zivilstand	-	Branche	-
Beruf	-		
Branche	-		

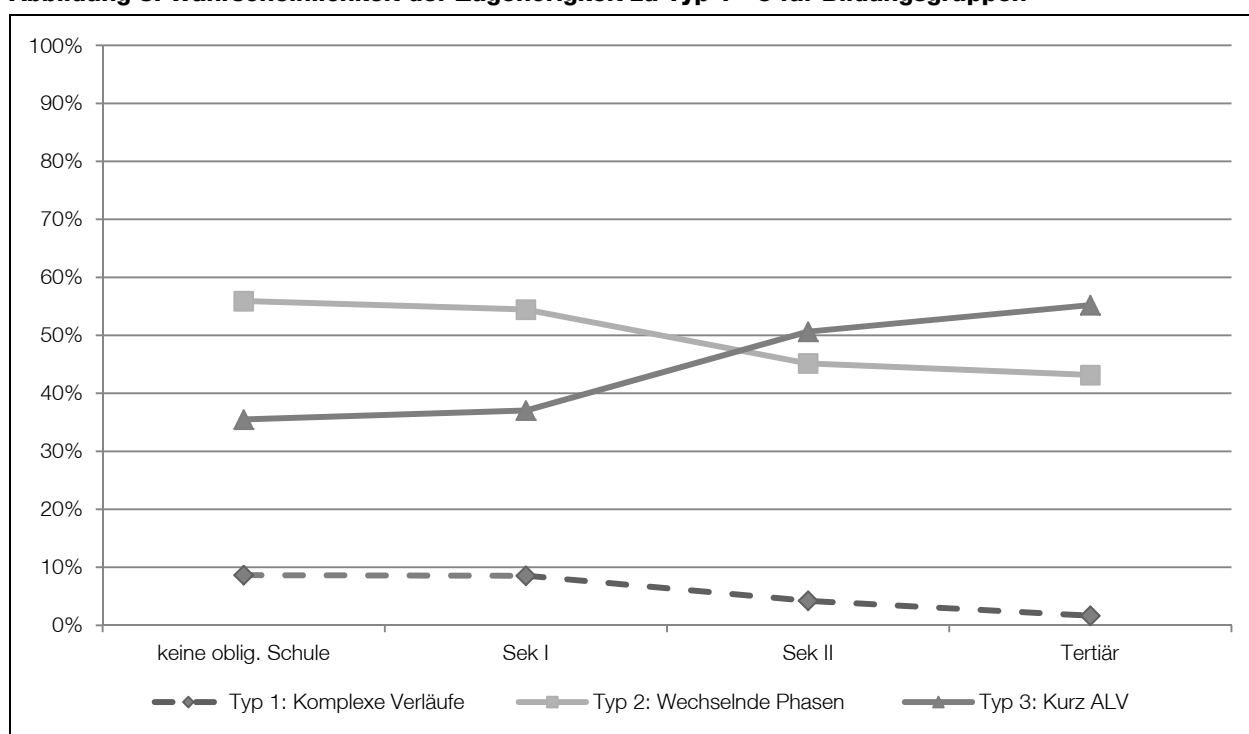
Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010

Bemerkungen: Referenzkategorien der Risiko- bzw. Resilienzmerkmale sind:

- Ausbildung: Sekundarstufe 2
- Kanton: Graubünden
- Nationalität: EU 2
- Zivilstand: verheiratet
- Gemeindetypologie: Industrielle und tertiäre Gemeinden
- Alter: 26 bis 45 Jahre

Ein besseres Verständnis für die modellierten Zusammenhänge kann gewonnen werden, wenn anhand der Modellparameter die geschätzten Wahrscheinlichkeiten der Zugehörigkeiten zu den jeweiligen Verlaufstypen ausgewiesen werden. In Abbildung 8 wurde dies für die Differenzierung entlang den höchsten Bildungsabschluss vorgenommen. Der Vergleich der drei Verlaufstypen veranschaulicht, dass die Wahrscheinlichkeit der Zugehörigkeit zu Typ 1 – Komplexe Verläufe gesamthaft gesehen deutlich geringer ist, als jene für Typ 2 und 3. Gleichzeitig steigt die Wahrscheinlichkeit von unproblematischen Verläufen mit zunehmendem Bildungsabschluss, während die Wahrscheinlichkeit für riskante Verläufe abnimmt. Auffallend ist die sprunghafte Veränderung bei Personen mit Berufsabschluss (Sek. II) im Vergleich zu Personen ohne Berufsabschluss (Sek. I und kein oblig. Schulabschluss.)

Abbildung 8: Wahrscheinlichkeit der Zugehörigkeit zu Typ 1 - 3 für Bildungsgruppen



Quelle: Modellschätzungen anhand SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010

In einem abschliessenden Schritt werden einzelne *Risikogruppen* eruiert. Anhand der Modellparameter werden Konstellationen ausgemacht, die ein klar erhöhtes bzw. verringertes Risiko von gewissen Verläufen aufweisen. Dabei werden grundsätzlich Kumulationen von mehreren Risiko- bzw. Resilienzmerkmalen gemäss den vorangehenden Analysen betrachtet. Dabei wurde die Beschreibung der Personengruppen auf persönliche Merkmale beschränkt, d.h. Gemeindetypologie und Wohnkanton wurden als erklärende Merkmale weggelassen, da sich diese auf den Kontext beziehen.

Dabei wird ersichtlich, dass die Kumulation mehrerer Risikomerkmale wie geringer Bildung, Herkunft aus einem EU8-Land, der Verlust des Lebenspartners und fortgeschrittenes Alter zu einer deutlichen Veränderung der vorhergesagten Risikowahrscheinlichkeit führen. Bei Personen, die obiger Beschreibung entsprechen, ist die Wahrscheinlichkeit komplexe Verläufe zu erfahren doppelt so hoch, wie die Wahrscheinlichkeit von Verläufen mit schneller Reintegration in den Arbeitsmarkt (vgl. Tabelle 14). Im Vergleich zu Personen mit durchschnittlichem Profil ist die Wahrscheinlichkeit von riskanten Verläufen demnach rund 3.5-mal höher.

Auch Schweizerinnen und Schweizer derselben Altersgruppe können als Risikogruppe bezeichnet werden, weil die Wahrscheinlichkeit der Zugehörigkeit zu Verlaufstyp 3 – Erwerb deutlich geringer als im Durchschnitt ausfällt. Junge Erwachsene (zwischen 18 und 25 Jahre) aus einem EU-8 Land und ohne Berufsabschluss können als letzte Risikogruppe ausgemacht werden.

Als Gruppe die im Vergleich zum Durchschnitt günstigere Wahrscheinlichkeiten der Zugehörigkeiten zu unproblematischen Verläufen aufweist, können junge Erwachsene (unter 23 Jahre) mit einem Abschluss auf Sekundarstufe II ausgemacht werden. Diese Gruppe hat ein sehr geringes Risiko riskante Verläufe zu erleben und gleichzeitig eine etwas erhöhte Chance für Verläufe mit Aufnahme einer Erwerbstätigkeit nach kurzer Zeit. Wie sich allerdings den Werten in Tabelle 14 entnehmen lässt, sind die Abweichungen zu Durchschnittsbeziehenden eher gering.

Tabelle 14: Vorhergesagte Wahrscheinlichkeiten für ausgewählte Gruppen

	Typ 1: Komplex	Typ 2: Wechselhaft	Typ 3: Erwerb	Gesamt
Gruppe				
kein obligatorischer Schulabschluss aus EU8, geschieden/verwitwet/getrennt, 46 - 65 Jahre	24.7%	62.4%	12.9%	100%
Abschluss auf Sekundarstufe I, aus EU8, zwischen 18 und 25 Jahren	13.4%	56.9%	29.6%	100%
zwischen 45 und 65 Jahre, Schweizer-/in, ledig	8.6%	50.6%	40.8%	100%
<i>Profil von Durchschnittsbeziehenden</i>				
lediger Mann, zwischen 26 und 45 Jahren, Schweizer, Gastgewerbe, mit Sekundarstufen II Abschluss.	7.6%	37.6%	54.7%	100%
Abschluss auf Sekundarstufe II, unter 23 Jahre	4.1%	38.1%	57.8%	100%

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010

Anmerkungen: Wahrscheinlichkeiten mit erhöhter Zugehörigkeit zu einem bestimmten Verlaufstyp sind **fett** markiert.

4.3.3 Risikoprofile für Eintritte über SH

Im Folgenden werden nun analog zu der vorhergehenden Analyse der Eintritte über die ALV die Risikoprofile für Eintritte über die Sozialhilfe näher betrachtet. Als Zielvariable dient hierzu die SH-Typologie, die folgende Ausprägungen besitzt:

- Typ 1: "Ohne Kontakt zum Arbeitsmarkt (Lange Phasen mit SH und ohne ALV-Bezüge/Erwerb)"
- Typ 2: "Sozialhilfe und IV-Rente"
- Typ 3: "Working Poor (Erwerb und SH)"
- Typ 4: „Kurze Phase mit SH, später erwerbstätig“

Während die ersten drei Typen eher problematische Verläufe in der Sozialhilfe beschreiben, stellt Typ 4 einen kurzen und folgenlosen also relativ unproblematischen Kontakt mit der Sozialhilfe dar. Aus diesem Grund wird, wie schon bei den Risikoprofilen zur ALV, bei den Auswertungen Typ 4 als Referenzkategorie verwendet.

In **Tabelle 15**, **Tabelle 16** und **Tabelle 17** wurden nun für die ersten drei Typen des Sozialhilfebezuges Risiko- und Resilienzmerkmale abgetragen. Der Fokus liegt dabei auf Verläufen ohne Kontakt zum Arbeitsmarkt (Typ 1), Verläufen mit Sozialhilfe und IV-Rente (Typ 3) und Konstellationen, die als Working Poor charakterisiert werden können (Typ 3). Die Merkmale mit prädiktiven Eigenschaften sind dabei in absteigender Reihenfolge entsprechend der Höhe des Risiko- bzw. Resilienzfaktors sortiert. Diese Werte widerspiegeln gleichzeitig die Prädiktorstärke der entsprechenden Merkmalausprägungen, womit Vergleiche möglich sind.

Für Personen ohne Kontakt zum Arbeitsmarkt (Typ 1) zeigt sich, dass der Wohnkanton eine entscheidende Rolle beim Verlauf des Sozialhilfebezugs spielt. So haben Personen aus den Kantonen Genf, Tessin, Neuenburg und Waadt ein erhöhtes Risiko für lange Phasen des Sozialhilfebezugs ohne Kontakt zum Arbeitsmarkt. Daneben sind auch Personen zwischen 46 und 65 also in der zweiten Hälfte ihres Erwerbslebens, eher vom Arbeitsmarkt ausgeschlossen und an die Sozialhilfe gebunden.

Auch bei den Resilienzmerkmalen für Personen ohne Kontakt zum Arbeitsmarkt bildet die kantonale Verortung eine zentrale Grösse. So ist es für Personen aus dem Kanton Wallis relativ unwahrscheinlich, dass sie bei einem Sozialhilfebezug dauerhaft aus dem Arbeitsmarkt ausscheiden. Ähnliches gilt auch für jüngere Personen unter 25 Jahren.

Tabelle 15: Typ 1 "Ohne Kontakt zum Arbeitsmarkt (Lange Phasen mit SH, ohne ALV-Bezüge/Erwerb)"

Risikomerkmale	Risikofaktor	Resilienzmerkmale	Resilienzfaktor
Kanton: Genf	2.3	Kanton: Wallis	1.5
Kanton: Tessin	2.3	Alter: 23 bis 25	1.5
Alter: 46 bis 65	1.7	Alter: unter 23 Jahre	1.2
Kanton: Neuenburg	1.5	Geschlecht:	-
Kanton: Waadt	1.4	Nationalität:	-
Geschlecht:	-	Gemeindetypologie:	-
Haushaltstyp:	-		
Nationalität:	-		
Gemeindetypologie:	-		

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010

Bemerkungen: Referenzkategorien der Risiko- bzw. Resilienzmerkmale sind:

Geschlecht: Männlich

Haushaltstyp: Andere

Kanton: St.Gallen

Alter: 26 bis 45

Nationalität: EU2

Gemeindetypologie: Industrielle und tertiäre Gemeinden

Das wichtigste Risikomerkmale dafür, neben dem Sozialhilfebezug gleichzeitig auch eine IV-Rente zu beziehen, bildet das Alter (vgl. **Tabelle 16**). So sind Personen im Alter zwischen 46 und 65 Jahren überdurchschnittlich oft von dieser Konstellation betroffen. Daneben bildet der Wohnkanton wiederum ein zentrales Risikomerkmale dafür, ob jemand mit Sozialhilfe gleichzeitig eine IV-Rente bezieht. So sind Personen aus dem Tessin oder aus St.Gallen überdurchschnittlich oft von dieser Leistungskombination betroffen. Zudem steigt auch bei Personen, die alleine leben, das Risiko eines SH-Bezugs bei gleichzeitiger IV-Rente.

Alleinerziehende Personen weisen hingegen nur selten gleichzeitig neben einem Sozialhilfebezug auch eine IV-Rente auf. Ein weiterer wichtiger Faktor, der die Wahrscheinlichkeit eines Kombibezugs von Sozialhilfe und IV-Rente verringert, bildet der Wohnkanton. So sind Personen aus dem Graubünden, aus Zürich und aus Waadt relativ selten in der genannten Situation. Schliesslich sind, wie auch beim Typ 1, jüngere Personen unter 25 kaum von der dargestellten Situation betroffen.

Tabelle 16: Typ 2 "Sozialhilfe und IV-Rente"

Risikomerkmale	Risikofaktor	Resilienzmerkmale	Resilienzfaktor
Alter: 46 bis 65	2.6	Haushaltstyp: Alleinerziehende mit 3+ Kinder	2.8
Kanton: Tessin	2.0	Haushaltstyp: Alleinerziehende mit 2 Kinder	2.5
Kanton: St. Gallen	1.7	Kanton: Graubünden	2.4
Haushaltstyp: Alleinlebend	1.3	Haushaltstyp: Alleinerziehend mit 1 Kind	2.0
Geschlecht:	-	Alter: 23 bis 25	1.9
Nationalität:	-	Alter: Unter 23 Jahre	1.5
Gemeindetypologie:	-	Kanton: Zürich	1.4
		Kanton: Waadt	1.3
		Geschlecht:	-
		Nationalität:	-
		Gemeindetypologie:	-

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010

Bemerkungen: Referenzkategorien der Risiko- bzw. Resilienzmerkmale sind:

Geschlecht: Männlich

Haushaltstyp: Nicht-Alleinlebende

Kanton: Aargau

Alter: 26 bis 45

Nationalität: EU17 angrenzend

Gemeindetypologie: suburbane Gemeinden

In **Tabelle 17** werden schliesslich die Risiko- und Resilienzmerkmale für Personen die neben dem Sozialhilfebezug gleichzeitig einer Erwerbsarbeit nachgehen (Typ 3: Working Poor) ausgewiesen. Hierbei zeigt sich, dass vor allem alleinerziehende Personen mit einem oder mehreren Kindern ein erhöhtes Working Poor-Risiko aufweisen. Mit steigender Kinderzahl erhöht sich dieses Risiko zusätzlich. Daneben sind auch Personen aus dem Kanton Neuenburg überdurchschnittlich häufig erwerbstätig bei gleichzeitigem Sozialhilfebezug. Zudem weisen Personen im Alter zwischen 46 und 65 Jahren sowie sehr junge Personen im Alter unter 23 Jahre ein erhöhtes Risiko auf, Working Poor zu sein. Schliesslich bildet auch das Geschlecht einen entscheidenden Faktor für das Risiko der Zugehörigkeit zu Typ 3. So sind Frauen im Vergleich zu Männern häufiger von der Situation Working Poor betroffen.

Der stärkste Resilienzfaktor für den Working Poor-Status bildet der Kanton Jura. Personen aus dem Kanton Jura weisen also relativ selten einen Sozialhilfebezug bei gleichzeitiger Erwerbsarbeit auf. Zudem bildet auch der Haushaltstyp eine wichtige Einflussgrösse, durch die das Risiko für den Status Working Poor vermindert wird. Dabei sind Paare ohne Kinder und nicht-alleinlebende Personen nur selten Working Poor.

Tabelle 17: Typ 3 "Working Poor (Erwerb und SH)"

Risikomerkmale	Risikofaktor	Resilienzmerkmale	Resilienzfaktor
Haushaltstyp: Alleinerziehende mit 3+ Kinder	1.9	Kanton: Jura	3.3
Haushaltstyp: Alleinerziehende mit 2 Kinder	1.6	Haushaltstyp: Paare ohne Kind	1.6
Kanton: Neuenburg	1.5	Haushaltstyp: Nicht-Alleinlebende	1.5
Haushaltstyp: Alleinerziehend mit 1 Kind	1.4	Geschlecht:	-
Alter: 46 bis 65	1.3	Alter:	-
Alter: unter 23 Jahre	1.3	Nationalität:	-
Geschlecht: weiblich	1.2	Gemeindetypologie:	-
Nationalität:	-		
Gemeindetypologie:	-		.

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010

Bemerkungen: Referenzkategorien der Risiko- bzw. Resilienzmerkmale sind:

Geschlecht: Männlich

Haushaltstyp: Paare mit 2 Kindern

Kanton: Freiburg

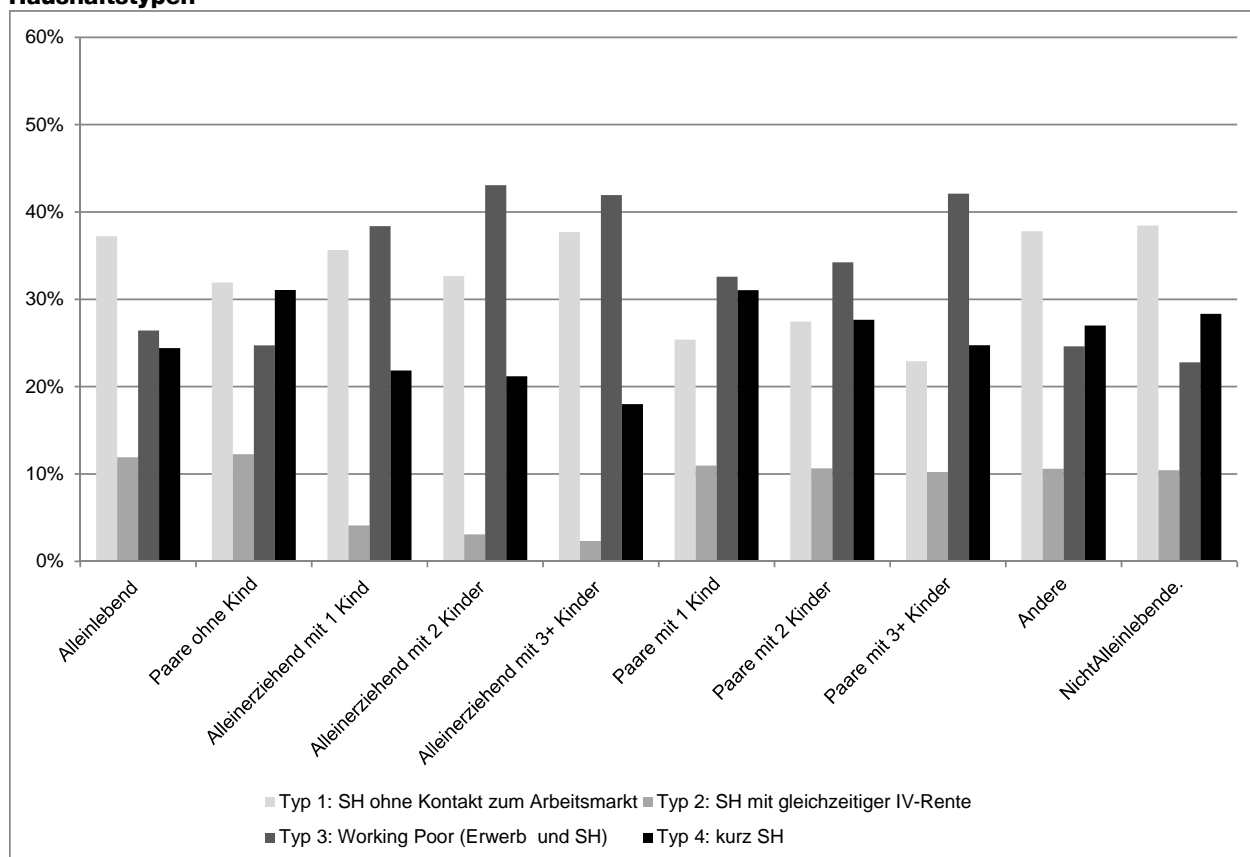
Alter: 26 bis 45

Nationalität: EU2

Gemeindetypologie: periurbane Gemeinden

Analog zur Auswertung der ALV wurde in einem nächsten Schritt für den Haushaltstyp als zweitwichtigstes Merkmal des Modelles (vgl. **Abbildung 7**) die Wahrscheinlichkeiten berechnet, zu einem der drei Typen zu gehören. Insgesamt betrachtet ist die Wahrscheinlichkeit der Zugehörigkeit zu Typ 2 (Sozialhilfe mit gleichzeitiger IV-Rente) wesentlich tiefer als bei den anderen Typen. Insbesondere bei alleinerziehenden Personen ist diese Konstellation besonders selten. Wird der unproblematische Sozialhilfebezug (Typ 4: kurz SH) betrachtet, so fällt auf, dass vor allem Paare mit einem oder keinem Kind eine erhöhte Wahrscheinlichkeit aufweisen, in diese Kategorie zu fallen. Alleinerziehende Personen oder Paare mit mehr als drei Kindern hingegen sind besonders von Verläufen mit Sozialhilfe bei gleichzeitiger Erwerbstätigkeit (Typ 3: Working Poor) betroffen. Die Chance, bei Sozialhilfebezug über eine längere Phase aus dem Arbeitsmarkt auszuscheiden, ist vor allem für Paare mit Kindern relativ tief.

Abbildung 9: Vorhergesagte Wahrscheinlichkeit der Zugehörigkeit zu Typ 1 - 4 für unterschiedliche Haushaltstypen



Quelle: Modellschätzungen anhand SHIVALV 2006-2010, AHV-IK 2006-2010.

Auf Basis der vorangegangenen Analysen wurden schliesslich verschiedene besondere Risikogruppen ausfindig gemacht, für die dann in Tabelle 18 die Wahrscheinlichkeiten der Zugehörigkeit zu einem der vier Typen berechnet wurden.

Dabei zeigt sich, dass vor allem allein lebende im Alter 46 bis 65 eine besonders hohe Wahrscheinlichkeit aufweisen, Sozialhilfe zu beziehen und dabei dauerhaft aus dem Arbeitsmarkt auszuscheiden. Mit 41.8% ist die Wahrscheinlichkeit sogar mehr als dreimal so hoch wie diejenige eines kurzen und somit unproblematischen Sozialhilfebezuges (Typ 4). Diese Gruppe scheint allgemein besonders gefährdet zu sein, im System der Sozialhilfe problematische Verläufe aufzuweisen. So ist auch die Wahrscheinlichkeit von Typ 2 (SH und IV-Rente) mit einem Fünftel überdurchschnittlich hoch. Die Wahrscheinlichkeit, nur kurz Sozialhilfe zu beziehen, ist mit 16.5% unterdurchschnittlich.

Die nächste Risikogruppe beinhaltet alleinerziehende Personen mit drei und mehr Kindern. Sie weisen mit 41.4 Prozent eine fast doppelt so hohe Wahrscheinlichkeit auf, in die Gruppe der Working Poor (Typ 3) zu fallen, wie die Durchschnittsgruppe. Gleichzeitig ist ihre Wahrscheinlichkeit, Sozialhilfe und gleichzeitig IV-Rente zu beziehen, stark vermindert gegenüber dem Durchschnittstyp.

Mit Personen im Alter zwischen 23 und 25 Jahren konnte eine Gruppe ausfindig gemacht werden, die mit 32.6 Prozent die höchste Wahrscheinlichkeit besitzt, nur kurz Sozialhilfe zu beziehen. Dabei muss jedoch berücksichtigt werden, dass die Wahrscheinlichkeit nicht wesentlich höher ist als diejenige des Durchschnittstypen. Dasselbe gilt auch für die Wahrscheinlichkeiten dieser Gruppe, einen Sozialhilfverlauf gemäss Typ 1 oder Typ 2 aufzuweisen, die beide nur wenig tiefer als der Durchschnitt sind.

Durchschnittliche Risiken für problematische Verläufe im System der Sozialen Sicherheit weisen alleinlebende Schweizer/innen im Alter zwischen 26 und 45 auf, ebenso Paare ohne Kinder.

Tabelle 18: Vorhergesagte Wahrscheinlichkeiten für ausgewählte Gruppen

Gruppe	Typ 1: SH ohne Kontakt zum Arbeitsmarkt	Typ 2: SH mit gleichzeitiger IV-Rente	Typ 3: Working Poor (Erwerb und SH)	Typ 4: kurz SH	Gesamt
alleinlebende 46 bis 65 Jährige	41.8%	20.7%	21.0%	16.5%	100%
alleinerziehende mit 3+ Kinder	37.8%	2.7%	41.4%	18.1%	100%
alleinlebender Schweizer/innen zwischen 26 und 45 Jahre	36.1%	13.8%	22.8%	27.3%	100%
23 bis 25 Jährige	29.1%	6.6%	31.8%	32.6%	100%
Paare ohne Kind	31.6%	13.5%	24.3%	30.6%	100%

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010

Bemerkungen: Wahrscheinlichkeiten mit hoher Abweichung zum Durchschnittstyp sind **fett** markiert.

4.3.4 Risikoprofile für Eintritte über IV

Als nächstes wird nun dieselbe Analyse wie in den vorangehenden beiden Abschnitten mit Verläufen im System der Invalidenversicherung durchgeführt. Die entsprechende Verlaufstypologie besitzt dabei die folgenden Ausprägungen:

- Typ 1 "Einfache Verläufe mit Vollrente"
- Typ 2 "Komplexe Verläufe mit offenem Ausgang (IV-Rente, IV und Erwerb sowie keine Leistung/kein Erwerb)"
- Typ 3 "Sehr komplexe Verläufe mit langen Kombinationsbezügen (IV, IVT und ALV)"
- Typ 4 "Einfache Verläufe, die schnell in Phasen mit IV-Rente und Erwerbstätigkeit münden"

Da Typ 1 ein relativ einfacher und häufiger Verlauf in der IV darstellt, wurde er als Referenz für die Modellrechnung verwendet.

Tabelle 19, Tabelle 20 und Tabelle 21 beinhalten die *Risiko- und Resilienzmerkmale* mit Fokus auf komplexe Verläufe mit offenem Ausgang (Typ 2), sehr komplexe Verläufe mit langen Kombinationsbezügen (Typ 3) und einfachen Verläufen (Typ 4). Die Merkmale mit prädiktiven Eigenschaften sind dabei in absteigender Reihenfolge entsprechend der Höhe des Risiko- bzw. Resilienzfaktors sortiert. Diese Werte widerspiegeln gleichzeitig die Prädiktorstärke der entsprechenden Merkmalausprägungen, was Vergleiche ermöglicht.

Tabelle 19 zeigt die entsprechenden Auswertungen für komplexe Verläufe in der IV (Typ 2). Dabei sind insbesondere Personen von diesen Konstellationen betroffen, die an einer physischen Krankheit leiden und/oder eine halbe Rente beziehen. Letzterer Effekt kann zu einem gewissen Grad als endogener Zusammenhang interpretiert werden, da die Wahrscheinlichkeit des Referenzverlaufs „IV-Rente und Erwerb“

stark eingeschränkt ist bei Zusprache einer Vollrente.

Anders sieht es aus, wenn die Resilienzmerkmale betrachtet werden, bei denen fast alle im Modell berücksichtigten Variablen ausser das Alter signifikante Effekte zeigen. Die entscheidendste Grösse bildet dabei mit einem Resilienzfaktor von 2.4 die Herkunft aus einem Land ausserhalb der EU. Zudem sind auch Personen mit einer ganzen oder dreiviertel Rente eher selten von komplexen Verläufen betroffen. Ähnliches trifft auch für Personen mit Geburtsgebrechen zu, die relativ selten komplexe Verläufe aufweisen. Des Weiteren sind schliesslich auch Personen aus dem Kanton Tessin weniger häufig von diesem Typ des IV-Bezuges betroffen. Schliesslich gilt es noch darauf zu verweisen, dass verheiratete Personen ebenfalls relativ selten komplexe Verläufe in der IV aufweisen.

Tabelle 19: Typ 2 "Komplexe Verläufe mit offenem Ausgang (IV-Rente, IV und Erwerb sowie keine Leistung/kein Erwerb)"

Risikomerkmal	Risikofaktor	Resilienzmerkmale	Resilienzfaktor
Gebrechensart: physisch erkrankt	1.5	Nationalität: Welt	2.4
IV-Rententeil: halbe Rente	1.5	IV-Rententeil: ganze Rente	2.2
Geschlecht:	-	Gebrechensart: Geburtsgebrechen	2.1
Kanton:	-	Geschlecht: weiblich	1.9
Zivilstand:	-	IV-Rententeil: dreiviertel	1.8
Alter:	-	Kanton: Tessin	1.5
Nationalität:	-	Zivilstand: verheiratet	1.4
		Alter:	-

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010

Bemerkungen: Referenzkategorien der Risiko- bzw. Resilienzmerkmale sind:

Geschlecht: Männlich

Kanton: Thurgau

Zivilstand: verwitwet

Alter: 46 bis 65

Nationalität: EU17 Süd

Rententeil: viertel Rente

Gebrechensart: psychisch

Werden nun die sehr komplexen Verläufe betrachtet, die sich durch lange Kombibezüge (IV, IVT, ALV bzw. SH) auszeichnen, so wird ersichtlich, dass primär der IV-Rententeil dafür ausschlaggebend ist, ob Personen in diesen Typ des IV-Bezugs fallen (vgl. Tabelle 20). So sind vor allem Personen mit einem relativ kleinen Bezug von IV von dieser Konstellation betroffen (halbe oder viertel Rente). Letzterer Effekt kann zu einem gewissen Grad als endogener Zusammenhang interpretiert werden (vgl. oben).

Ein Blick auf die Resilienzmerkmale zeigt, dass Personen mit einer ganzen Rente nur selten sehr komplexe Verläufe aufweisen. So ist der Resilienzfaktor für dieses Merkmal mit 5.5 äusserst hoch. Auch hier kann von einem grossenteils endogenen Zusammenhang zwischen Rententeil und Möglichkeit zur Erwerbstätigkeit gesprochen werden. Zudem sind auch verheiratete Personen und Frauen nur sehr selten Typ 3 der IV-Typologie zuzurechnen.

Tabelle 20: Typ 3 "Sehr komplexe Verläufe mit langen Kombinationsbezügen (IV, IVT und ALV)"

Risikomerkmale	Risikofaktor	Resilienzmerkmale	Resilienzfaktor
IV-Rententeil: halbe Rente	3.2	IV-Rententeil: ganze Rente	5.5
IV-Rententeil: viertel Rente	3.0	Zivilstand: verheiratet	3.1
Kanton: Bern	2.6	Geschlecht: weiblich	2.1
Geschlecht:	-	Kanton:	-
Zivilstand:	-	Alter:	-
Alter:	-	Nationalität:	-
Nationalität:	-	Gebrechensart:	-
Gebrechensart:	-		

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010

Bemerkungen: Referenzkategorien der Risiko- bzw. Resilienzmerkmale sind:

Geschlecht: Männlich

Kanton: Basel-Stadt

Zivilstand: ledig

Alter: unter 23 Jahre

Nationalität: EU17 Süd

Rententeil: dreiviertel Rente

Gebrechensart: physisch Unfall

Schliesslich sollen auch noch die Risiko- und Resilienzmerkmale für den vierten und letzten Typ des IV-Bezugs näher betrachtet werden. Dabei handelt es sich um einfache Verläufe, die entweder eine IV-Rente mit Erwerbstätigkeit oder eine ausschliessliche Erwerbstätigkeit nach sich ziehen. Tabelle 21 zeigt, dass von diesen sogenannten einfachen Verläufen vorwiegend Personen mit einer halben oder einer Viertelrente betroffen sind, also Personen mit eher kleinen Rententeilen. Zudem zeigen auch eher Schweizer/innen und Ledige Personen dieses Muster des IV-Bezuges.

Analog zu den sehr komplexen Verläufen bildet wiederum der Bezug einer ganzen IV-Rente das wichtigste Resilienzmerkmal, auch hier ist von einem endogenen Zusammenhang zu sprechen. Die zweite wichtige Grösse mit einem hohen Resilienzfaktor neben dem Rententeil stellt die Nationalität einer Person dar. So haben Personen, die von ausserhalb der EU in die Schweiz kamen, eher selten einfache Verläufe in der IV. Etwas weniger wichtig als Resilienzmerkmal für Typ 4 ist das Geschlecht, sodass Frauen signifikant weniger häufig in den Typ 4 fallen als Männer. Dasselbe gilt schliesslich auch für Personen, die verheiratet oder psychisch erkrankt sind.

Tabelle 21: Typ 4 "Einfache Verläufe, die schnell in Phasen mit IV-Rente und/oder Erwerbstätigkeit münden"

Risikomerkmale	Risikofaktor	Resilienzmerkmale	Resilienzfaktor
IV-Rententeil: halbe Rente	3.7	IV-Rententeil: ganze Rente	4.5
IV-Rententeil: viertel Rente	2.8	Nationalität: Welt	3.2
Nationalität: Schweiz	1.8	Geschlecht: weiblich	2.2
Zivilstand: ledig	1.6	Zivilstand: verheiratet	1.8
Geschlecht:	-	Gebrechensart: psychisch	1.4
Kanton:	-	Kanton:	-
Alter:	-	Alter:	-
Gebrechensart:	-		

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010

Bemerkungen: Referenzkategorien der Risiko- bzw. Resilienzmerkmale sind:

Geschlecht: Männlich

Kanton: Zug

Zivilstand: verwitwet

Alter: 46 bis 65

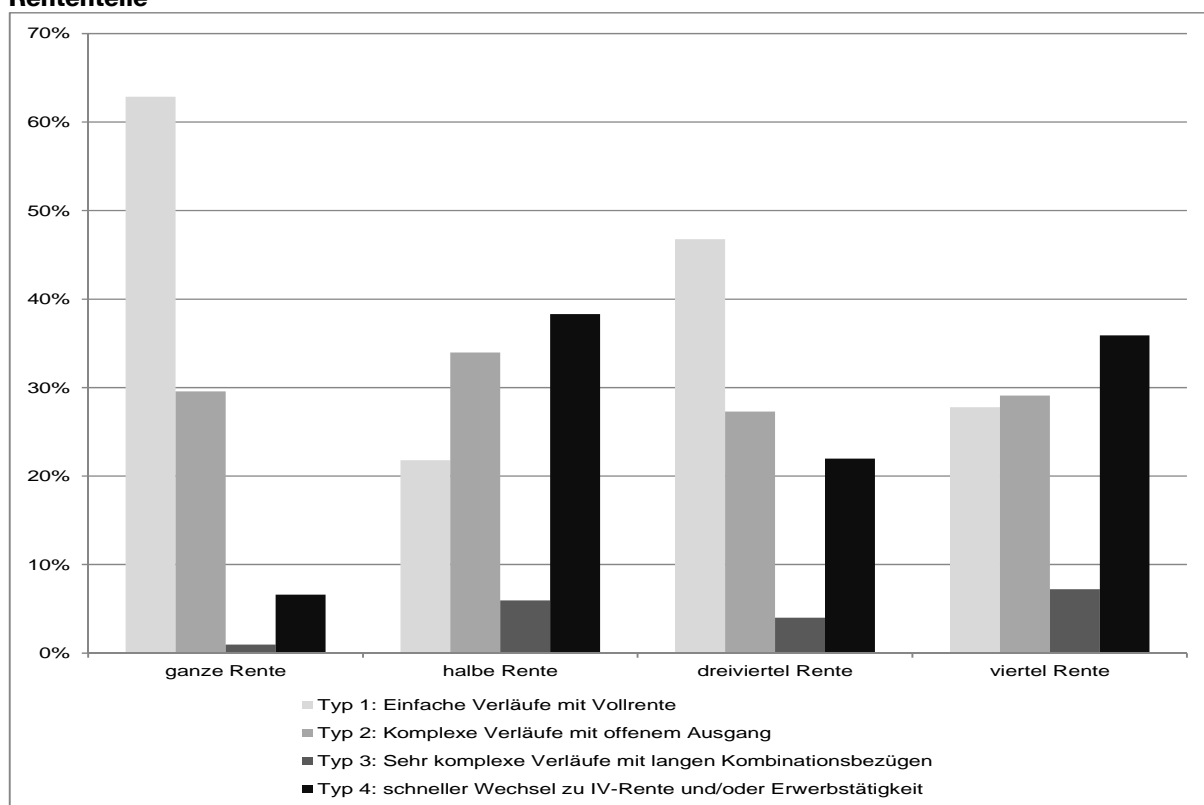
Nationalität: EU17 angrenzend

Rententeil: dreiviertel Rente

Gebrechensart: physischer Unfall

Nach der Darstellung der Risiko- und Resilienzmerkmale wurden in einem weiteren Schritt für das wichtigste Merkmal im IV-Modell den Rententeil (vgl. **Abbildung 7**) in Abbildung 10 die Wahrscheinlichkeiten berechnet zu einem der vier Typen des IV-Modells zu gehören. Dabei zeigt sich, sehr Komplexe Verläufe (Typ 3) nur relativ selten vorkommen und bei Personen mit einer viertel Rente am häufigsten sind. Einfache Verläufe sind hingegen wesentlich häufiger und dies insbesondere bei Personen die eine ganze oder eine dreiviertel Rente beziehen, also bei Personen mit einem relativ hohen Rententeil. Bei Personen mit einem tiefen Rententeil (halbe und viertel Rente) spielt hingegen Typ 4 eine entscheidende Rolle. Sie Wechseln also schnell in die IV-Rente oder steigen wieder in die Erwerbstätigkeit ein. Vergleicht man hierzu die Personen mit einer ganzen Rente, so zeigt sich eine sehr tiefe Wahrscheinlichkeit für die Zugehörigkeit zu Typ 4. Die komplexen Verläufe mit offenem Ausgang (Typ 3) sind schliesslich mehr oder weniger gleichmässig über die Abstufungen des Umfangs der IV-Rente verteilt. Einzig Personen, die eine halbe Rente beziehen besitzen eine leicht höhere Wahrscheinlichkeit der Zugehörigkeit zu diesem Typ.

Abbildung 10: Vorhergesagte Wahrscheinlichkeit der Zugehörigkeit zu Typ 1 - 4 für unterschiedliche IV-Rententeile



Quelle: Modellschätzungen anhand SHIVALV 2006-2010, AHV-IK 2006-2010.

Abschliessend wurden wiederum die Wahrscheinlichkeiten für Gruppen mit spezifischen Merkmalskombinationen berechnet, einen der vier Typen des Verlaufs nach Erstbezug einer IV-Rente aufzuweisen. Wie Tabelle 22 zeigt, besitzen verheiratete Frauen mit einem Geburtsgebrechen und voller IV-Rente, die von ausserhalb der EU in die Schweiz kamen, mit 92.1% die höchste Wahrscheinlichkeit, einfache Verläufe mit Vollrente (Typ 1) aufzuweisen. Diese Wahrscheinlichkeit ist dabei fast doppelt so hoch wie diejenige des Durchschnittstypen. Die übrigen 7.9 Prozent verteilen sich primär auf Typ 2, also komplexe Verläufe mit offenem Ausgang. Dabei ist die Wahrscheinlichkeit der Zugehörigkeit zu dieser Gruppe unter der entsprechenden Wahrscheinlichkeit aller anderen Gruppen.

Die nächste Gruppe, für die die Wahrscheinlichkeiten der Zugehörigkeit zu den Verlaufstypen berechnet wurde, sind psychisch erkrankte verheiratete Frauen mit voller IV-Rente aus einem Land ausserhalb der EU. Diese Gruppe besitzt mit 88.8 Prozent eine fast doppelte so hohe Wahrscheinlichkeit für einfache Verläufe (Typ 1) wie die Durchschnittsgruppe. Auf der anderen Seite sind für diese Personen komplexe Verläufe (Typ 2) und Verläufe mit schnellem Wechsel zu IV-Rente oder Erwerbstätigkeit (Typ 4) relativ unwahrscheinlich.

Anschliessend wurden verheiratete Frauen mit einer vollen IV-Rente näher betrachtet. Mit 72.5 Prozent besitzen diese Personen eine vergleichsweise hohe Wahrscheinlichkeit für die Zugehörigkeit zu Typ 1 und mit 22.4 Prozent eine relativ tiefe Wahrscheinlichkeit für Typ 2. Dabei sind diese Effekt jedoch nicht mehr so eindeutig, wie bei den zuvor behandelten Gruppen.

Als nächstes wurden Personen betrachtet, die physisch erkrankt sind und eine volle IV-Rente beziehen. Sie weisen mit 7.2 und 39.0 Prozent relativ häufig sehr komplexe Verläufe (Typ 3) und Verläufe mit schnellem Wechsel in die IV-Rente oder in die Erwerbsarbeit auf (Typ 4). Eher unwahrscheinlich sind für diese Personen

hingegen Verläufe gemäss Typ 1. Analoges wie für die eben behandelte Gruppe gilt auch für Personen die eine halbe IV-Rente beziehen. So besitzen sie relativ hohe Werte für Typ 3 und Typ 4. Für Typ 1 ist die Wahrscheinlichkeit hingegen vergleichsweise tief.

Schliesslich wurden als letzte Gruppe ledige Schweizer/-innen mit einer halben IV-Rente näher betrachtet. Hier zeigt sich vor allem, dass diese Personen mit 54.2 Prozent eine äusserst hohe Wahrscheinlichkeit aufweisen, schnell in eine IV-Rente und/oder in die Erwerbsarbeit zu wechseln (Typ 4). Zudem ist ein einfacher Verlauf, bei dem eine Vollrente bezogen wird, für diese Gruppe mit 11.1 Prozent relativ unwahrscheinlich. Daneben ist der Wert für Typ mit 7.0 Prozent relativ hoch.

Tabelle 22: Vorhergesagte Wahrscheinlichkeiten für ausgewählte Gruppen

Gruppe	Typ 1: Einfache Verläufe mit Vollrente	Typ 2: Komplexe Verläufe mit offenem Ausgang	Typ 3: Sehr komplexe Verläufe mit langen Kombinationsb ezügen	Typ 4: schneller Wechsel zu IV- Rente oder Erwerbstätigke it	Gesamt
verheiratete Frau mit einem Geburtsgebrechen und voller IV-Rente von ausserhalb der EU	92.1%	5.1%	0.9%	1.8%	100%
psychisch erkrankte verheiratete Frau mit voller IV- Rente von ausserhalb der EU	88.8%	9.4%	0.9%	0.9%	100%
verheiratete Frau mit voller IV- Rente	72.5%	22.4%	0.7%	4.4%	100%
physisch erkrankter verheirateter Schweizer mit voller IV-Rente zwischen 46 und 65 Jahren	48.3%	41.0%	0.8%	9.9%	100%
physisch erkrankt mit voller IV- Rente	19.8%	34.1%	7.2%	39.0%	100%
halbe IV-Rente	23.0%	31.2%	8.0%	37.7%	100%
ledige Schweizer/-innen mit halber IV-Rente	11.1%	27.7%	7.0%	54.2%	100%

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010

Bemerkungen: Wahrscheinlichkeiten mit hoher Abweichung zum Durchschnittstyp sind **fett** markiert.

4.3.5 Risikoprofile für Eintritte ins Gesamtsystem

Schliesslich wurden in einem letzten Schritt die Risikoprofile analog zu den einzelnen Systemen der sozialen Sicherheit für die Verlaufstypologie des Gesamtsystems gebildet. Die Typologie für das Gesamtsystem besitzt dabei die folgenden Ausprägungen:

- Typ 1: „Mit IV-Rente“
- Typ 2: „Existenzsicherung primär durch Sozialhilfe“
- Typ 3: „Am Rand des Arbeitsmarktes (ALV, Erwerb und Phasen ohne Erwerb und Leistungen)“
- Typ 4: „Kurze Unterstützungsphase zu Beginn, danach Erwerb“

Da Typ 4 nur kurz mit den Systemen der sozialen Sicherheit in Kontakt kommt und schnell wieder mit einer Erwerbsarbeit beginnt, kann er als relativ unproblematisch betrachtet werden. Aus diesem Grunde wurde er bei den Modellrechnungen als Referenzkategorie verwendet.

In Tabelle 25, Tabelle 24 und Tabelle 23 werden nun wiederum die *Risiko- und Resilienzmerkmale* mit Fokus auf Verläufe mit IV-Bezügen (Typ 1), Verläufe mit einer Existenzsicherung primär über die Sozialhilfe (Typ 2) und Verläufe am Rand des Arbeitsmarktes (Typ 3) ausgewiesen. Die Merkmale mit prädiktiven Eigenschaften sind dabei in absteigender Reihenfolge entsprechend der Höhe des Risiko- bzw. Resilienzfaktors sortiert. Diese Werte widerspiegeln gleichzeitig die Prädiktorstärke der entsprechenden Merkmalausprägungen, was Vergleiche ermöglicht.

Tabelle 23 bildet die Risiko- und Resilienzmerkmale für den Bezug einer IV-Rente (Typ 1) ab. Dabei weisen Personen zwischen 46 und 65 Jahren mit einem Faktor von 4.4 mit Abstand das höchste Risiko für Verläufe mit IV-Rente auf. Daneben bildet wiederum der Wohnkanton eine zentrale Grösse beim Risiko für diesen Verlaufstyp. Personen aus den Kantonen Appenzell Innerrhoden, Tessin, St. Gallen, Basel-Landschaft und Basel-Stadt. Des Weiteren ist die Nationalität als zentrale Risikogrösse für Verläufe mit IV-Bezug. Vor allem Personen, die von ausserhalb der EU in die Schweiz kommen, besitzen mit einem Risikofaktor von 1.8 ein relativ hohes Risiko für diesen Verlaufstyp. Daneben sind Schweizer/innen und Personen aus den EU17 Süd Staaten relativ häufig von IV-Renten betroffen. Schliesslich beziehen Personen, die verwitwet, geschieden oder getrennt sind, im Vergleich zu verheirateten Personen öfter eine IV-Rente.

Bei den Resilienzmerkmalen zeigt sich, dass vor allem Personen aus dem Kanton Zug eher selten von Verläufen mit IV-Renten betroffen sind. Daneben können auch die Kantone Zürich und Waadt als Resilienzmerkmale festgemacht werden. Zudem beziehen Personen aus den EU 17 Nord Staaten mit einem Resilienzfaktor von 1.6 eher selten eine IV-Rente. Darüber hinaus bildet der Zivilstand ein weiteres Resilienzmerkmal für den IV-Bezug. So sind ledige Personen seltener als verheiratete Personen IV-Bezüger. Schliesslich folgen Frauen und Personen zwischen 26 und 45 Jahren mit einem Resilienzfaktor von jeweils 1.2 eher selten einem Verlauf gemäss diesem Typ.

Tabelle 23: Typ 1 "Mit IV-Rente"

Risikomerkmale	Risikofaktor	Resilienzmerkmale	Resilienzfaktor
Alter: 46 bis 65	4.4	Kanton: Zug	2.0
Kanton: Appenzell I	1.9	Nationalität: EU 17 Nord	1.6
Nationalität: Welt	1.8	Kanton: Zürich	1.4
Kanton: Tessin	1.5	Kanton: Waadt	1.3
Kanton: St. Gallen	1.4	Zivilstand: ledig	1.3
Kanton: Basel-Landschaft	1.4	Geschlecht: weiblich	1.2
Kanton: Basel-Stadt	1.3	Alter: 26 bis 45	1.2
Nationalität: Schweiz	1.2		
Zivilstand: verwitwet, geschieden, getrennt	1.2		
Nationalität: EU 17 süd	1.2		
Geschlecht:	-		

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010

Bemerkungen: Referenzkategorien der Risiko- bzw. Resilienzmerkmale sind:

Geschlecht: Männlich

Kanton: Schaffhausen

Zivilstand: verheiratet

Alter: 18 bis 25

Nationalität: EU 17 angrenzend

Wird der nächste Typ der Typologie für das Gesamtsystem (Typ 2) in Tabelle 24 betrachtet, fällt auf, dass vor allem Personen, die aus einer Partnerschaft kommen, die also verwitwet, geschieden oder getrennt sind, ein erhöhtes Risiko aufweisen, ihre Existenz primär über die Sozialhilfe zu sichern. Wie schon bei Verläufen am Rande des Arbeitsmarktes spielt auch bei der Sozialhilfe der Wohnkanton eine bedeutende Rolle. So besitzen Personen aus den Kantonen Neuenburg, Basel-Stadt, Genf, Waadt, Bern, Solothurn und Zürich ein erhöhtes Risiko für einen Verlauf gemäss Typ 2. Schliesslich kann auch die Nationalität als entscheidender Risikofaktor beim Sozialhilfebezug festgemacht werden. Personen von ausserhalb der EU und Personen aus den EU8-Staaten besitzen somit ein erhöhtes Risiko, in diesen Verlaufstyp zu fallen.

Auch bei den Resilienzmerkmalen bildet die kantonale Verortung eine zentrale Rolle ein. So sichern Personen aus Nidwalden, Graubünden, Thurgau, Wallis und Jura eher selten ihre Existenz primär durch die Sozialhilfe. Daneben besitzen Personen aus den nördlichen Staaten der EU17 ein tiefes Risiko von Verläufen mit Sozialhilfebezügen. Zwei weitere Gruppen, die sich eher selten grösstenteils auf Sozialhilfe stützen, bilden verheiratete Personen und Personen zwischen 26 und 45 Jahren. Schliesslich muss noch erwähnt werden, dass Frauen im Vergleich zu Männern eher selten Sozialhilfe beziehen. Der entsprechende Resilienzfaktor ist jedoch mit 1.2 äusserst tief. Zudem muss hierbei berücksichtigt werden, dass bei der Sozialhilfe in Partnerschaften häufig der Mann als Dossierträger auftritt und Frauen dadurch unterrepräsentiert sein dürften.

Tabelle 24: Typ 2 "Existenzsicherung primär durch Sozialhilfe"

Risikomerkmale	Risikofaktor	Resilienzmerkmale	Resilienzfaktor
Zivilstand: verwitwet, geschieden, getrennt	3.6	Kanton: Nidwalden	1.7
Kanton: Neuenburg	3.3	Nationalität: EU 17 Nord	1.7
Nationalität: Welt	3.1	Kanton: Graubünden	1.6
Kanton: Basel-Stadt	2.9	Kanton: Thurgau	1.6
Nationalität: EU 8	2.7	Zivilstand: verheiratet	1.4
Kanton: Genf	2.5	Kanton: Wallis	1.4
Kanton: Waadt	2.3	Kanton: Jura	1.4
Kanton: Bern	2.0	Alter: 26 bis 45	1.3
Kanton: Solothurn	1.2	Geschlecht: weiblich	1.2
Kanton: Zürich	1.2		
Geschlecht:	-		
Alter:	-		

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010

Bemerkungen: Referenzkategorien der Risiko- bzw. Resilienzmerkmale sind:

Geschlecht: Männlich

Kanton: Tessin

Zivilstand: ledig

Alter: 46 bis 65

Nationalität: EU 17 angrenzend

In Tabelle 25 wird ersichtlich, dass vor allem der Wohnkanton ein zentraler Risikofaktor für Verläufe am Rand des Arbeitsmarktes (Typ 3) darstellt. So sind Personen aus den Kantonen Genf, Tessin, Basel-Stadt, Wallis, Neuenburg und Waadt besonders häufig nur teilweise Erwerbstätig. Dabei fällt auf, dass Kantone aus der Westschweiz relativ stark vertreten sind. Neben der regionalen Verortung spielt auch das Alter eine zentrale Rolle bei der Zuteilung zu Typ 3. Personen zwischen 46 und 65 Jahren sind somit verstärkt am Rande des Arbeitsmarktes anzutreffen. Des Weiteren sind auch Personen, die von ausserhalb der EU in die Schweiz eingereist sind, überdurchschnittlich häufig von dieser Konstellation betroffen. Schliesslich kann auch hinsichtlich des Geschlechtes ein erhöhtes Risiko von Frauen für diesen Verlaufstyp festgemacht werden. Mit einem Risikofaktor von 1.1 fällt dieser Effekt jedoch äusserst schwach aus.

Wie Tabelle 25 zeigt, sind Schweizer verhältnismässig nur selten von Verläufen gemäss Typ 3 betroffen sondern Weisen eher kurze Unterstützungsphasen (Typ 4) auf. Dasselbe gilt auf einem etwas tieferen Niveau auch für Personen, die aus den angrenzenden Ländern der EU 17 stammen. Daneben spielt wiederum der Wohnkanton eine zentrale Rolle beim Risiko von Verläufen am Rande des Arbeitsmarktes. So sind Personen aus den Kantonen Nidwalden, Zug, Schwyz, Thurgau und Aargau seltener von diesem Verlauf betroffen. Schliesslich bilden auch der Zivilstand und das Alter Resilienzfaktoren für Typ 3. So weisen ledige und junge Personen zwischen 18 und 23 relativ selten Verläufe gemäss Typ 3 auf. Dabei muss aber berücksichtigt werden, dass der Resilienfaktor des Alters mit einem Wert von 1.2 relativ niedrig ist.

Tabelle 25: Typ 3 "Am Rand des Arbeitsmarktes (ALV, Erwerb und Phasen ohne Erwerb und Leistungen)

Risikomerkmale	Risikofaktor	Resilienzmerkmale	Resilienzfaktor
Kanton: Genf	2.0	Nationalität: Schweiz	1.7
Kanton: Tessin	1.8	Kanton: Nidwalden	1.4
Alter: 46 bis 65	1.7	Zivilstand: ledig	1.3
Nationalität: Welt	1.4	Kanton: Zug	1.2
Kanton: Basel-Stadt	1.3	Alter: 18 bis 25	1.2
Kanton: Wallis	1.3	Nationalität: EU 17 angrenzend	1.2
Kanton: Neuenburg	1.3	Kanton: Schwyz	1.2
Kanton: Waadt	1.2	Kanton: Thurgau	1.1
Geschlecht: weiblich	1.1	Kanton: Aargau	1.1
Zivilstand:	-	Geschlecht:	-

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010

Bemerkungen: Referenzkategorien der Risiko- bzw. Resilienzmerkmale sind:

Geschlecht: Männlich

Kanton: Basel-Landschaft

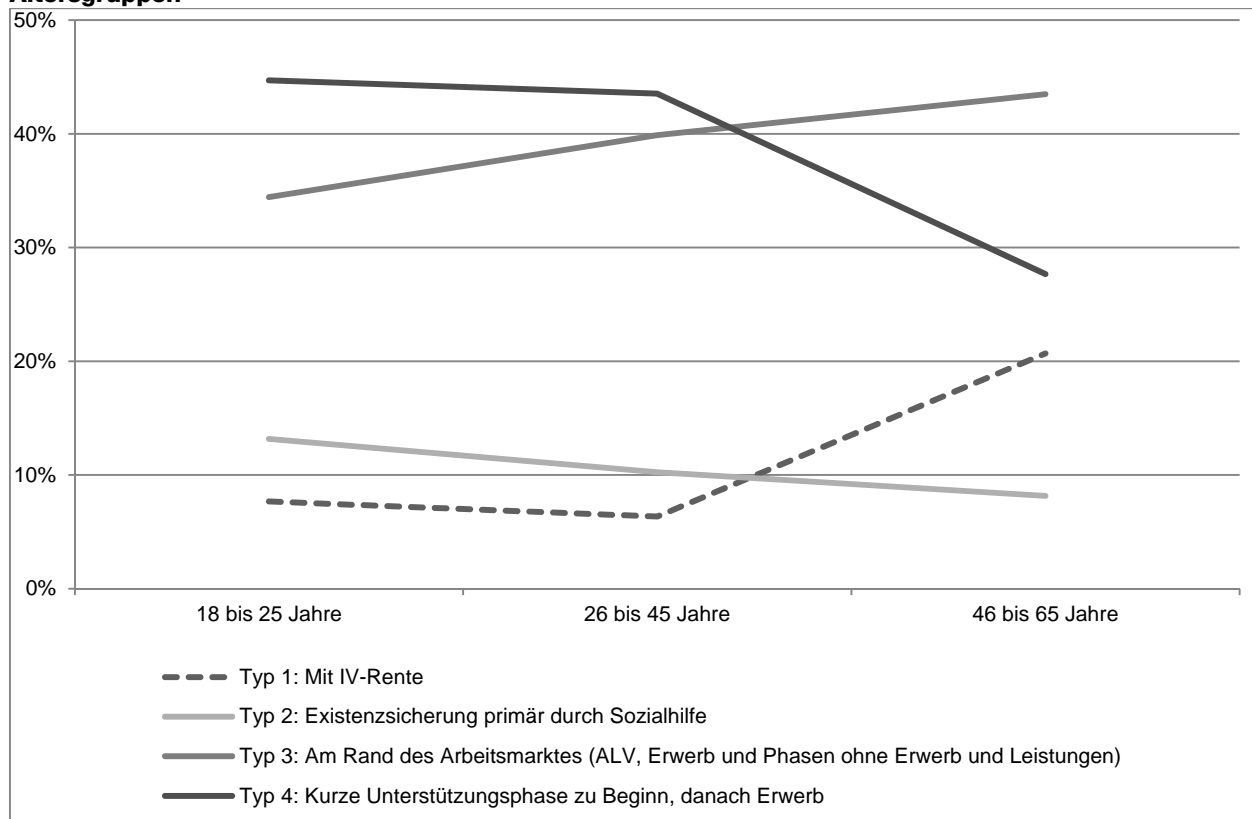
Zivilstand: verwitwet, geschieden, getrennt

Alter: 26 bis 45

Nationalität: EU 17 Nord

In einem nächsten Schritt wurden wiederum für die wichtigste Variable des Modelles, dem Alter, die Wahrscheinlichkeiten berechnet, zu einem der vier Typen zu gehören (vgl. Abbildung 11). Insgesamt betrachtet kann gesagt werden, dass die Wahrscheinlichkeiten einer IV-Rente (Typ 1) oder einer Existenzsicherung primär über die Sozialhilfe (Typ 2) über alle Altersgruppen hinweg vergleichsweise niedrig sind. Der grösste Teil der behandelten Personen kommt nur für kurze Zeit mit einem der Systeme der sozialen Sicherheit in Kontakt (Typ 4) oder hält sich am Rande des Arbeitsmarktes auf (Typ 3). Werden nun die vier Typen gesondert nach den Alterskategorien betrachtet, so zeigt sich, dass die Wahrscheinlichkeit für kurze Unterstützungsphasen mit anschliessender Erwerbstätigkeit mit zunehmendem Alter abnimmt. Älteren Personen fällt es also schwerer, nach einem Bezug von Sozialleistungen schnell wieder in die Erwerbstätigkeit zurückzukehren. Eine gespiegelte Verlaufskurve jedoch auf einem tieferen Niveau kann für Personen mit IV-Bezug festgestellt werden. So nimmt mit steigendem Alter die Wahrscheinlichkeit zu, dass Personen Verläufe gemäss Typ 1 aufweisen. Bei Typ 2 der Existenzsicherung primär über die Sozialhilfe zeigt sich, dass ältere Personen seltener ihre Existenz primär über die Sozialhilfe sichern. Dabei sind jedoch die Unterschiede zwischen den Altersgruppen nur relativ klein. Schliesslich müssen die Wahrscheinlichkeiten der einzelnen Altersgruppen für Verläufe am Rande des Arbeitsmarktes (Typ 3) näher betrachtet werden. Hier zeigt sich, dass mit steigendem Alter auch die Wahrscheinlichkeit für diesen Verlauf zunimmt.

Abbildung 11: Vorhergesagte Wahrscheinlichkeit der Zugehörigkeit zu Typ 1 - 4 für unterschiedliche Altersgruppen



Quelle: Modellschätzungen anhand SHIVALV 2006-2010, AHV-IK 2006-2010.

In Tabelle 26 sind zum Abschluss der Risikoprofile für die Typologie des Gesamtsystems wiederum einzelne Gruppen mit besonderen Risikokonstellationen abgebildet. Dabei zeigt sich, dass verwitwete, geschiedene oder getrennte Schweizer/-innen überdurchschnittlich oft IV-Rente beziehen (Typ 1). Mit 22.2% ist die Wahrscheinlichkeit dieses Verlaufes fast viermal so hoch wie diejenige des Durchschnittstypen. Daneben weisen diese Personen eine vergleichsweise tiefe Wahrscheinlichkeit für den unproblematischen Verlauf gemäss Typ 4 auf, bei dem auf einen kurzen Leistungsbezug relativ schnell wieder Erwerbstätigkeit folgt. Schliesslich ist die Wahrscheinlichkeit für einen Verlauf mit Sozialhilfebezug (Typ 2) im Vergleich zur Durchschnittsgruppe relativ hoch. Bei diesen Personen kann also insgesamt von einer eher problematischen Gruppe gesprochen werden.

Die nächste Gruppe die näher betrachtet wurde, bilden verwitwete, geschiedene oder getrennte Personen aus einem Land ausserhalb der EU. Hierbei zeigt sich eine überaus starke Tendenz für eine Existenzsicherung primär über die Sozialhilfe (Typ 2). Die Wahrscheinlichkeit für diesen Verlauf ist mit 40.5 % mehr als viermal so hoch wie bei der Durchschnittsgruppe. Zudem kann für diese Gruppe eine tiefe Wahrscheinlichkeit eines unproblematischen Verlaufes gemäss Typ 4 festgestellt werden. Somit handelt es sich auch hierbei wiederum um eine eher problematische Gruppe.

Eine weitere Gruppe, deren Verläufe im Gesamtsystem näher betrachtet wurden, bilden Frauen zwischen 46 und 65 Jahren aus einem Land ausserhalb der EU. Dabei zeigt sich vor allem eine erhöhte Wahrscheinlichkeit für Verläufe am Rande des Arbeitsmarktes (Typ 3). Daneben weist diese Gruppe mit 17.7% eine verhältnismässig tiefe Wahrscheinlichkeit für die unproblematischen Verläufe gemäss Typ 4 auf.

Betrachtet man nun ledige Schweizer/-innen zwischen 18 und 25 Jahren, so zeigt sich ein umgekehrtes Bild gegenüber der zuvor beschriebenen Gruppe. Sie besitzt einerseits die tiefste Wahrscheinlichkeit für einen Verlauf am Rande des Arbeitsmarktes und andererseits die höchste Wahrscheinlichkeit für den unproblematischen Verlauf nach Typ 4. Dabei muss jedoch berücksichtigt werden, dass die Abweichungen vom Durchschnittstyp jeweils nur sehr klein sind. Zudem besitzt diese Gruppe der jungen Erwachsenen mit einem Wert von 13.0% auch einen leicht erhöhten Wert für den Bezug von Sozialhilfeleistungen.

In einem weiteren Schritt wurden die Wahrscheinlichkeiten der Zugehörigkeit zu einem der vier Verlaufstypen von verheirateten Frauen zwischen 26 und 45 Jahren aus der EU 17 Nord berechnet. Hierbei zeigen sich äusserst geringe Wahrscheinlichkeiten für IV-Renten und Sozialhilfebezüge. Sie weisen also nur selten Verläufe auf, die weit vom Arbeitsmarkt entfernt sind. Jedoch kann eine erhöhte Wahrscheinlichkeit für Verläufe am Rande des Arbeitsmarktes festgestellt werden.

Schliesslich wurden als letzte Gruppe ledige Frauen aus der EU 17 Nord zwischen 26 und 45 Jahren genauer betrachtet. Bei dieser Gruppe liegt vor allem eine sehr tiefe Wahrscheinlichkeit für Verläufe mit einer IV-Rente vor. Die Wahrscheinlichkeiten für die anderen drei Verlaufstypen unterscheiden sich hingegen nicht wesentlich von der Durchschnittsgruppe.

Tabelle 26: Vorhergesagte Wahrscheinlichkeiten für bestimmte Gruppen

Gruppe	Typ 1 Mit IV-Rente	Typ 2 Existenzsicherung primär durch Sozialhilfe	Typ 3 Am Rand des Arbeitsmarktes	Typ 4 Kurze Unterstützungsphase zu Beginn, danach Erwerb	Gesamt
verwitwete, geschiedene oder getrennte Schweizer/-innen zwischen 46 und 65 Jahren	22.2%	21.0%	33.2%	23.6%	100%
verwitwete, geschiedene oder getrennte Personen von ausserhalb der EU	7.9%	40.5%	34.6%	17.0%	100%
Frauen zwischen 46 und 65 Jahren aus ausserhalb der EU	16.4%	13.7%	52.1%	17.7%	100%
26 bis 45 jährige, ledige Schweizerinnen	5.7%	9.2%	33.6%	51.5%	100%
ledige Schweizer/-innen zwischen 18 und 25	7.6%	13.0%	27.5%	51.9%	100%
verheiratete Frauen zwischen 26 und 45 Jahren aus der EU 17 Nord	2.6%	2.9%	55.9%	38.5%	100%
ledige Frauen aus der EU 17 Nord zwischen 26 und 45 Jahren	2.4%	4.8%	48.7%	44.1%	100%

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010

Bemerkungen: Wahrscheinlichkeiten mit hoher Abweichung zum Durchschnittstyp sind **fett** markiert.

4.4 Modell zur Prognose der Risikostruktur

Ziel dieses Arbeitsschrittes ist eine Prognose der Risikostruktur anhand von Personen, die nach 2006 mit dem System der sozialen Sicherheit in Kontakt gekommen sind. Dafür arbeiten wir mit einem Risikomodell, dessen Grundlage in Abschnitt 4.4.1 beschrieben ist. Im selben Abschnitt wird ferner die anhand der prognostizierten Werte festzustellende Veränderung der Risikostruktur besprochen. In Abschnitt 4.4.2 erfolgt eine Beurteilung des Modells hinsichtlich Güte der Anpassung an die Daten und der Schätzparameter.

4.4.1 Prognose der Risikostruktur für die Eintrittskohorten 2007 bis 2010

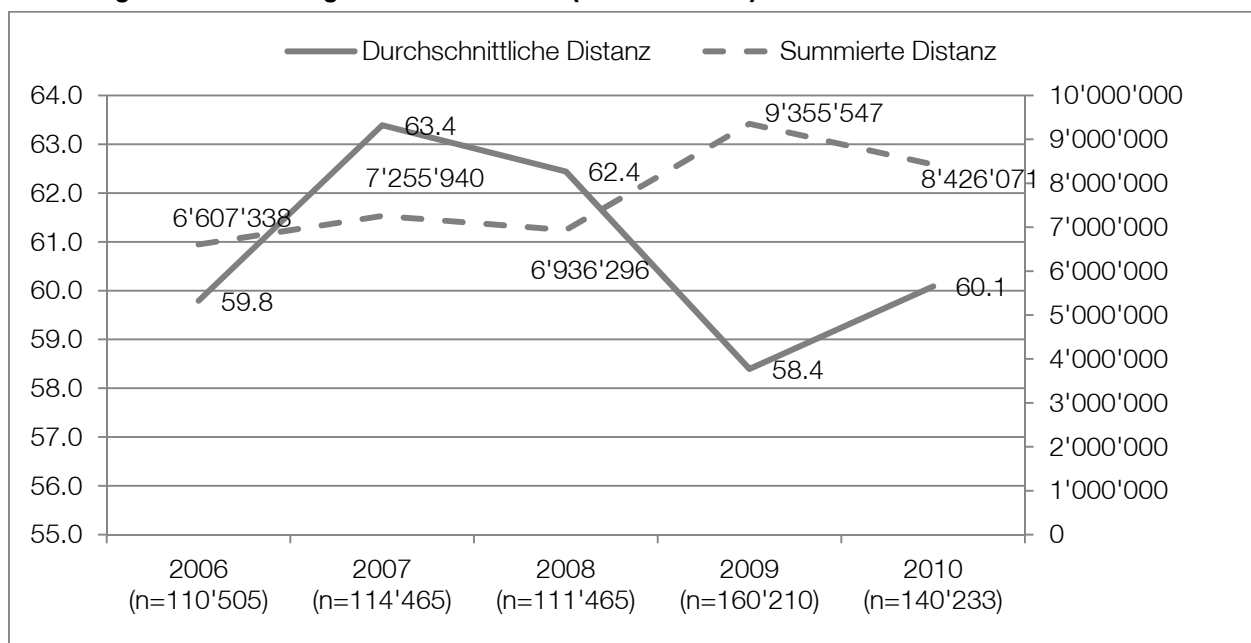
Für die Entwicklung des Modells stützen wir uns auf die beobachteten Verläufe der Kohorte 2006 und die Verdichtung der Verläufe über den im Abschnitt 4.2.1 vorgestellten Erwerbsintegrationsindex. Mit diesem numerischen Indikator der Nähe eines Verlaufes zum Arbeitsmarkt als abhängige Variable lässt sich ein Regressionsmodell auf Basis der Merkmale bei Eintritt berechnen. Die Situation bei Eintritt wird dabei durch folgende Merkmale abgebildet:

- Wohnkanton
- Demographische Merkmale (Geschlecht, Zivilstand, Alter, Nationalität)
- Bezug von Leistungen aus einem oder mehreren Systemen der Sozialen Sicherheit
- Erwerbstätigkeit

Damit ist der Zusammenhang zwischen Eintrittssituation und der erwarteten Distanz zur Erwerbstätigkeit modelliert und es kann eine Prognose für spätere Eintrittskohorten erfolgen. Dafür wird für jede Jahreskohorten, d.h. für jedes in den Daten abgebildete Jahr (2007 bis 2010) jene Fälle eruiert, die zum ersten Mal mit dem System der Sozialen Sicherheit in Kontakt gekommen sind. Auf Basis ihrer Merkmale bei Eintritt wird die Distanz zur Erwerbstätigkeit prognostiziert.

In Abbildung 12 sind die prognostizierten Werte zusammen mit den gemessenen Werten der Eintrittskohorte 2006 abgebildet. Höhere Werte entsprechen einer weiteren Distanz zur Erwerbstätigkeit, die mit längeren Perioden mit Leistungsbezügen verbunden sind. Damit ist eine Abschätzung der Gesamtbelastung des Systems der sozialen Sicherheit möglich. Veränderungen werden durch zwei simultan wirkende Kräfte begünstigt. Erstens ist davon auszugehen, dass eine Zunahme der Zahl an Leistungsbeziehenden mit einer stärkeren Belastung der Sozialen Sicherheit einhergeht. Für eine erste globale Betrachtung eignet sich daher die Interpretation der gestrichelten Linie (► Summierte Distanz). Dabei ist deutlich zu erkennen, wie sich die Zunahme der Fälle im Jahr 2009 in einer markant höheren summierten Distanz zur Erwerbstätigkeit widerspiegelt. Veränderungen ergeben sich jedoch nicht nur aus der reinen Zahl von Neubeziehenden. Vielmehr ist zweitens die veränderte Zusammensetzung von Neubeziehenden zu berücksichtigen. Dies ist aus der Betrachtung der durchgezogenen Linie (► Durchschnittliche Distanz) besser zu erkennen. So ist etwa die Zunahme der Zahl der Neubeziehenden im Jahr 2009 vor allem auf eine Zunahme bei der ALV zurückzuführen, bei welchen Verläufe erwartungsgemäss näher am Arbeitsmarkt sind. Dies äussert sich in einer tieferen durchschnittlichen Distanz. Auch berücksichtigt diese Zahl Veränderungen der demographischen Zusammensetzung. Wenn etwa der Anteil an Personen mit Risikomerkmale zunimmt (geringe Bildung, fortgeschrittenes Alter, Ausländer aus bestimmten Regionen), äussert sich dies in einer höheren prognostizierten durchschnittlichen Distanz. Alles in allem schwankt die mittlere erwartete Distanz zur Erwerbstätigkeit über die vier betrachteten Jahre nur geringfügig.

Abbildung 12: Veränderung der Risikostruktur (2006 bis 2010)



Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010, Berechnungen BFH
Bemerkungen: Es ist die erwartete Distanz zur Erwerbstätigkeit abgebildet. Die Werte für das Jahr 2006 wurden empirisch beobachtet. Die Jahre 2007-2010 sind prognostiziert.

Für die Prognose folgen wir der Annahme, dass das anhand der Kohorte 2006 erarbeitete Risikomodell eine gewisse Stabilität über die Zeit aufweist, d.h. dass mit Datengrundlage 2006 gefundene Zusammenhangsmuster zu späteren Zeitpunkten Gültigkeit besitzt. Das mag für kürzere Zeiträume plausibel sein. Wie bei allen sozialwissenschaftlichen Phänomenen vermuten wir aber auch hier, dass gefundene Zusammenhänge stark von gesamtgesellschaftlichen Entwicklungen beeinflusst sind und sich daraus, über längere Zeiträume betrachtet, Verschiebungen von Zusammenhangsmustern ergeben können. Analysen mit neueren Daten zur Stabilität des Modells scheinen uns aus diesem Grund unerlässlich.

Eine Erweiterung der Analysen sehen wir ferner darin, die Distanz zur Erwerbstätigkeit über kürzere Zeiträume zu betrachten (zwei oder drei Jahre) und somit kurzfristige Prognosen vorzunehmen.

4.4.2 Beurteilung der Modellgüte und der Schätzparameter

Nachfolgend werden die wichtigsten Ergebnisse der Modellschätzung mit Datenbasis Kohorte 2006 besprochen. Die Tabelle mit den vollständigen Ergebnissen befindet sich im Anhang (7.4).

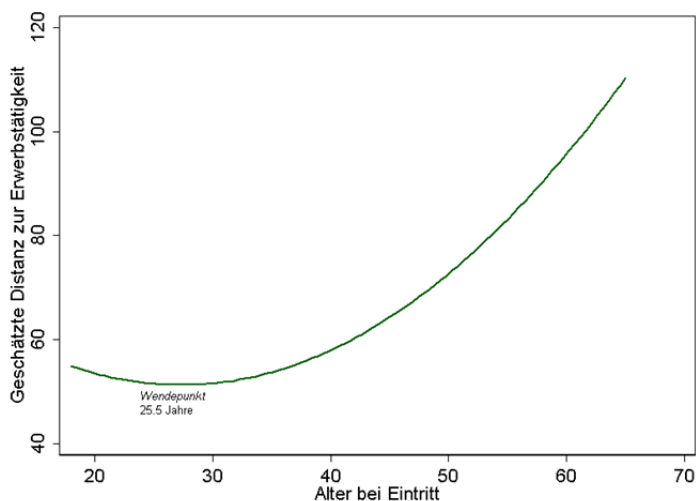
Die Güte des Modells lässt sich über den R-squared-Wert ablesen. Diesem lässt sich entnehmen, dass 41% der Varianz der Zielvariablen anhand des Risikomodells erklärt werden kann. Dies ist für sozialwissenschaftliche Verhältnisse ein vergleichsweise hoher Wert. Dabei gilt es anzumerken, dass insbesondere der Zustand bei Eintritt einen hohen Teil der Varianzaufklärung ausmacht. Wenn diese Variable aus dem Modell ausgeschlossen wird, senkt sich der Anteil der erklärten Varianz auf 12%.

Aus den Effektschätzern selber kann gelesen werden, welche Merkmalsausprägungen Verläufe begünstigen, die näher an der Erwerbstätigkeit sind bzw. eher mit einem Ausschluss aus dem Arbeitsmarkt einhergehen. So stehen positive Werte für ein erhöhtes Risiko (Distanz zur Erwerbstätigkeit ist grösser). Entsprechend stehen negative Werte für ein geringes Risiko (Distanz zur Erwerbstätigkeit kleiner). Die Ergebnisse des Risikomodells können somit als Ergänzung zu den Analysen in Abschnitt 4.3 verstanden werden, welche die individuelle Verlaufsebene fokussierten.

Wie bereits erwähnt besitzt der *Eintrittszustand* (Leistungsbezüge und Erwerbstätigkeit) selber bereits hohen Erklärwert. Referenzkategorie bilden hierbei Verläufe, die mit einem Bezug bei der ALV starten. Wenn dieser mit einer Erwerbstätigkeit einhergeht, dann reduziert sich die vorhergesagte Distanz zur Erwerbstätigkeit bereits um einen Wert von -16. Alle übrigen Zustände gehen mit höherer Distanz zur Erwerbstätigkeit einher. Besonders hoch ist der Wert bei einem Eintritt über Invalidenrente (+68), Invalidenrente und Erwerb kombiniert (+67), bei Leistungskombinationen ohne ALV (+66) bzw. mit Erwerb (+64).

Ein weiterer bedeutsamer Faktor bezüglich der erwarteten Distanz zur Erwerbstätigkeit stellt das *Alter* dar. Dabei zeigt sich, dass sich der Alterseffekt am besten mittels Hinzufügen eines quadratischen Terms modellieren lässt. Wie aus Abbildung 12 ersichtlich wird, ist der Zusammenhang keineswegs linear. Minimales Risiko weisen Personen mit 25.5 Jahren auf, danach nimmt das Risiko bis Anfang 40 nur leicht zu. Ab dem 40. Lebensjahr steigt die vorhergesagte Distanz zur Erwerbstätigkeit jedoch markant an.

Abbildung 13: Effekt des Alters auf die vorhergesagte Distanz zur Erwerbstätigkeit



Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010

Ebenfalls ein Effekt ist bezüglich der *Nationalität* zu erkennen, wobei hier auf die EU27/EFTA-Kategorisierung zurückgegriffen wird. Schweizerinnen und Schweizer (Referenz) weisen eine geringere Distanz zur Erwerbstätigkeit auf, während Verläufe von Personen aus den EU 8-Länder bzw. ausserhalb der EU im Vergleich zu den übrigen Kategorien die grösste Distanz zur Erwerbstätigkeit aufweisen.

Erstaunlicherweise sind auch im *kantonalen* Vergleich Unterschiede zu erkennen. Personen, die in Uri mit dem System der sozialen Sicherheit in Kontakt kommen, weisen im Mittel eine geringere Distanz zur Erwerbstätigkeit auf (-5). Dies trifft auch für weitere Kantone der Innerschweiz wie Nidwalden (-4) und Zug (-3) zu. Mit einer grösseren Distanz zur Erwerbstätigkeit ist im Tessin (+10) zu rechnen.

Mit geringerer Effektstärke gehen Unterschiede bezüglich *Zivilstand* und *Geschlecht* einher. Ledige Personen weisen die geringste und getrennt Lebendende die höchste (+5.6) Distanz zur Erwerbstätigkeit auf. Der Geschlechterunterschied beträgt +2, d.h. Frauen weisen ein etwas höheres Risiko auf als Männer.

4.5 Indikatoren für die Systeme der Sozialen Sicherheit

Bestehende Reporte zum System der Sozialen Sicherheit fokussieren häufig auf eine Betrachtung im Querschnitt. So werden Quoten zu bestimmten Zeitpunkten berechnet, indem die Zahl der Leistungsbeziehenden in Bezug zur Gesamtbevölkerung gesetzt wird (beispielsweise Arbeitslosenquote, Sozialhilfequote). Bisher wenig Berücksichtigung findet die Betrachtung im Längsschnitt. Dazugehören Aussagen zur Dauer von Bezügen, insbesondere von Kombinationsbezügen, wie es erst mit einer gesamtsystemischen Sicht möglich ist, und Aussagen zur Zahl beobachteter Wechsel.

Deswegen werden in diesem Abschnitt mögliche Indikatoren vorgestellt, auf deren Grundlage ein Monitoring mit Indikatoren im Längsschnitt ergänzt werden könnte. Berechnungsgrundlage sind die Verlaufsanalysen der vorangehenden Abschnitte. Berechnet werden Verweildauer- und Mobilitätsindikatoren sowie die mittlere Distanz zur Erwerbstätigkeit. Die *Verweildauerindikatoren* geben jeweils an, wie viele Quartale sich Personen im Mittel über einen bestimmten Zeitraum in einem der 16 definierten Zustände befanden (beispielsweise wie viele Erwerbsquartale bzw. Quartale mit Leistungsbezügen beobachtet werden konnten).

Mobilitätsindikatoren wiederum stehen für die beobachteten Bewegungen. Die mittlere *Zahl der Statuszustände* gibt an, wie viele unterschiedliche Zustände während eines bestimmten Zeitraums erfahren werden, während die mittlere *Zahl der Statuswechsel* Aufschluss zur generellen Zahl der Wechsel gibt. Falls etwa viele Wechsel zwischen denselben Zuständen erfolgen (beispielsweise ALV und Erwerb), so fällt die Zahl der Statuswechsel höher aus wie die Zahl der Statuszustände. Der *mittleren Distanz zur Erwerbstätigkeit* kann schliesslich entnommen werden, wie nahe bzw. wie fern Verläufe vom Arbeitsmarkt verlaufen.

Die hier präsentierten Ergebnisse beziehen sich auf die Eintrittskohorte 2006, die sich – analog zu den bisher gezeigten Auswertungen – auf einen Beobachtungszeitraum von vier Jahre beziehen. Andere Zeiträume wären denkbar. Anhand der gezeigten Indikatoren lassen sich die Bedeutung einzelner Leistungssysteme sowie von Kombinationsbezügen schnell erfassen. Somit werden Gruppen- und intertemporale Vergleiche möglich. Hier ist der Vergleich beispielhaft für Subgruppen nach Eintrittssystemen vorgenommen.

4.5.1 Verlaufsindikatoren

Im Vergleich zeigt sich, dass der Zustand „Keine Sozialleistungen, nicht erwerbstätig“ bei Eintritten über die Sozialhilfe im Mittel häufiger vorkommt, während dies bei Eintritten über IV-Rente am seltensten ist. Gleichzeitig ist die mittlere beobachtete Dauer der Erwerbstätigkeit bei Eintritten über ALV markant höher. Im Übrigen zeigt sich, dass die mittlere Dauer von Leistungskombinationen über vier Jahre betrachtet eher kurz ausfällt, was darauf hindeutet, dass diese eher von überbrückender Bedeutung sind.

Interessant wäre es nun zu sehen, wie sich die Werte einer späteren Eintrittskohorte von den hier gezeigten unterscheiden. Damit könnten strukturelle Veränderungen der Systeme der Sozialen Sicherheit sichtbar gemacht werden.

Tabelle 27: Verlaufsindikatoren für Gesamtsystem und nach Eintrittssystemen

Eintritt über	Alle Systeme	ALV	IV	SH
Zustand	Mittlere Verweildauer in Quartalen; Mittelwert (Stdabw.)			
Keine Sozialleistungen, nicht erwerbstätig	1.79 (3.47)	1.74 (3.4)	0.83 (2.77)	2.38 (3.95)
Erwerbstätig	8.22 (5.57)	9.79 (4.84)	0.12 (0.99)	4.36 (5.03)
Erwerbstätig + alle möglichen Kombination	0.01 (0.18)	0.00 (0.09)	0.02 (0.27)	0.03 (0.35)
Erwerbstätig, ALV + weitere Leistungskombi	0.01 (0.16)	0.00 (0.13)	0.032 (0.38)	0.00 (0.14)
ALV, Sozialhilfe und Erwerb	0.04 (0.32)	0.03 (0.29)	0.00 (0.02)	0.14 (0.62)
IV-Rente und Erwerb	0.40 (2.19)	0.03 (0.52)	4.87 (6.56)	0.11 (0.95)
Sozialhilfe und Erwerb	0.50 (1.82)	0.13 (0.80)	0.00 (0.08)	2.58 (3.60)
IV-Taggelder und Erwerb	0.05 (0.57)	0.02 (0.38)	0.01 (0.17)	0.02 (0.34)
ALV und Erwerb	1.23 (2.07)	1.55 (2.23)	0.00 (0.09)	0.21 (0.84)
ALV	1.83 (2.25)	2.30 (2.31)	0.01 (0.16)	0.40 (1.19)
ALV und Sozialhilfe	0.08 (0.51)	0.07 (0.45)	0.00 (0.07)	0.34 (1.02)
IV-Taggelder	0.12 (0.91)	0.03 (0.46)	0.01 (0.25)	0.06 (0.59)
ALV + weitere Leistungskombi	0.02 (0.28)	0.01 (0.23)	0.12 (0.80)	0.01 (0.20)
Sozialhilfe	0.88 (2.76)	0.21 (1.16)	0.01 (0.21)	4.58 (5.04)
Leistungskombi (ohne ALV)	0.04 (0.45)	0.01 (0.16)	0.11 (0.86)	0.18 (0.91)
IV-Rente	0.79 (3.15)	0.06 (0.68)	9.85 (6.81)	0.56 (2.23)

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010

Bemerkungen: Dargestellt sind Mittelwerte und in Klammer die Standardabweichung

4.5.2 Mobilitätsindikatoren und Distanz zur Erwerbstätigkeit

Den Mobilitätsindikatoren lässt sich entnehmen, wie „bewegt“ Verläufe im Mittel ablaufen. Damit kann Aufschluss zur generellen Komplexität von Verläufen erlangt werden. Je höher die Werte ausfallen, desto weniger nachhaltig scheinen die gefundenen Lösungen zu sein, so dass ein erneuter Wechsel nötig wird. Wie aus Tabelle 28 ersichtlich wird, erreichen Verläufe bei Eintritt über die Sozialhilfe die höchste beobachtete Komplexität. Es folgen Verläufe mit Eintritten über ALV. Die Distanz zur Erwerbstätigkeit zeigt, dass ALV-Verläufe näher am Arbeitsmarkt verlaufen als Verläufe mit Eintritten über Sozialhilfe und IV.

Tabelle 28: Mobilitätsindikatoren und Distanz zur Erwerbstätigkeit

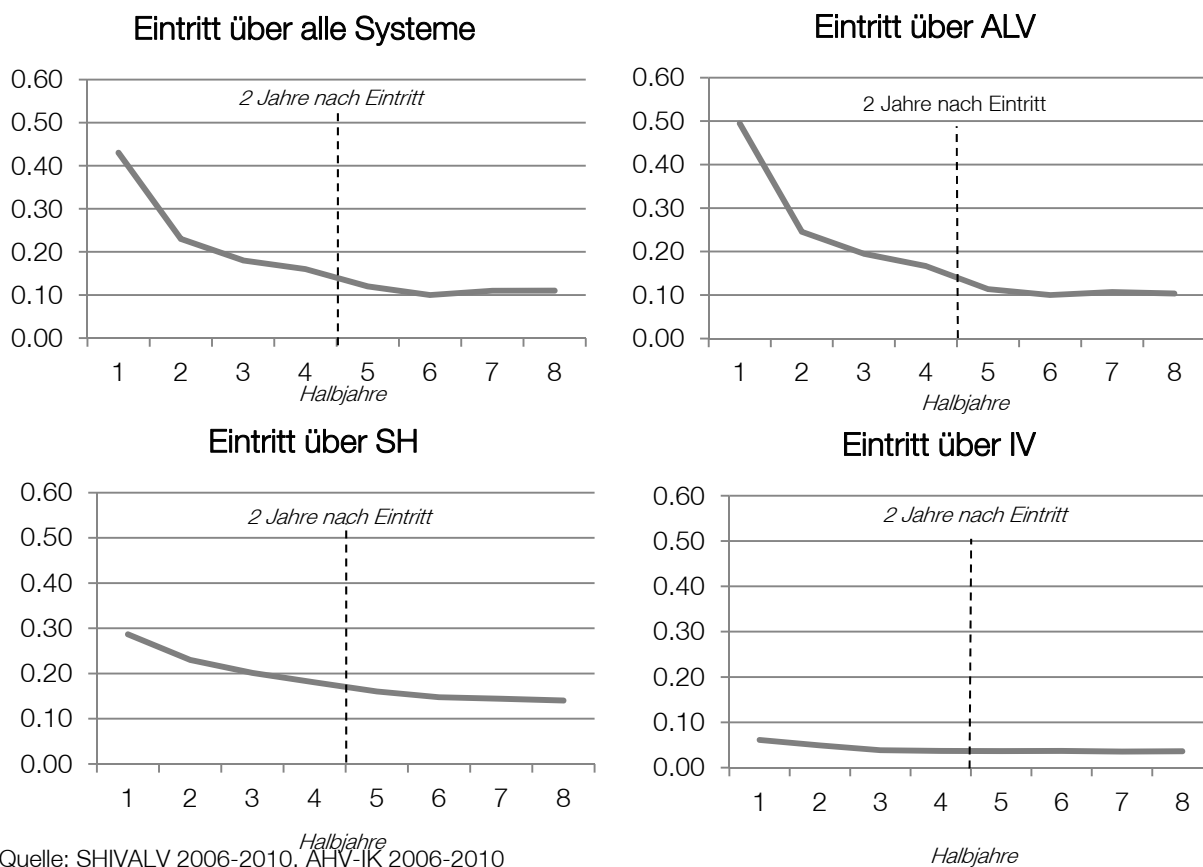
Eintritt über	Mittelwert (Stdabw.)			
	Alle Systeme	ALV	IV	SH
Anzahl Statuszustände	2.73 (0.98)	2.79 (0.87)	1.46 (0.67)	2.96 (1.21)
Anzahl Statuswechsel	2.62 (1.00)	2.75 (0.98)	0.60 (-0.00)	2.85 (1.08)
Distanz zur Erwerbstätigkeit	59.79 (44.83)	46.56 (38.06)	127.00 (7.92)	92.65 (40.45)

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010

Bemerkungen: Dargestellt sind Mittelwerte und in Klammer die Standardabweichung

Vertiefende Auswertungen zeigen, wie sich die Zahl der Wechsel über den Beobachtungszeitraum von vier Jahren entwickelt. Abbildung 14 visualisiert das Ergebnis unterschieden nach Eintrittssystem. Gesamthaft gesehen, sind die meisten Veränderungen im ersten Jahr nach Eintritt zu beobachten. In den Jahren zwei und drei nimmt die Zahl der Wechsel stetig ab, bis eine Stabilisierung mit sehr wenigen Wechseln im vierten Jahr zu beobachten ist. Die Komplexität von Verläufen gemessen an der Zahl der Wechsel unterscheidet sich je nach Eintrittssystem markant. ALV-Verläufe sind vor allem im ersten Jahr nach Eintritt mit vielen Wechseln verbunden, danach scheint sich eine gewisse Konsolidierung einzustellen, die sich in einer Abnahme der Bewegungen innerhalb der Verläufe zu späteren Zeitpunkten äussert.

Abbildung 14: Zahl der Zustandswechsel nach Eintritt in der zeitlichen Entwicklung



Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010

Bemerkungen: y-Achse = Mittlere Zahl der Zustandswechsel im Vergleich von Quartal t zu t+1, x-Achse = Zeit nach Eintritt (Halbjahresschritte)

Bewegungen von Verläufen mit Eintritten über die Sozialhilfe verteilen sich etwas gleichmässiger über den Beobachtungszeitraum, wenn auch bei dieser Gruppe erhöhte Bewegungsintensität im ersten Jahr festgestellt werden kann. Diese erreicht allerdings das Niveau der ALV-Verläufe nicht. Schliesslich sind IV-Verläufe wesentlich weniger bewegt.

4.6 Weiterführende Analysen

Die vorgestellten Auswertungen basieren auf Analysemethoden, die wir aus der Vielzahl der vorhandenen Methoden der modernen Statistik ausgewählt und hinsichtlich ihres Gewinnes an Erkenntnis am besten eingestuft haben. Mit der Anwendung dieser ausgewählten Methoden ist der in den Administrativdaten liegende volle Informationsgehalt allerdings noch nicht vollends sichtbar gemacht worden.

Erweiterungspotential sehen wir an drei Stellen:

- Erstens ist als Ergänzung zu den Risikogruppen auf Grundlage der Regressionsmodelle (vgl. Abschnitt 4.3) der Einsatz von *Assoziationsanalysen* möglich. Dies ist eine Korrelationsanalyse, die mit mehr als zwei Merkmalen durchgeführt werden kann. Dabei werden Konstellationen gesucht, die häufig zusammen mit bestimmten Verläufen auftreten. Mit dieser Auswertungsmethode wird das Gewicht stärker auf die Bedeutung von Merkmalskombinationen gelegt. Allenfalls könnten damit zusätzlich spezielle Konstellationen ermittelt werden, die typischerweise gewissen Risikogruppen angehören.
- Zweitens können Auswertungen zur Rolle von Veränderung der Situation von Personen und damit zur Bedeutung weiterer Ereignisse während des beobachteten Zeitraums von vier Jahren vorgenommen werden. Bisher fokussieren die Analysen zur Erkennung von Risikomerkmale und -gruppen auf Merkmale des Leistungsbezugs und der Erwerbsbeteiligung, auf Ereignisse bezüglich derselben sowie auf die Situation der Personen bei Eintritt in das System der Sozialen Sicherheit. Die weiteren Analysen könnten Auswertungen zu *Veränderungen von Personenmerkmalen* beinhalten, beispielsweise die Veränderung des Haushaltstyps durch die Geburt eines Kindes.
- Drittens wären mit Fokus auf Verläufe mit Eintritt über die ALV *Auswertungen unter Einbezug der AVAM-ASAL-Daten* möglich, die Aussagen zu Bedeutung von gesprochenen Massnahmen (Weiterbildung, Umschulung und Beschäftigung) zulassen.

Im Rahmen des Berichts an die Forschungskommission der Berner Fachhochschule, welcher den Abschluss des Projekts Data Mining darstellt, werden die Arbeiten an den oben erwähnten weiteren Analysen weitergeführt.

5 Entwicklung eines Monitoring-Tools

In diesem Kapitel wird das Auswertungstool beschrieben, das im Rahmen des Projekts Data Mining zuhanden des SECO entwickelt worden ist. Dieses besteht in einem sogenannten Cube auf Basis der SQL Server Analysis Services 2008 R2, mit welchem die im Projekt aufbereiteten Datengrundlagen aus einer Datenbank abgefragt werden können. Dadurch ergeben sich Auswertungsmöglichkeiten, die für ein Monitoring der Risikostruktur der ALV-Beziehenden geeignet sind.

In Abschnitt 5.1 werden zuerst die verwendeten Datengrundlagen und aufbereiteten Variablen beschrieben und anschliessend in Abschnitt 5.2 die möglichen Auswertungen erläutert.

5.1 Datengrundlagen und Variablen

Zur Erstellung des Auswertungstools, im Folgenden kurz Cube genannt, wurden im Kern dieselben Originaldatensätze verwendet wie für die Erstellung der bisher dargestellten Analysen:

- SHIVALV 2005 bis 2010
- IK-Daten 2005 bis 2010
- Separate AVAM/ASAL bzw. PLASTA/SIPAC-Daten 2005 bis 2010

Im Verlauf der bisherigen Analysen wurden die Originaldatensätze jedoch bereits zu Arbeitsdatensätzen aufbereitet. Diese wurden in einer relationalen Datenbank gespeichert und als Basis für den Cube genutzt. Es sind dies die folgenden Datensätze:

- Episodendatensatz: Darin werden Merkmale jeder Bezugs-(bzw. Nichtbezugs-)episode einer einzelnen Person definiert.
- Verlaufsdatensatz Eintrittskohorte 2006: Darin wird für alle Neubeziehenden des Jahres 2006 (mit vorangehendem zweijährigem Nichtbezug) der Verlaufstyp bestimmt und es werden Indikatoren des Verlaufs berechnet.
- Personendatensatz: Darin sind für alle Personenmerkmale der in den Originaldatensätzen vorkommenden Personen enthalten, die aus den Registerdaten der Sozialen Sicherheit ermittelbar sind. Dabei werden die Merkmale aus dem Verlaufsdatensatz übernommen für die Eintrittskohorte 2006. Für spätere Eintrittskohorten ist zudem eine Schätzung der Distanz zur Erwerbstätigkeit aufgrund der Eintrittsmerkmale enthalten.

In den nachfolgenden Tabellen wird beschrieben, welche Merkmale aus den einzelnen Datensätzen in den Cube übernommen wurden, und welche zusätzlichen Recodierungen zur Übernahme notwendig waren. Aus Sicht des Cubes handelt es sich bei den Merkmalen der Teile 1 bis 3 um so genannte Fakten (eine Anzahl quantifizierender, numerischer Daten). Die in Teil 4 beschriebenen Personenmerkmale werden im Cube als so genannte Dimensionen abgebildet, welche die Fakten charakterisieren. Die Fakten können bei der Analyse anhand der Dimensionen aggregiert und/oder aufgeschlüsselt werden. Die in Teil 1 angesprochenen Typologien sind teilweise ebenfalls als Dimension abgebildet, welche ein zugehöriges Merkmal mit der Anzahl Personen charakterisiert.

Die Merkmale in Teil 1 werden als retrospektiv bezeichnet, da sie in Quartal 17 des Verlaufs von Neueintritten gemessen werden und sich daher auf die vergangenen 16 Quartale beziehen.

Teil 1 – Retrospektive Merkmale

<i>Merkmale</i>	<i>Zeitliche Verortung</i>	<i>Zeitraum der Aggregation</i>	<i>Datensatz</i>
Typologie(n) - Typologie_alleSys (grob) - TypologieALV - TypologieIV - TypologieIVT - TypologieSH Bei gleichzeitigem Eintritt über mehrere Systeme gibt es eine Hierarchie: ALV-SH-IV-IVT <div style="border: 1px solid black; padding: 2px; display: inline-block; margin-left: 20px;">Detail- typ</div>	QuartalNummer ¹¹ =17	- Quartal - Jahr	Verlaufsdatensatz Eintrittskohorte 2006
Verlaufs- und Mobilitätsindikatoren - ALVdauer2 (Quartale ALV) - ALVERdauer (Quartale ALV und Erwerb) - kombi2dauer (Quartale ALV + weitere Leistungskomb) - kombi2ERdauer (Quartale Erwerbstätig, ALV + weitere Leistungskombi) - IVdauer2 (Quartale IV-Rente) - IVdauerER (Quartale IV-Rente und Erwerb) - ALVSHdauer (Quartale ALV und Sozialhilfe) - ALVSHERdauer (Quartale ALV, Sozialhilfe und Erwerb) - kombi1dauer (Quartale Leistungskombi (ohne ALV)) - kombiERdauer (Quartale Erwerbstätige + Leistungskombi (ohne ALV)) - SHERdauer (Quartale Sozialhilfe und Erwerb) - SHdauer2 (Quartale Sozialhilfe) - IVTdauer2 (Quartale IV-Taggelder) - IVTERdauer (Quartale IV-Taggelder und Erwerb) - NoRecdauer (Quartale keine Sozialleistungen, nicht erwerbstätig) - ERdauer2 (Quartale Erwerbstätig) - epinum (Anzahl Episoden) - elenum (Anzahl unterschiedlicher Zustände)	QuartalNummer=17	- Quartal - Jahr	Verlaufsdatensatz Eintrittskohorte 2006

Teil 2 – Deskriptive Auswertung laufender Bezüge

<i>Merkmale</i>	<i>Zeitliche Verortung</i>	<i>Zeitraum der Aggregation</i>	<i>Datensatz¹²</i>
Wie lange sind Personen durchschnittlich in einem System (ALV, SH, IV, IVT), in den vorangehenden 24 Monaten <i>Ergänzung:</i> - Ausgehend von Monat/Quartal=0 wird der Zeitraum -1 bis und mit -24 betrachtet - Keine Werte vor Quartal 1 2007	ab 1. Quartal 2007	Quartal und Jahr	Episoden
Anzahl Personen in System + Differenzierung nach Systemen (ALV,SH,IV und IVT)	ab 1. Quartal 2005	Pro Jahr und Quartal	Personen
Eintritt in System 0->1 + Differenzierung nach Systemen (ALV,SH,IV)	ab 1. Quartal 2007	Pro Jahr und Quartal	Episoden

¹¹ QuartalNummer bezieht sich auf die relative Zeit aus dem Verlaufsdatensatz.

¹² Bei der Beobachtungseinheit „Episoden“ können einzelne Personen mehrmals in der gleichen Berechnung vorkommen.

und IVT)			
Neueintritte in das Gesamtsystem: ¹³ (a) über alle Systeme und (b) nach Systemen (ALV, SH, IV, IVT) <i>Ergänzung:</i> – Keine Werte vor Quartal 1 2007	ab 1. Quartal 2007	Pro Jahr und Quartal	Episoden
Austritt aus System 1->0 + Differenzierung nach Systemen (ALV,SH,IV und IVT)	QuartalNummer=2	Pro Jahr und Quartal	Episoden
Wechsel in anderes System -> ein System 1->0 und eine 1 in einem anderen System a ALV->SH b ALV ->IVT c ALV->IV d ALV->Erwerb e ALV->Alles Nichts (alles 0)	QuartalNummer=2	Pro Jahr und Quartal	Episoden

Zur Erstellung der in Teil 3 enthaltenen Risikoprognose vgl. Abschnitt 4.4 und Anhang 7.1.

Teil 3 - Risikoprognose

<i>Merkmale</i>	<i>Zeitliche Verortung</i>	<i>Zeitraum der Aggregation</i>	<i>Datensatz</i>
Seqdist (Erwerbsintegrationsindex) + Differenzierung nach Systemen (ALV,SH,IV und IVT)	ab 1. Quartal 2005	Quartal und Jahr	Verlaufsdatensatz Eintrittskohorte 2006
Anhand des bestehenden Modells auf Basis von Personenmerkmalen bei Eintritt in geschätzter Wert für den Verlaufstyp + Differenzierung nach Systemen (ALV,SH,IV und IVT)	ab 1. Quartal 2005	Quartal und Jahr	Personendatensatz, Eintrittskohorten ¹⁴ 2007, 2008, 2009 und 2010

Im nachfolgenden Teil 4 sind nur harmonisierte Merkmale aufgeführt (vgl. Abschnitt 3.6.3). Es fehlen daher noch die persönlichen Merkmale aus den einzelnen Systemen (ALV: RAV-Region, Beruf, Branche, Ausbildungsstand) sowie Ereignisse bezüglich persönlicher Merkmale.

¹³ In den 24 vorangehenden Monaten ist kein Bezug aus den 4 Systemen IV, ALV oder SH erfolgt.

¹⁴ Eintrittskohorte= Ohne Bezug 24 Monate vor aktuellem Bezug. Nach 24 Monaten ist ein erneuter Neueintritt möglich

Teil 4 - Persönliche Merkmale

<i>Merkmale</i>	<i>Aggregation</i>	<i>Datensatz</i>	<i>Bemerkung</i>
Person		Personendaten satz	Wird im Zusammenhang mit der Verknüpfung der persönlichen Merkmale genutzt, steht aber nicht direkt zur Analyse zur Verfügung.
Alter zum Zeitpunkt der Erstbeobachtung	Jahr	Personendaten satz	Bleibt für eine Person konstant über alle Beobachtungen.
Alter zum Zeitpunkt der Beobachtung	Jahr	Personendaten satz	
ALV Gemeinde		Personendaten satz	Nur bei Personen mit ALV Eintrag vorhanden.
Geschlecht		Personendaten satz	
Kanton		Personendaten satz	
Nation		Personendaten satz	Staat mit Staatscodes gemäss BFS
Nationalität		Personendaten satz	CH/Ausländer

5.2 Auswertungsmöglichkeiten

Für das Data Mining Projekt wurde auf Basis der SQL Server Analysis Services ein Cube erstellt, welcher die Abfrage und Analyse von Teilen der aufbereiteten Daten erlaubt. Der Cube wird auf dem Server vmklotho.bfh.ch in der Entwicklungsumgebung der BFH TI betrieben. Der Zugriff kann über alle Programme erfolgen, die Analysis Services als Datenquelle unterstützen. Für normale Endbenutzer eignet sich Microsoft Office Excel in der Version 2007 oder höher gut als Frontend für die freie Abfrage auf dem Cube. Excel ist von Microsoft innerhalb des Business Intelligence Angebots auch explizit als Frontend für Endbenutzer vorgesehen. Neben seiner breiten Verfügbarkeit bietet Excel den Vorteil, dass die Benutzer mit einer weitgehend vertrauten Benutzeroberfläche arbeiten können, die der von Pivot-Tabellen sehr ähnlich ist. Die abgefragten Auswertungen können normal in Excel weiter bearbeitet werden und z.B. zu Grafiken aufbereitet werden.

Im Projekt wurden zwei Möglichkeiten zur Auswertung des Cubes genutzt:

- Erstens Abfragen mittels Pivot-Tabellen über die Verbindung „externe Daten“ von Excel.
- Zweitens Erstellen von Berichten mit Grafiken, u.a. Kartendarstellungen mittels SQL-Server Reporting Services.

Der Zugriff auf den Cube kann direkt mit der auf dem Server vmklotho.bfh.ch installierten Excel Version erfolgen. Da aber nur zwei Personen gleichzeitig auf dem Server angemeldet sein können, wurde auch der Zugriff auf den Analysis Services Cube innerhalb des BFH Netzes freigegeben (von aussen ist eine VPN Verbindung nötig). Für den Zugriff wird ein gültiger und entsprechend berechtigter Windowsbenutzer auf dem Server benötigt. Für den Zugriff von einem anderen Rechner aus müssen deshalb einige Vorbereitungen getroffen werden, damit die Authentifizierung gegenüber dem Server richtig funktioniert. Im Anhang werden diese Vorbereitungen sowie der eigentliche Zugriff auf dem Cube von Excel aus dokumentiert. Der Zugriff auf

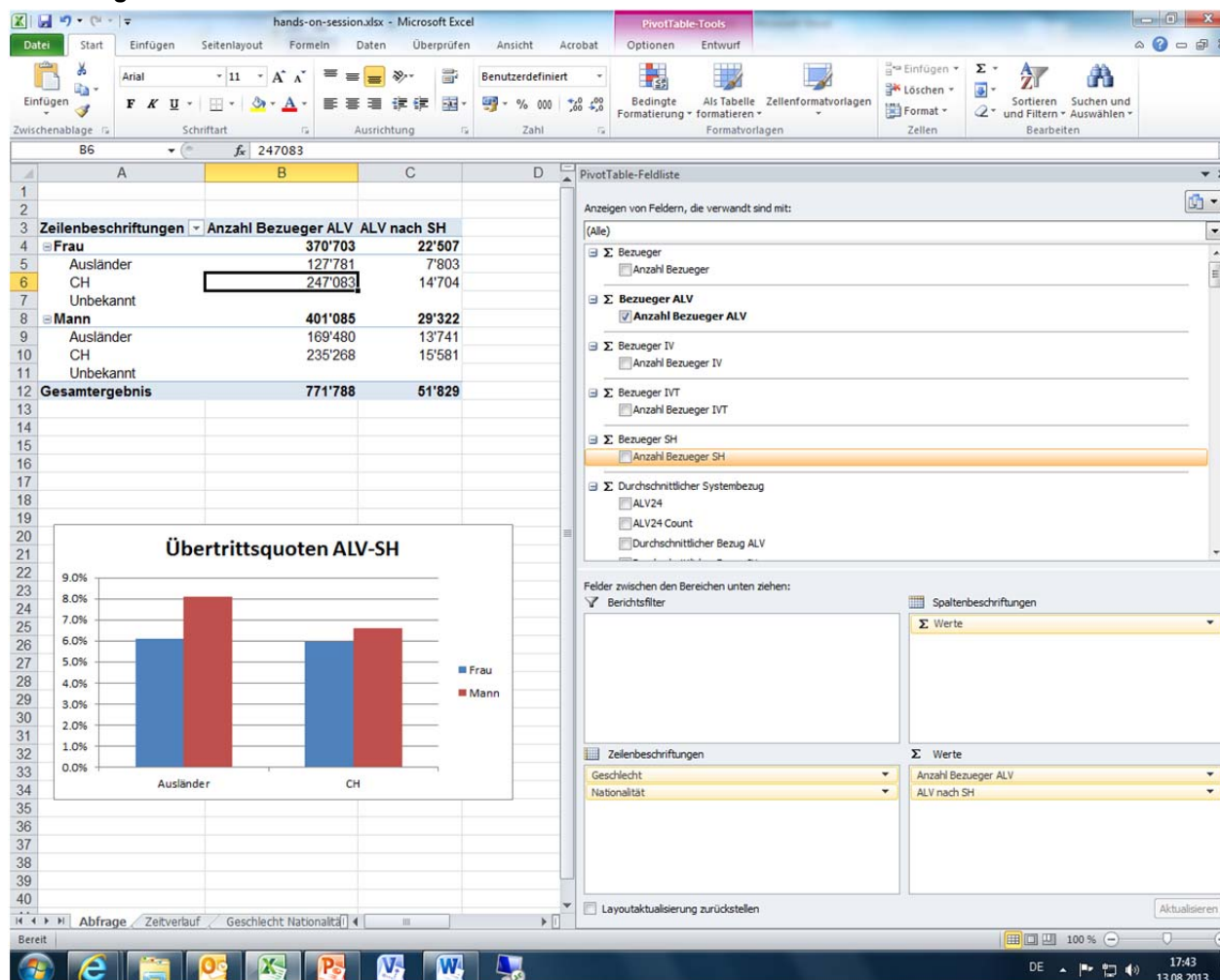
die Berichte erfolgt über den Zugriff auf die Reporting Services Webseite, die auf vmklotho.bfh.ch läuft. Auch hier ist ein gültiger Windowsbenutzer für die Authentifizierung nötig.

5.2.1 Erstellen von Pivot-Tabellen

Eine der Hauptstärken des Monitoring-tools liegt in der Gesamtsicht auf die Systeme der Sozialen Sicherheit. Damit können Übergänge zwischen den Systemen sowie in die Erwerbsarbeit mit grosser Differenziertheit beschrieben werden. Daher wird als erstes Auswertungsbeispiel hier der Übertritt von der ALV in die SH gezeigt, welcher nach den soziodemografischen Merkmalen Geschlecht und Nationalität betrachtet wird. Die dargestellte Abfrage mittels Pivot-Tabelle bezieht sich auf alle im Zeitraum 2005 bis 2010 vorhandenen Personen mit Bezug von ALV bzw. mit Wechsel von ALV zu SH. Zusätzliche Filter auf einzelne Jahre wären im auf der rechten Seite von Abbildung 10 gezeigten Auswahlmenu möglich.

Links unten in Abbildung 10 werden die in der Pivot-Tabelle dargestellten Daten in einer selbst zu gestaltenden Grafik weiterverarbeitet. Die im Beispiel gewählte Darstellung als Balkendiagramm zeigt auf, dass die Übertrittsquoten von ALV zu SH für Ausländerinnen und Schweizerinnen gleich hoch sind (6%), während sie bei Ausländern gegenüber Schweizern deutlich erhöht sind (8% vs 6.5%).

Abbildung 15: Übertritte von ALV zu SH nach Geschlecht und Nationalität



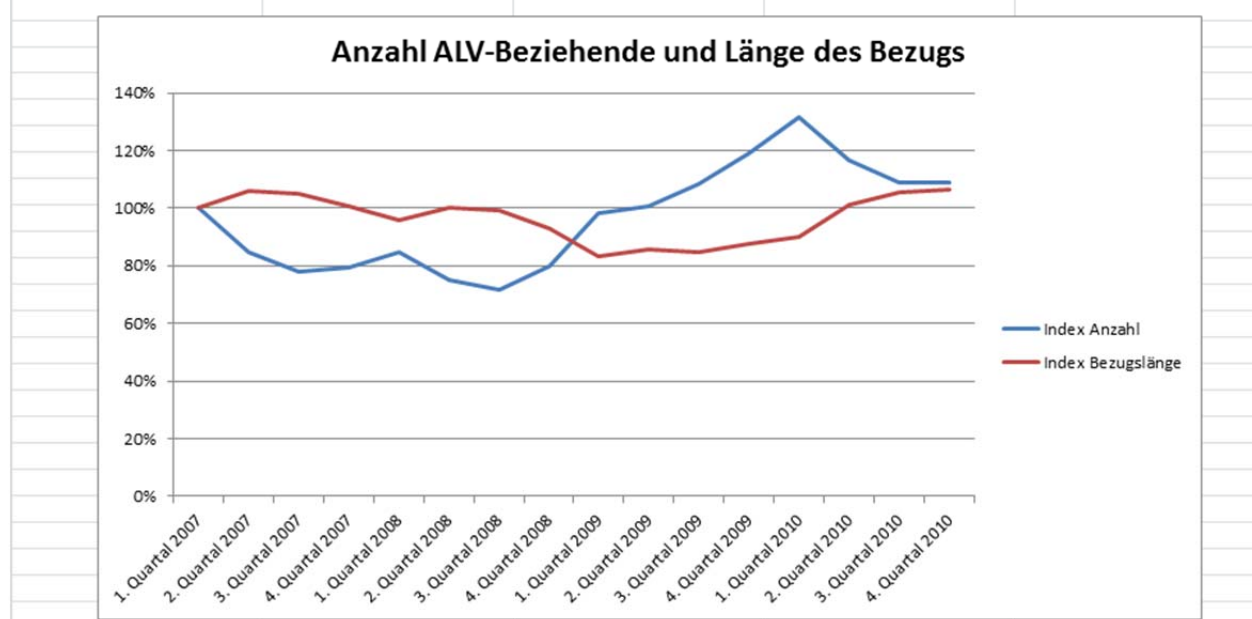
Quelle: SHIVALV 2006-2010.

Eine weitere Stärke des Monitoring-tools besteht in der gleichzeitigen Darstellung mehrerer Systemindikatoren von Einzelsystemen. Als zweites Beispiel wird eine Abfrage zur Länge des ALV-Bezugs und der Anzahl der ALV-Beziehenden im Zeitraum 2007 bis 2010 gezeigt. Da diese beiden Grössen eine unterschiedliche Skala aufweisen, werden die beiden Grössen indexiert angegeben (1. Quartal 2007 = 100%). Die Bezugsdauer wird über die letzten 24 Monate/8 Quartale gemessen. Deshalb sind in Abbildung 11 für 2005 und 2006 keine Werte angegeben.

Wie im ersten Beispiel werden auch hier die mittels Pivot-Tabelle abgefragten Daten in einer selbst in Excel gestalteten Grafik weiterverarbeitet. Der konjunkturbedingte Anstieg der Anzahl ALV-Bezüger/innen ab dem letzten Quartal 2008 ist deutlich sichtbar. Hingegen zeigt sich, dass zum gleichen Zeitpunkt die Bezugslänge unter den Referenzwert des 1. Quartals 2007 fällt. In der Hochkonjunkturphasen sind also weniger Bezüger/innen mit längeren Bezügen zu beobachten, in Rezessionsphasen mehr Bezüger/innen mit kürzeren Bezugsdauern.

Abbildung 16: Anzahl ALV-Beziehende und Länge des ALV-Bezugs im Zeitverlauf

Zeilenbeschriftungen	Durchschnittlicher Bezug	Anzahl Bezueger ALV	Index Anzahl	Index Bezugslänge
1. Quartal 2007	2.62	129'342	100%	100%
2. Quartal 2007	2.78	109'695	85%	106%
3. Quartal 2007	2.74	100'880	78%	105%
4. Quartal 2007	2.64	102'653	79%	101%
1. Quartal 2008	2.5	109'652	85%	96%
2. Quartal 2008	2.62	97'035	75%	100%
3. Quartal 2008	2.6	92'503	72%	99%
4. Quartal 2008	2.43	103'052	80%	93%
1. Quartal 2009	2.19	127'188	98%	84%
2. Quartal 2009	2.24	130'234	101%	86%
3. Quartal 2009	2.22	140'193	108%	85%
4. Quartal 2009	2.29	153'654	119%	88%
1. Quartal 2010	2.36	170'271	132%	90%
2. Quartal 2010	2.65	150'834	117%	101%
3. Quartal 2010	2.76	141'142	109%	105%
4. Quartal 2010	2.79	140'840	109%	106%
Gesamtergebnis	2.51	771'788	597%	90%



Quelle: SHIVALV 2006-2010

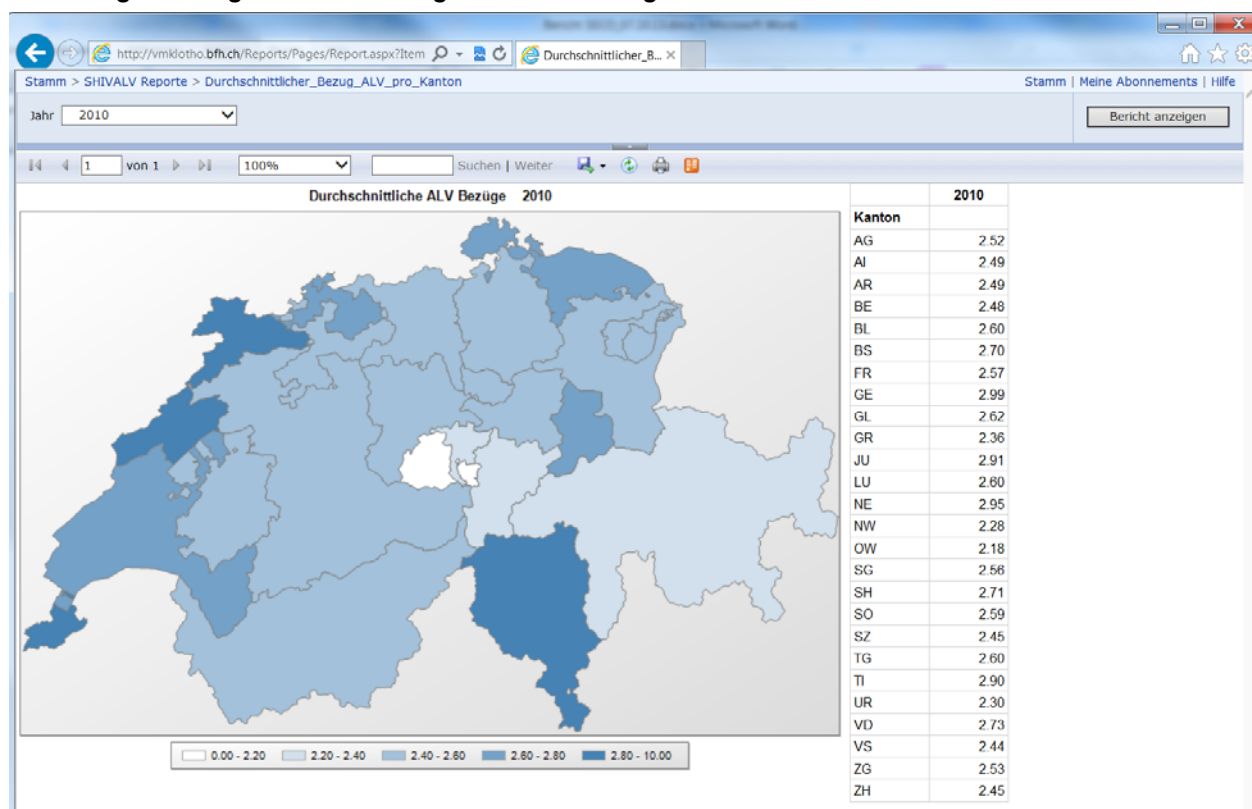
5.2.2 Berichte mit Grafiken aus SQL-Server Reporting Services

Neben der freien Abfrage des Cube mittels Excel bietet das Monitoring-tool auch die Möglichkeit der Erstellung und des Abrufs vordefinierter Berichte. Erstellt werden die Definitionen der Berichte mittels der SQL Server Reporting Services. Durch Veröffentlichung der Definitionen auf dem Webinterface der Reporting Services Veröffentlicht können die Benutzer die Berichte über einen Webbrowser abrufen. Gedacht sind die Reporting Services für die zentrale Definition und Ablage regelmässig benötigter Berichte. Diese können auch mittels Filter Parametrisiert werden. In der aktuellen Version des Monitoring Tools werden Berichte mit Grafiken und Karten definiert.

Abbildung 17 zeigt ein Beispiel mit einem Report, der die durchschnittlichen ALV Bezüge in Quartalen über die letzten 8 Quartale aggregiert auf Ebene Kanton darstellt. Die Aggregation der Daten erfolgt in Cube anhand des Kanton Merkmals. Die Daten für die Kantonsgrenzen werden von den Reporting Services aus einer relationalen Datenbank gelesen und mit den Daten aus dem Cube verknüpft. Als Ergänzung zur Karte werden die Daten in einer Tabelle dargestellt. Das im Bericht angezeigte Jahr kann über einen Filter eingestellt werden.

Abbildung 13 zeigt ein analoges Beispiel für einen Report mit einer Karte, die den Durchschnitt der Distanz zur Erwerbstätigkeit pro Kanton darstellt.

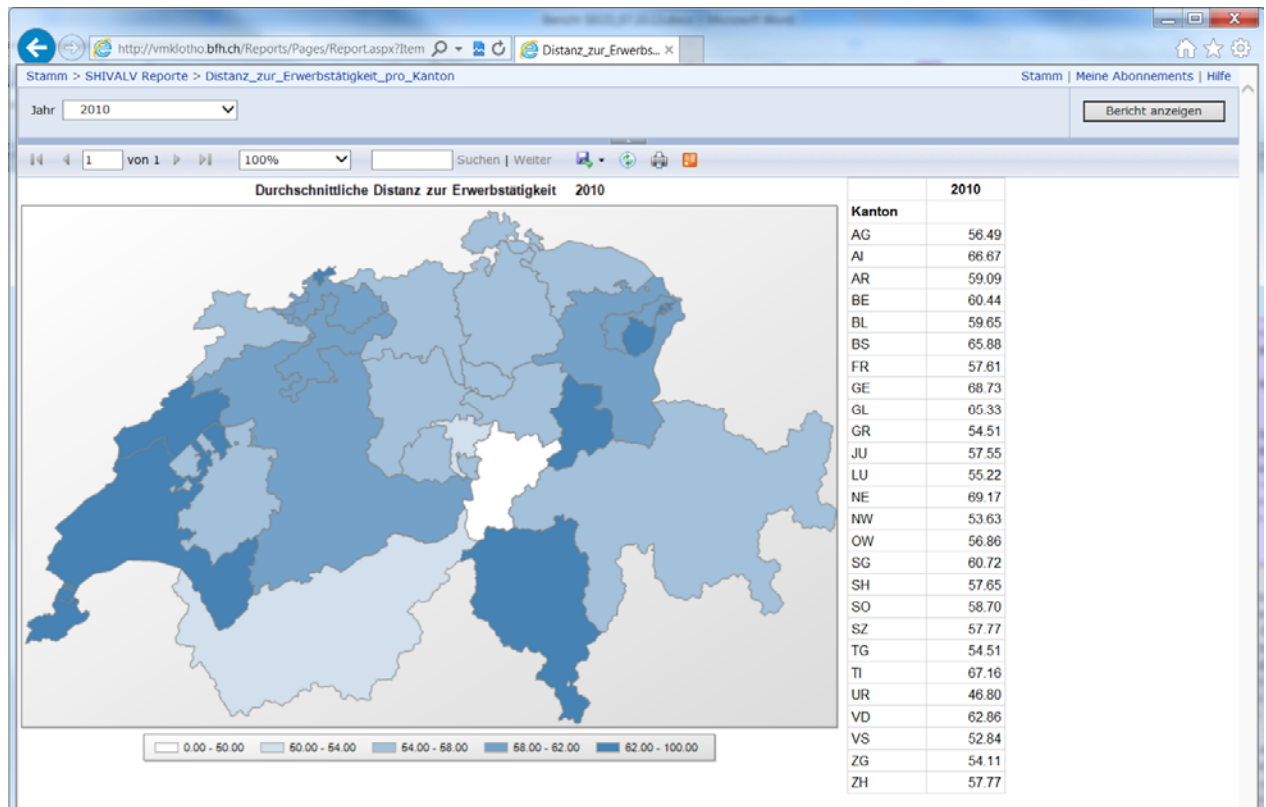
Abbildung 17: Länge des ALV-Bezugs im Kantonsvergleich



Quelle: SHIVALV 2006-2010

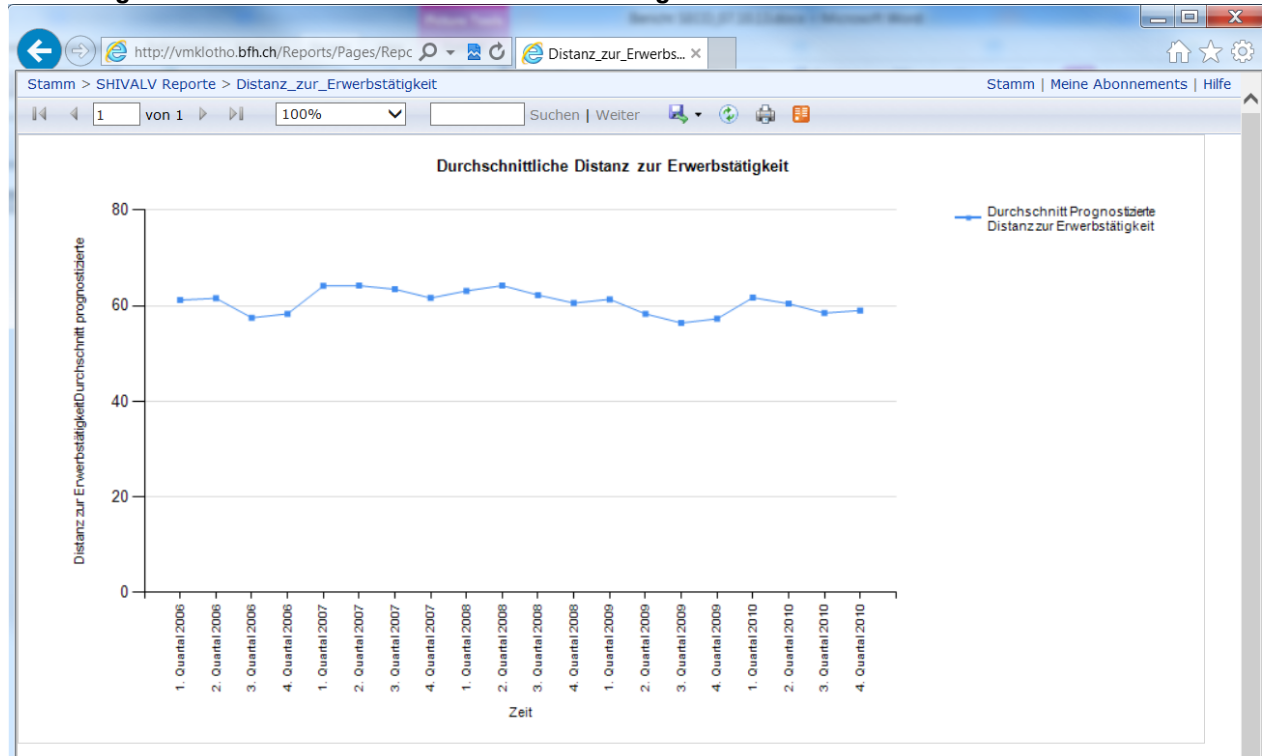
Neben der Darstellung von Karten und Tabellen erlauben die Reporting Services auch das Einbinden von verschiedenen Typen von Diagrammen. In Abbildung 19 ist ein Beispiel zu sehen, das den Gesamtdurchschnitt der Distanz zur Erwerbstätigkeit als Zeitreihe in einem Liniendiagramm darstellt.

Abbildung 18: Risikoindikator Distanz zur Erwerbstätigkeit im Kantonsvergleich



Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010

Abbildung 19: Risikoindikator Distanz zur Erwerbstätigkeit als Zeitreihe



Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010

6 Literaturverzeichnis

Backhaus, K., et al. (2000). *Multivariate Analysemethoden - Eine anwendungsorientierte Einführung*, Springer, Berlin.

Bolliger Christian, Tobias Fritschi; Renate Salzgeber, Pascale Zürcher, Oliver Hümbelin, (2012) : *Eingliederung vor Rente: Evaluation der Früherfassung, der Frühintervention und der Integrationsmassnahmen in der IV*. Bern: BSV.

Fluder, Robert, Renate Salzgeber, Tobias Fritschi, Pfiffner Roger, Oliver Hümbelin, Herbert Ruckstuhl, Urs Germann, Kilian Koch (2013): *Verläufe und Profile von neuen IV-Rentenbeziehenden 2010. Analysen anhand von SHIVALV-Daten 2005-2010*. Bern: Berner Fachhochschule.

Fluder Robert, Renate Salzgeber, Tobias Fritschi, Thomas Graf (2012): *Verläufe, Risikoprofile und Einflussfaktoren für die berufliche Integration von neuen ALV-Beziehenden (noch unveröffentlicht)*.

Fluder, R., Graf, T., Ruder, R., Salzgeber, R. (2009). *Quantifizierung der Übergänge zwischen Systemen der Sozialen Sicherheit (IV, ALV und Sozialhilfe)*, BSV, Forschungsbericht Nr. 1/09.

Morgan S.L. und C.Winship (2007): *"Counterfactuals and Causal Inference"* Cambridge: Cambridge University Press

Petersohn, H. (2005). *Data Mining: Verfahren, Prozesse, Anwendungsarchitektur*, Oldenbourg, 2005.

Salzgeber, Renate; Fritschi, Tobias; Graf, Thomas (2010): *Zeitliche Verläufe in der Sozialhilfe*. Bern: BFH.

Scherer, S. und J. Brüderl (2010): *Sequenzdatenanalyse*. In: Wolf, C. und H. Best (Hrsg.): *Handbuch der sozialwissenschaftlichen Datenanalyse*. Wiesbaden: VS Verlag.

Wiedenbeck, M. und C. Züll (2010): *Clusteranalyse*. In: Wolf, C. und H. Best (Hrsg.): *Handbuch der sozialwissenschaftlichen Datenanalyse*. Wiesbaden: VS Verlag.

7 Anhang

7.1 Rechnerische Grundlagen des Erwerbsintegrationsindex


Ausgangslage für das Bilden des Erwerbsintegrationsindex bildet die Frage, wie sich Verläufe einfach miteinander vergleichen lassen. Damit einher geht die Frage, wie sich die Komplexität eines realen Verlaufes numerisch auf sinnvolle Weise verdichten lässt. Fündig geworden sind wir bei den Methoden der Sequenzdatenanalyse, die etwa bei der Biographie- oder der DNA-Forschung Anwendung findet. Der Kerngedanke des Vergleiches besteht darin, die Distanz zwischen zwei Sequenzen zu bestimmen. Die Vorgehensweise wird nachfolgend schrittweise erläutert.

Zur **Veranschaulichung** sind dafür unten drei Sequenzen abgebildet. Buchstabe und Zahl geben jeweils den Zustand bezüglich sozialer Sicherung und Erwerbstatus für ein Quartal wieder. **Seq 0** ist eine idealtypische Sequenz mit durchgängiger Erwerbstätigkeit. **Seq 1** ist ein Verlauf, der mit einem Quartal mit ALV-Bezug beginnt und anschliessend in eine Periode mit fortdauernder Erwerbstätigkeit mündet. **Seq 2** ist eine Sequenz für einen Verlauf mit durchgehender IV-Rente.

- **Seq 0:** X1X1X1 X1X1X1 X1X1X1 X1X1X1 X1X1X1 X1X1X1 X1X1X1
- **Seq 1:** A0X1
- **Seq 2:** I0


Bemerkung: X1=Erwerb, A0=ALV-Bezug, I0=IV-Rente

Für den Erwerbsintegrationsindex werden nun alle empirisch beobachteten Sequenzen mit dem idealtypischen Verlauf einer durchgängigen Erwerbstätigkeit verglichen (Seq 0). Rechnerisch geschieht dies mit der **Levensthein-Distanz**. Dabei wird gezählt, wie viele Operationen benötigt werden, um eine Sequenz in eine andere zu überführen. Zugelassen sind **Indel**- und **Substitutionsschritte**. So werden mit Indelschritten Episoden so lange verschoben, bis sich die Sequenzen im paarweisen Vergleich entsprechen. Damit wird der Ähnlichkeit von Abfolgen Rechnung getragen. Ein Substitutionsschritt beinhaltet ein „Austauschen“ einer Sequenz, um eine paarweise Entsprechung herbeizuführen.

- **Seq 1:** A0A0X1
- **Seq 4:** A0  X1

Indelschritt

- **Seq 1:** A0A0X1
- **Seq 4:** A0A0X1

 Substitutionsschritt

Die Kosten dieser Schritte lassen sich festlegen und widerspiegeln, wie „Ähnlichkeit“ bzw. „Distanz“ vom Algorithmus gehandhabt wird. Zur **Festlegung der Substitutionskosten** haben wir die empirisch beobachtete Wahrscheinlichkeit berechnet, von einem definierten Ausgangszustand aus in die Erwerbstätigkeit zu gelangen (Übertritts-Wahrscheinlichkeit). Ausgehend von den Werten in Tabelle 29 wurden die Substitutionskosten über folgende Beziehung festgelegt:

$$\text{Substitutionskosten}_j = 10 - \frac{\text{Pr}(\text{Erwerbstätig})_j}{10}$$

Es resultiert ein Kostenvektor der Länge j , wobei j für die Anzahl unterschiedlicher Zustände steht. Die Kosten für den Übertritt von der Erwerbstätigkeit in die Erwerbstätigkeit selber werden dabei auf null gesetzt. Je „näher“ sich ein Zustand an der Erwerbstätigkeit befindet, desto geringer sind die Substitutionskosten, so ist der Zustand „ALV und Erwerb“ etwa näher an der Erwerbstätigkeit als der Zustand „IV-Rente“.

Tabelle 29: Übertritts-Wahrscheinlichkeit, Zielzustand; Erwerbstätig

<i>Status_{t0}</i>	<i>Pr(Erwerbstätig)_{t0+1}</i>
ALV	21.09
ALV und Erwerb	37.88
ALV + weitere Leistungskombi	1.23
Erwerbstätig, ALV + weitere Leistungskombi	2.23
ALV und Sozialhilfe	5.40
ALV, Sozialhilfe und Erwerb	12.00
IV-Rente	0.01
IV-Rente und Erwerb	0.39
Sozialhilfe	2.09
Sozialhilfe und Erwerb	15.03
IV-Taggelder	4.26
IV-Taggelder und Erwerb	13.93
Leistungskombi (ohne ALV)	0.23
Erwerbstätig + alle möglichen Kombination	0.42
Keine Sozialleistungen, nicht erwerbst.	10.78
Erwerbstätig	93.17

Die **Festlegung der Indelkosten** ist in Abstimmung der Substitutionskosten geschehen. Fay, Kohler und Luniak (2006: 450) empfehlen die Indelkosten mindestens halb so hoch anzusetzen, wie den höchsten Substitutionskostenwert, wenn die Position eines Elementes innerhalb einer Sequenz von Bedeutung ist (im vorliegenden Fall wäre dies eine 5). Falls die relative Position wichtiger ist und ergo Verläufe mit ähnlichen Abfolgen näher beieinanderzuliegen kommen sollen, dann empfehlen die Autoren, die Indelkosten auf 1/10 des höchsten Substitutionskostenwerts anzusetzen (im vorliegenden Fall wäre dies gleich eins). Unsere Testauswertungen haben ergeben, dass Indelkosten von vier zu einer sinnvollen Differenzierung der berechneten Distanzvariable führen.

Der **Needleman-Wunsch Algorithmus** sucht schliesslich die minimal mögliche Distanz. Ergebnis des Rechnungsschrittes ist ein Distanzmass, das die Komplexität eines einzelnen Verlaufes in einer Zahl verdichtet. Hohe Werte stehen dabei für Verläufe die „weit“ entfernt von der Erwerbstätigkeit sind, während tiefe Werte für Verläufe „nahe“ am Arbeitsmarkt stehen.

Für die eingangs erwähnten Beispielsequenzen resultieren folgende Distanzwerte (Dist)

- **Seq 0:** X1X1X1 X1X1X1 X1X1X1 X1X1X1 X1X1X1 X1X1X1 X1X1X1 => Dist=0
- **Seq 1:** A0X1 => Dist=8
- **Seq 2:** I0 => Dist=128

Bemerkung: X1=Erwerb, A0=ALV-Bezug, I0=IV-Rente

Basis für die Übertrittswahrscheinlichkeit mit Fokus auf Übertritte in die Erwerbstätigkeit (vgl. Tabelle 29) ist die unten dargestellte Transitionsmatrix, die Übertrittswahrscheinlichkeiten für alle erfassten Zustände beinhaltet.

Tabelle 30: Transitionsmatrix

		A0	A1	kombi2	kombi2_er	C0	C1	I0	I1	S0	S1	T0	T1	kombi1	kombi1_er	X0	X1	
1	ALV	A0	57.99	10.86	0.18	0.01	0.53	0.08	0.03	0.01	0.56	0.15	0.17	0.02	0	0	8.32	21.09
2	ALV und Erwerb	A1	10.33	49.39	0.02	0.06	0.15	0.27	0	0.02	0.07	0.25	0.02	0.06	0	0	1.48	37.88
3	Leistungskombi (ohne ALV)	kombi2	6.99	0.61	58.74	3.5	0.19	0	15.78	4.63	0	0.05	5.81	0.66	0.28	0.05	1.47	1.23
4	Erwerbstätig, ALV + weitere Leistungskombi	kombi2_er	2.23	5.41	8.44	51.59	0	0.16	1.43	21.34	0	0	1.43	4.94	0.32	0.16	0.32	2.23
5	ALV und Sozialhilfe	C0	13.53	2.33	0.16	0	47.65	6.03	0.02	0.01	13.94	8.24	0.03	0.01	0.55	0.06	2.01	5.4
6	ALV, Sozialhilfe und Erwerb	C1	3.68	10.89	0	0.05	10.97	37.76	0	0.03	3.02	20.85	0	0.05	0.03	0.19	0.48	12
13	IV-Rente	I0	0.01	0	0.14	0.01	0	0	97.07	1.49	0.01	0	0.03	0.00	0.24	0.01	0.97	0.01
14	IV-Rente und Erwerb	I1	0	0	0.35	0.2	0	0	4.56	94.12	0	0	0.01	0.04	0.03	0.11	0.19	0.39
17	Sozialhilfe	S0	0.17	0.02	0.01	0	1.08	0.1	0.43	0.02	81.33	7.39	0.07	0.00	1.41	0.05	5.84	2.09
18	Sozialhilfe und Erwerb	S1	0.26	0.28	0	0	1.33	1.18	0.04	0.13	8.9	71.08	0	0.02	0.1	0.46	1.19	15.03
19	IV-Taggelder	T0	2.94	0.47	0.99	0.04	0.07	0	0.38	0.2	0.67	0.07	76.84	4.25	0.47	0.06	8.29	4.26
20	IV-Taggelder und Erwerb	T1	1.18	1.2	0.22	0.57	0.04	0.02	0.06	1.39	0.12	0.31	6.04	72.71	0.04	0.39	1.79	13.93
20	ALV + weitere Leistungskombi	kombi1	0.02	0	0.67	0.05	0.48	0.07	20.72	0.51	6.04	0.81	4.36	0.23	63.38	1.71	0.74	0.23
21	Erwerbstätig, ALV + weitere Leistungskombi	kombi1_er	0	0	0.1	0.52	0.21	0.52	3.46	19.71	1.36	5.45	0.84	4.40	7.23	55.66	0.1	0.42
23	Keine Sozialleistungen, nicht erwerbst.	X0	2.9	0.46	0.01	0	0.1	0.02	0.56	0.04	1.87	0.25	0.22	0.01	0.02	0	82.77	10.78
24	Erwerbstätig	X1	1.8	1.89	0	0	0.04	0.03	0.01	0.09	0.11	0.27	0.01	0.02	0	0	2.56	93.17

7.2 Auswertung¹⁵ der Clusterergebnisse anhand der für die Clusteranalyse verwendeten Merkmale

¹⁵ Gezeigt werden Screenshots der SPSS-Modeller Auswertung. Auf der linken Seite ist für alle Inputvariablen die relative Verteilung unterschieden nach Cluster dargestellt. Auf der rechten Seite sind die Unterschiede über die Mediane sowie die 25% und 75% Perzentile gezeigt (metrische Variablen). Bei kategorialen Variablen sind unterschiedliche Auftretenshäufigkeiten über die Größe der Kreise zu erkennen. Die Variablenbezeichnungen sind im Abschnitt 5.1 (Teil 1) beschrieben.

Eintritte über ALV

Cluster

Bedeutsamkeit der Eingabe (Prädiktor)

1.0 0.8 0.6 0.4 0.2 0.0

Cluster	Cluster-2	Cluster-3	Cluster-1
Beschriftung			
Beschreibung			
Größe	47.4% (40337)	46.4% (39601)	6.2% (5264)
Input	(min) elenum 	(min) elenum 	(min) elenum
	(min) epinum 	(min) epinum 	(min) epinum
	(min) lengthA0 	(min) lengthA0 	(min) lengthA0
	(min) lengthA1 	(min) lengthA1 	(min) lengthA1
	(min) lengthC0 	(min) lengthC0 	(min) lengthC0
	(min) lengthC1 	(min) lengthC1 	(min) lengthC1
	(min) lengthI0 	(min) lengthI0 	(min) lengthI0
	(min) lengthI1 	(min) lengthI1 	(min) lengthI1
	(min) lengthkombi1 	(min) lengthkombi1 	(min) lengthkombi1
	(min) lengthkombi2 	(min) lengthkombi2 	(min) lengthkombi2
	(min) lengthkombi2_er 	(min) lengthkombi2_er 	(min) lengthkombi2_er
	(min) lengthS0 	(min) lengthS0 	(min) lengthS0
	(min) lengthS1 	(min) lengthS1 	(min) lengthS1
	(min) lengthT0 	(min) lengthT0 	(min) lengthT0
	(min) lengthT1 	(min) lengthT1 	(min) lengthT1
	(min) lengthX0 	(min) lengthX0 	(min) lengthX0
	(min) lengthX1 	(min) lengthX1 	(min) lengthX1
	(min) seqdist 	(min) seqdist 	(min) seqdist
	(min) zustand5 	(min) zustand5 	(min) zustand5
	(min) zustandend 	(min) zustandend 	(min) zustandend
	(min) zustandstart 	(min) zustandstart 	(min) zustandstart
	(min) lengthkombi1_er 	(min) lengthkombi1_er 	(min) lengthkombi1_er

Cluster-Vergleich

Cluster-2 Cluster-1 Cluster-3



7.3 Variablen aus einzelnen Systemen

ASAL/AVAM-Daten		
ASAL_Monat	ALV-BEZUG-MONATE (Januar-Dezember; 1=JA, 2=NEIN)	"XXXXXXXXXXXX" (Jan-Dez; 1=JA, 2=NEIN)
RHFristMIN	ALV-BEGINN RAHMENFRIST (YYYYMM) - Minimum	Jahr (YYYY), Monat (MM)
RHFristMAX	ALV-BEGINN RAHMENFRIST (YYYYMM) - Maximum	Jahr (YYYY), Monat (MM)
ausgest	ALV-AUSGESTEUERT	1=JA, 0=Nein
sbn2000	ALV-BERUF SBN2000 (AVAM)	Beruf gemäss Berufsnomenklatur 2000
nogacd	ALV-BRANCHENCODE NOGA (AVAM)	Branchencode gemäss NOGA
ausbild	ALV-AUSBILDUNGSSTUFE (AVAM)	Ausbildungscodes AVAM (Beilage)
ALV_Taggeld	ALV-Bruttotaggeld in CH (Jahr)	ALV-Taggeld in CHF pro Jahr
PID_SECO	PID_SECO	ID seco für weitere Verknüpfungen
ALV_GMDE	ALV-Gemeindecode BFS	
IV-Daten		
IV_Monat	IV-BEZUG-MONATE (Januar-Dezember; 1=JA, 2=NEIN)	"XXXXXXXXXXXX" (Jan-Dez; 1=JA, 2=NEIN)
wohn_IV	IV-Wohnort	"7777777777" = unbekannt
detat	IV-REGISTEREINTRAG (1. Rentenzahlung; YYYY)	11=Schweiz, 12=Ausreise Ausland, 21=Einreise Schweiz
pinv	IV-INVALIDITÄTSGRAD (in %)	Grad in %
cfrt	IV-RENTENTEIL (Ganze, halbe, viertel - Rente)	1=Ganze Rente, 2=Halbe Rente, 3=DreiViertel, 4=Viertel
cinf	IV-GEBRECHENSCODE	Liste der Gebrechenscode
mpr	IV-Rente in CHF (Monat)	IV-Rente des letzten Bezugsmonats
SH-Daten		
(gültige Daten für SHIVALV: Antragsteller, 18-65 J., mit ID=OK, ohne Personen im Flüchtlings, Asylwesen Variable SH_v0405 < 14)		
SH_Monat	SOZIALHILFE-BEZUG-MONATE (Januar-Dezember; 1=JA, 2=NEIN)	"XXXXXXXXXXXX" (Jan-Dez; 1=JA, 2=NEIN)
v1506_eff	SH-BEZUGSBEGINN (YYYYMM)	Jahr (YYYY), Monat (MM)
UE_typ_kind	SH-Unterstützungseinheit Typ	1=Alleinlebende 2=Paare ohne Kind 3=Alleinerziehende mit 1 Kind 4=Alleinerziehende mit 2 Kinder 5=Alleinerziehende mit 3+ Kinder 6=Paare mit 1 Kind 7=Paare mit 2 Kinder 8=Paare mit 3+ Kinder 9=Andere 10=Nicht-Alleinlebende -3=Missing
P_UE	SH-ANZAHL PERSONEN IN Unterstützungseinheit	Anzahl Personen
P_HH	SH-ANZAHL PERSONEN IN HAUSHALT insgesamt	Anzahl Personen
SH_Betrag	SH-Betrag in CHF (Jahr)	Sozialhilfebezug in CHF (Jahr), Missing=fehlender Wert
v0101_commune	FALLBEARBEITUNG - GEMEINDE BFS	BFS_Code
v0206_bfs	WOHNSITZ (ZIVILRECHTLICH) - GEMEINDE BFS	BFS_Code
v0301_bfs	AUFENTHALTSORT - GEMEINDE BFS	BFS_Code
v07011	ERWERBSSITUATION - ANTRAGSTELLER	Fragebogen SH
v0703	BESCHÄFTIGUNGSGRAD - ANTRAGSTELLER	Fragebogen SH
v0710	LETZTER BERUF - ANTRAGSTELLER	Berufscodes 8-Steller (Beilage)
v0711	BRANCHE - ANTRAGSTELLER	Liste SH-Statistik (Beilage)
v0713	AUSBILDUNG - ANTRAGSTELLER	Fragebogen SH

Quelle: SHIVALV



7.4 Ergebnis Parameterschätzung Risikomodell

In Tabelle 31 sind die für das Risikomodell verwendeten Merkmale aufgelistet sowie Effekt-Schätzer (Coef.), Standardabweichung, t-Wert, der daraus ermittelte p-Wert und das dazugehörige 95%-Konfidenzintervall.

Tabelle 31: Risikomodell alle Eintrittssysteme (Zielvariable Distanz zur Erwerbstätigkeit)

	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
Kanton (Referenz: Zürich)¹⁶					Untere	obere Grenze
Bern	1.03	0.40	2.59	0.009	0.25	1.80
Uri	-5.12	2.13	-2.4	0.016	-9.30	-0.94
Nidwalden	-4.24	1.86	-2.29	0.022	-7.88	-0.61
Zug	-3.06	0.97	-3.15	0.002	-4.97	-1.16
Freiburg	2.17	0.63	3.47	0.001	0.94	3.40
Basel-Stadt	5.71	0.62	9.17	0	4.49	6.92
Basel-Landschaft	1.59	0.62	2.58	0.01	0.38	2.80
Appenzell	5.72	3.04	1.88	0.06	-0.23	11.67
St.Gallen	2.49	0.51	4.87	0	1.49	3.49
Tessin	10.48	0.54	19.26	0	9.42	11.55
Waadt	4.19	0.41	10.33	0	3.39	4.98
Neuenburg	6.10	0.65	9.41	0	4.83	7.37
Genf	10.83	0.45	24.2	0	9.95	11.71
Geschlecht (Referenz: Männer)						
Frauen	2.13	0.21	10.01	0	1.71	2.54
Zivilstand (Referenz: ledig)						
Verheiratet	2.66	0.29	9.28	0	2.10	3.22
Verwitwet	2.52	1.19	2.11	0.035	0.18	4.86
Geschieden	2.95	0.42	7.06	0	2.13	3.77
Getrennt	5.63	0.78	7.24	0	4.10	7.15
Alter						
Alter	-1.05	0.06	-16.41	0	-1.18	-0.92
Alter Quadrat	0.02	0.00	24.67	0	0.02	0.02
Nationalität EU/ EFTA 27 (Referenz: Schweiz)						
EU 17/EFTA angrenzend	7.19	0.36	19.94	0	6.49	7.90
EU 17/EFTA Nord	10.34	1.20	8.61	0	7.99	12.69
EU 17 Süd	5.91	0.42	14	0	5.08	6.74
EU 2	7.79	2.82	2.76	0.006	2.25	13.32

¹⁶ Kantone ohne statistisch signifikanten Unterschied sind nicht dargestellt: Luzern, Schwyz, Obwalden, Glarus, Solothurn, Schaffhausen, Graubünden, Aargau, Thurgau, Wallis und Jura



EU 8	13.19	1.56	8.48	0	10.14	16.24
Rest der Welt	13.30	0.30	45.01	0	12.72	13.88

Zustand bei Eintritt (Referenz: ALV)

ALV und Erwerb	-16.07	0.25	-65.11	0	-16.55	-15.58
kombi2	47.79	3.50	13.67	0	40.94	54.64
kombi2_er	55.88	8.88	6.3	0	38.48	73.28
ALV und SH	25.59	1.02	24.97	0	23.58	27.59
ALV, SH und Er	11.54	1.63	7.09	0	8.35	14.73
IV	67.60	0.56	120.45	0	66.50	68.70
IV und Erwerb	67.46	0.70	95.69	0	66.08	68.84
SH	51.72	0.39	132.01	0	50.95	52.49
SH und Erwerb	20.68	0.49	41.9	0	19.72	21.65
IVT ¹⁷	62.99	0.87	72.66	0	61.29	64.69
IVT und Erwerb	45.09	1.45	31.14	0	42.25	47.93
kombi1	66.15	3.92	16.87	0	58.47	73.84
kombi1_er	64.14	8.10	7.92	0	48.26	80.02
Konstante	52.39	1.14	45.83	0	50.15	54.64

Number of obs	=	109819
F(51,109767)	=	1508.22
Prob > F	=	0
R-squared	=	0.41
Adj R-squared	=	0.41
Root MSE	=	34.37

Quelle: SHIVALV 2006-2010, AHV-IK 2006-2010, AVAM/ASAL 2005-2010

Bemerkungen

Zustand bei Eintritt

kombi2= ALV + weitere Leistungskombi, kombi2_er= Erwerbstätig, ALV + weitere Leistungskombi, kombi1= Leistungskombi (ohne ALV), kombi1_er= Erwerbstätig + alle möglichen Kombination von Leistungsbezügen

Nationalität EU-27/EFTA

EU 17 angrenzend: Deutschland, Frankreich, Italien, Österreich und Liechtenstein

EU 17 Nord: Belgien, Dänemark, Finnland, UK, Irland, Island, Luxemburg, Niederlande, Norwegen, Schweden

EU 17 Süd: Griechenland, Malta, Portugal, Spanien, Zypern

EU 2: Bulgarien, Rumänien

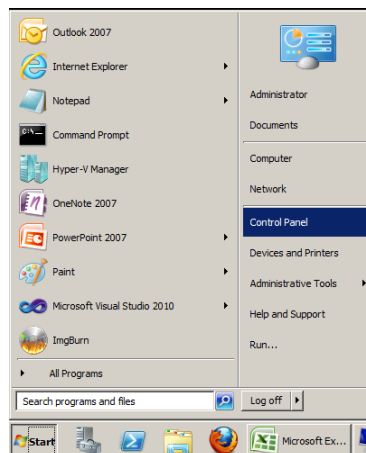
EU 8: Estland, Lettland, Litauen, Polen, Slowakei, Slowenien, Tschechien, Ungarn

¹⁷ Sollte nicht interpretiert werden, weil die bestehende Datenlage nicht alle Eintritte über IV-Taggeld beinhaltet.

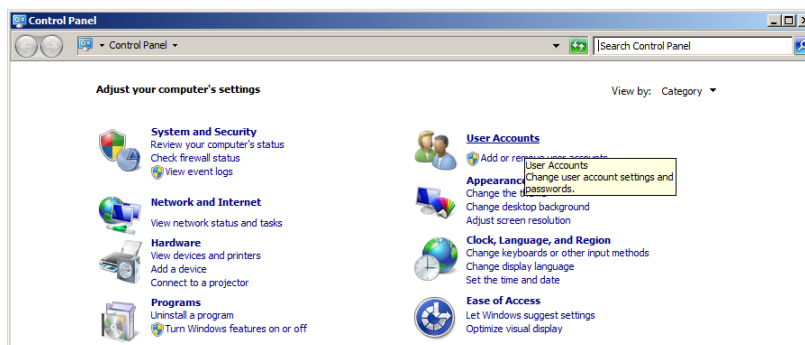
7.5 Anleitung zur Einrichtung des CUBE-Zugriffs

Der Zugriff auf den Analysis Services Server auf vmklotho.bfh.ch ist mittels Windows Authentifizierung gesichert. Es wird ein gültiger Windowsbenutzer auf dem Server vmklotho.bfh.ch benötigt, der mindestens Leserechte auf dem Cube besitzt, um auf diesen zugreifen zu können. Da der Server vmklotho.bfh.ch nicht in der BFH Domäne ist, müssen auf dem Computer, von dem aus der Zugriff auf den Cube erfolgt, Vorbereitungen getroffen werden, damit die Windows Authentifizierung gegenüber dem Server korrekt erfolgt. Eine Variante ist, dass der verwendete Benutzer auf dem zugreifenden Computer den gleichen Benutzernamen und das gleiche Passwort wie der Serverbenutzer hat. Die ist meist aber nicht möglich. Deshalb müssen in den meisten Fällen die Anmeldeinformationen des Benutzers auf vmklotho.bfh.ch im Credential Manager auf dem zugreifenden Computer hinterlegt werden. Dadurch werden beim Zugriff auf vmklotho.bfh.ch nicht die Anmeldeinformationen des angemeldeten Benutzers sondern die hinterlegten Informationen des Serverbenutzers verwendet. Nachfolgend wird Schritt für Schritt gezeigt, wie die Anmeldeinformationen hinterlegt werden können. Jeder Benutzer muss dies nur einmal tun. Die nachfolgenden Schritte beziehen sich auf Windows 7, sollen mit kleinen Abweichungen aber auch auf Windows XP und Windows 8 umsetzbar sein.

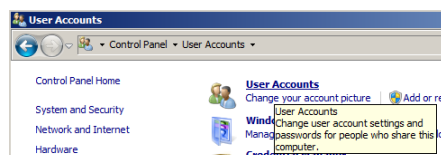
1. Als erstes muss die *Systemsteuerung* auf dem lokalen Rechner geöffnet werden.



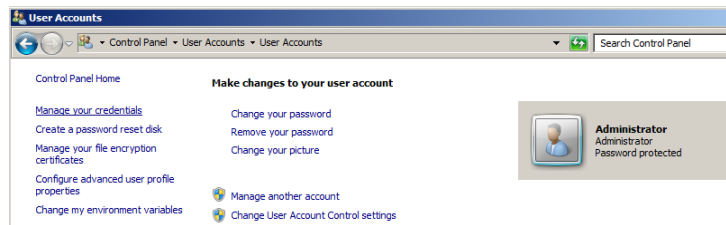
2. In der Systemsteuerung wählen wir Punkt *Benutzerkonten* aus.



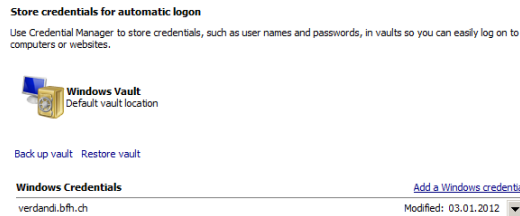
3. In der nun angezeigten Ansicht klicken wir erneut auf den Punkt *Benutzerkonten*.



4. Es sollten die Einstellung des aktuell angemeldeten Benutzers angezeigt werden. Unter den Optionen auf der linken Seite klicken wir auf den Punkt *Eigene Anmeldeinformationen Verwalten*.



- Es wird die Verwaltung der gespeicherten Anmeldeinformationen angezeigt. Zuerst steht eine Liste mit den Windows Anmeldeinformationen. Hier klicken wir auf die Option *Windows-Anmeldeinformation hinzufügen*.



- Im nun angezeigten Dialog können wir die Anmeldeinformationen für eine bestimmte Internetseite hinterlegen können. Diese Anmeldeinformationen werden dann immer anstatt der Informationen des angemeldeten Benutzers verwendet, wenn die entsprechende Seite besucht wird. Wir geben folgende Informationen ein und klicken dann auf *OK*, um die Anmeldeinformationen zu hinterlegen.

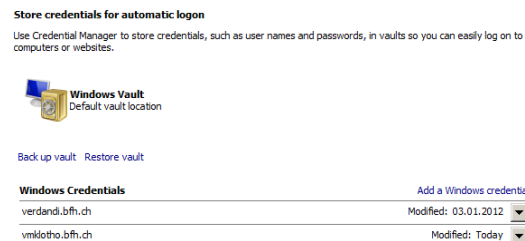
Internet oder Netzwerkadresse: vmklotho.bfh.ch

Benutzername: Anmeldeiname des Benutzers auf dem Server

Kenntwort: Passwort des Benutzers auf dem Server



- In der Liste der Windows Anmeldeinformationen wolle ein neuer Eintrag für vmklotho.bfh.ch stehen. Damit sind die Vorbereitungen abgeschlossen und die Verwaltung der Anmeldeinformationen kann geschlossen werden.



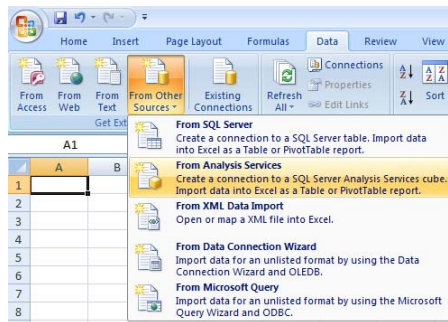
Zugriff auf den Cube mit Excel

Nachfolgend wird zuerst gezeigt, wie mit Excel eine neue Verbindung zum Cube hergestellt werden kann. Im zweiten Schritt wird gezeigt wie eine bereits gespeicherte Verbindung verwendet werden kann. Voraussetzung für den erfolgreichen Zugriff auf den Cube ist entweder, dass man auf dem Server vmklotho.bfh.ch selbst arbeitet oder die Windows Anmeldeinformationen für einen Serverbenutzer auf dem lokalen Rechner hinterlegt hat.

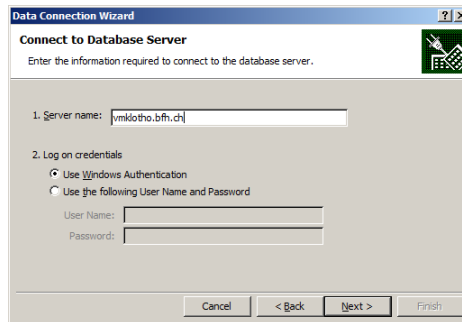
Einrichten der Verbindung

Hier wird gezeigt, wie in Excel eine neue Daten Verbindung angelegt und geöffnet wird

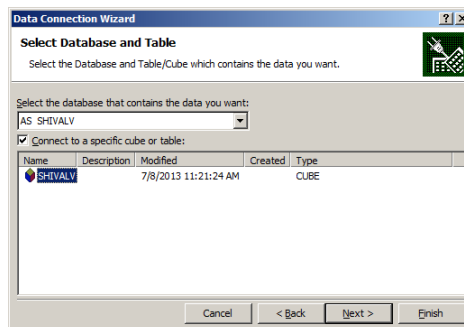
- Wir öffnen ein neues Dokument in Excel und wechseln in das *Data* Tab. Hier klicken wir auf den Punkt *From Other Sources* und im erscheinenden Menü auf *From Analysis Services*.



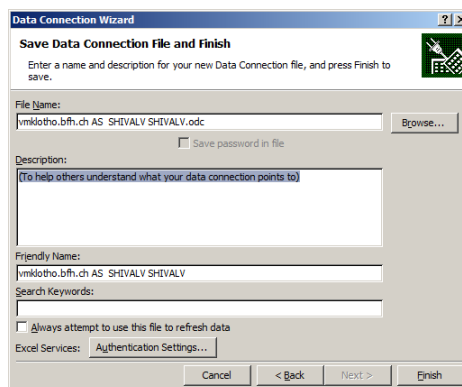
2. Es wird ein Assistent zur Einrichtung der Daten Verbindung angezeigt. Im ersten Dialog geben wir den Servernamen *vmklotho.bfh.ch* an und klicken auf *Next>*. Die Option *Windows Authentication* muss unbedingt beibehalten werden!



3. Es erscheint ein Auswahldialog für die Datenbank und den Cube, der für die Verbindung verwendet werden sollen. Hier ist die *AS SHIVALV* Datenbank und der *SHIVALV* Cube auszuwählen. Anschliessend klicken wir auf *Next>*.

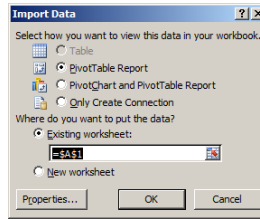


4. Im letzten Dialog des Assistenten kann noch der Name angegeben werden, unter dem die Datenquelle gespeichert wird. Optional kann auch eine Beschreibung und Schlüsselwörter hinterlegt werden. Wir übernehmen die vorgeschlagenen Werte und klicken auf *Finish*. Dadurch wird die Daten Verbindung gespeichert und nach schliessen des Assistenten direkt für die weitere Verwendung ausgewählt.

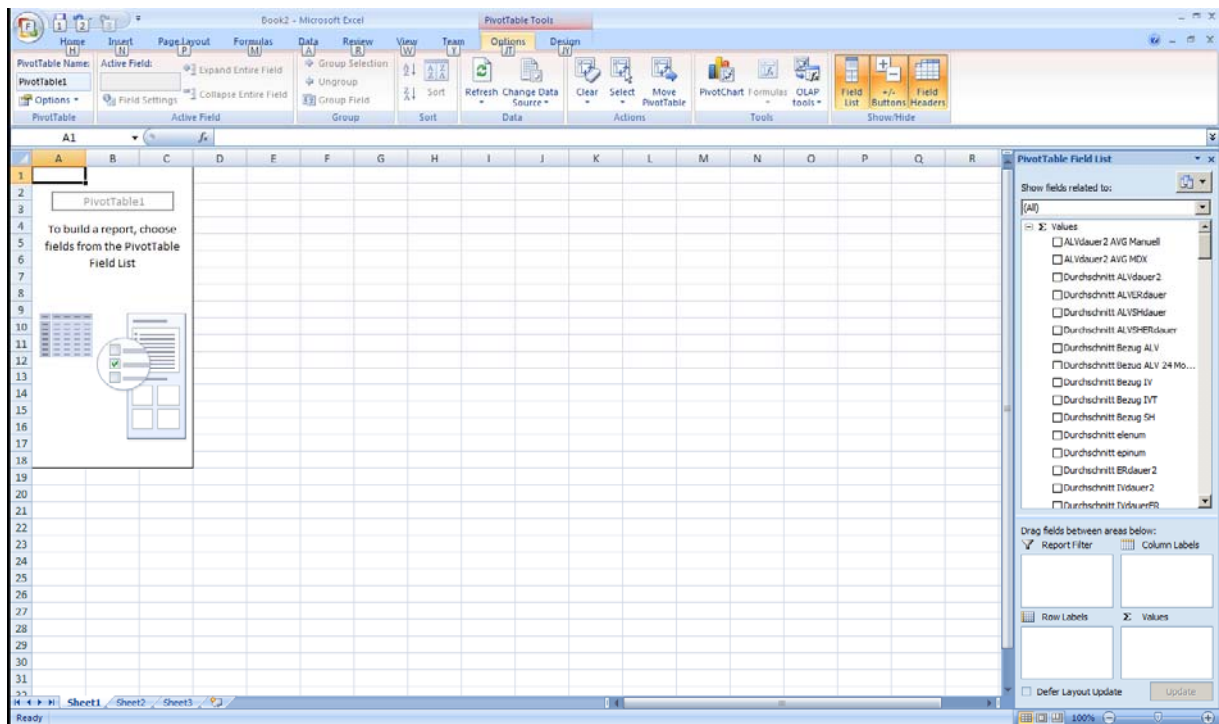


5. Excel zeigt die Optionen für die Verwendung der Daten Verbindung an. In unserem Fall wählen wir die Option *PivotTable Report*. Den Ort für das Einfügen des Reports behalten wir

bei und klicken auf *OK*.



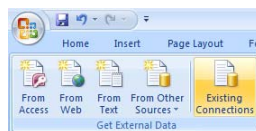
6. In Excel wurde eine Pivot Tabelle eingefügt, die auf den Daten aus dem SHIVALV Cue basiert. Über die Optionen auf der rechten Seite kann ausgewählt werden, welche Dimensionen und Measures in der Tabelle dargestellt werden und welche Dimensionen als zusätzliche Filter dienen sollen. Über das Ribbon am oberen Rand können weitere Einstellungen für die Pivot Tabelle vorgenommen und z.B. auch ein Pivot Chart ergänzt werden.



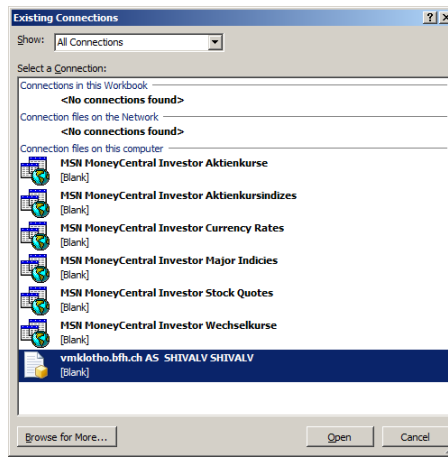
Öffnen einer bestehenden Verbindung

Wenn bereits einmal eine Daten Verbindung zum Cube erstellt und gespeichert wurde, so kann diese direkt geöffnet werden, ohne diese neu konfigurieren zu müssen.

1. Im *Data* Tab von Excel klicken wir auf die Option *Existing Connections*.



2. Es wird ein Auswahldialog mit allen am Default Speicherort gefundenen Daten Verbindungen angezeigt. Ist die Datenverbindung an einem anderen Ort gespeichert worden, kann über die Option *Browse for More* danach gesucht werden. In unserem Fall wählen wir die bestehende Verbindung für den SHIVALV Cube auf vmklotho.bfh.ch aus der Liste aus und klicken auf *Open*.



3. Wie nach der Erstellung einer neuen Verbindung wird der Dialog zur Auswahl der Verwendung angezeigt. Wir übernehmen die Option *PivotTable Report* und klicken auf *OK*, um diesen an den vorgegebenen Koordinaten einzufügen.



Tabelle 32: Tabelle mit Wertelabels von ausgewählten Variablen

Nationalität	Alter
0 CH	0 unter 23 Jahre
1 EU 17 angrenzend	1 23 bis 25
2 EU 17 Nord	2 26 bis 45
3 EU 17 Süd	3 46 bis 65
4 EU 2	
5 EU 8	
6 Welt.	
Gemeindetypologie	Beruf
1 Zentren	1 land- und forstwirtschaftliche Berufe der Tierzucht
2 suburbane Gemeinden	2 Produktionsberufe in der Industrie und im Gewerbe ohne Bau
3 einkommensstarke Gemeinden	3 technische Berufe sowie Informatikberufe
4 periurbane Gemeinden	4 Berufe des Bau- und Ausbaugewerbes und des Bergbaus
5 touristische Gemeinden	5 Handels- und Verkehrsberufe
6 industrielle und tertiäre Gemeinden	6 Berufe des Gastgewerbes und Berufe zur Erbringung persönlicher Dienstleistungen
7 ländliche Pendlergemeinden	7 Berufe des Managements und der Administration, des Bank- und Versicherungsgewerbes und des Rechtswesens
8 agrar-gemischte Gemeinden	8 Gesundheits-, Lehr- und Kulturberufe Wissenschaftler
9 agrarische Gemeinden.	9 nicht klassifizierbare Angaben
Branche	
1 Land- und Forstwirtschaft	11 Grundstücks- und Wohnungswesen
2 Verarbeitende Gewerbe; Herstellung von Waren	12 freiberufliche wissenschaftliche und technische Dienstleistungen
3 Energieversorgung	13 sonstige wissenschaftliche Dienstleistungen
4 Wasserversorgung und Abfallentsorgung	14 öffentliche Verwaltung, Verteidigung; Sozialversicherung
5 Baugewerbe, Bau	15 Erziehung und Unterricht
6 Handel, Instandhaltung und Reparatur von Motorfahrzeugen	16 Gesundheits- und Sozialwesen
7 Verkehr und Lagerei	17 Kunst, Unterhaltung und Erholung
8 Gastgewerbe, Beherbergung und Gastronomie	18 sonstige Dienstleistungen
9 Information und Kommunikation	19 private Haushalte mit Hauspersonal ohne ausgeprägten Schwerpunkt
10 Finanz- und Versicherungsdienstleistungen	20 exterritoriale Organisationen und Körperschaften

