

Von SAS nach SPSS und umgekehrt – Umwandlung von Dateien inklusive Formaten bzw. Wertelabels

Timm Lochmann, Carina Ortseifen

Universität Heidelberg
Universitätsrechenzentrum
Im Neuenheimer Feld 293
69120 Heidelberg

Timm.Lochmann@urz.uni-heidelberg.de
Carina.Ortseifen@urz.uni-heidelberg.de

Zusammenfassung

An vielen Universitäten wird für die Durchführung von statistischen Analysen sowohl die SAS als auch die SPSS Software verwendet. Manchmal sollen beide Systeme auf die gleichen Datensätze angewendet werden. Dazu ist es erforderlich, die Daten zwischen den beiden Programmen SAS und SPSS austauschen zu können. Neben kommerzieller Software wie DBMS/COPY und STAT/TRANSFER bieten die beiden Systeme selbst Prozeduren zur Konvertierung. Die verschiedenen Möglichkeiten zur Datenübertragung inklusive eines eigenen Makros werden dargelegt, insbesondere wird geprüft, ob und wie Variablen- und Wertelabels (Formate) korrekt übertragen werden.

Keywords: Datenaustausch, SAS, SPSS, Label, Format, fehlende Werte, missing value, Engine SPSS, PROC CONVERT, GET SAS, DBMS/COPY, STAT/TRANSFER, Makro %SPSS2SAS_F

1 Einleitung

Für die Durchführung von statistischen Analysen wird an der Universität Heidelberg, wie vermutlich auch an anderen Universitäten, sowohl die SAS als

auch die SPSS Software eingesetzt. Manchmal sollen beide Systeme dabei auf die gleichen Datensätze angewendet werden, um die jeweilige Stärken der beiden Programmpakete nutzen zu können.

Erfasst man die Daten dabei beispielsweise mit SPSS, möchte anschließend aber eine Analyse mit SAS durchführen, müssen die Daten zunächst konvertiert werden. Neben kommerziellen Transferprogrammen wie DBMS/COPY und STAT/TRANSFER bieten sowohl SAS als auch SPSS Möglichkeiten, diese Konvertierung durchzuführen. Von besonderem Interesse ist dabei, ob vorhandene SAS Labels und Formate bzw. SPSS Variablen- und Wertelabels korrekt übertragen werden.

Um die verschiedenen Möglichkeiten einigermaßen systematisch vergleichen zu können, werden zunächst in Abschnitt 2 die Unterschiede zwischen SAS- und SPSS-Dateien erläutert. Diese Unterschiede erlauben es, anhand geeigneter Beispiel-Datensätze im 3. Abschnitt die Übertragung von SAS nach SPSS und im 4. Abschnitt von SPSS nach SAS zu untersuchen. Am Ende finden Sie in Abschnitt 5 eine tabellarische Zusammenfassung aller Übertragungswege mit den jeweiligen Problemen. Die verwendeten Beispieldaten werden im Anhang aufgeführt.

Die folgenden Ausführungen beziehen sich, solange nichts anderes im Text vermerkt wird, auf die aktuellen Versionen der beiden Programmpakete: SPSS 10.0.7 und 10.1.3 deutsch und SAS 8.2 unter Windows NT 4.0. Die beiden Programme STAT/TRANSFER (Version 6.1, <http://www.stattransfer.com/>) und DBMS/COPY (Version 7.0, <http://www.conceptual.com/>) lagen uns nur als Demoverionen vor.

2 Unterschiede zwischen SAS 8 und SPSS 10 Dateien

Folgende Merkmale werden hinsichtlich unterschiedlicher Definition im SAS und SPSS System untersucht: Variablenamen, Variablentypen, (Variablen-) Label, Formate (Wertelabel), Datenwerte und die Darstellung fehlender Werte.

2.1 Variablenamen

In der SAS Version 8 können Variablenamen eine Länge von 1 bis 32 Zeichen haben. Das erste Zeichen muss ein Buchstabe oder ein Underscore (_) sein. Umlaute, Leer- und Sonderzeichen (mit Ausnahme von _) sind nicht erlaubt.

Variablenamen in SPSS 10 haben eine Länge zwischen 1 und 8 Zeichen. Das erste Zeichen muss ein Buchstabe sein. Ansonsten sind Buchstaben, Ziffern,

Punkte, Umlaute und folgende Sonderzeichen erlaubt: @, #, _ und \$. Nicht erlaubt sind lediglich Leerzeichen und die Sonderzeichen !, ?, ', +, - und *.

Problematisch können bei der Übertragung daher werden:

SAS → SPSS Lange Variablenamen, _ als erstes Zeichen
 SPSS → SAS Sonderzeichen und Umlaute in Variablenamen

2.2 Variablentypen

Während die SAS Software nur zwischen numerischen und Textvariablen unterscheidet, Datumsangaben, Kommazahlen etc. intern durch Formate steuert, unterscheidet das SPSS System numerische Variablen, Punkt/Komma, wissenschaftliche Notation, Datum, Währung und Text.

Problematisch können bei der Übertragung daher werden:

SPSS → SAS Datumsvariablen, Punkt/Komma, wissenschaftliche Notation, Währung.

2.3 Variablenlabel

Variablenlabel oder Label können in beiden Systemen zwischen 1 und 256 Zeichen lang sein. Alle Zeichen sind erlaubt.

Probleme sind bei der Übertragung nicht zu erwarten.

2.4 Formate / Wertelabels

Das SAS System unterscheidet die systemeigenen Formate (und Informate), wie Datumsformate oder Formate für Kommazahlen von den mit der Prozedur FORMAT erzeugten Benutzer-definierten Formaten. Alle Formate (und Informate) werden in SAS Katalogen gespeichert. Die Zuordnung eines Formats zu einer Variablen kann entweder permanent im Datenschnitt oder in einem PROC DATASETS Schritt erfolgen oder temporär in einem analytischen Prozedurschritt. Die Formatnamen sind zwischen 1 und 8 Zeichen lang. SAS unterscheidet zusätzlich noch zwischen Formaten für numerische und Formaten für Textvariablen. Die Formate weisen entweder Einzelwerten neue Werte oder einem ganzen Bereich (Range) einen neuen Wert zu. Bei der Definition von Bereichen können LOW, HIGH und OTHER verwendet werden.

Im SPSS System werden die Formate als Wertelabels bezeichnet. Sie sind direkt mit der Datei verbunden, können für die Anzeige und Analyse ein- und ausgeschaltet werden. Die Zuweisung kann nur „eins zu eins“ erfolgen, d.h. je einem Wert kann ein Wertelabel zugewiesen werden; Ranges sind somit nicht

möglich. Allerdings darf mehreren Werten das gleiche Wertelabel zugewiesen werden.

Problematisch können bei der Übertragung daher werden:

- SAS → SPSS SAS Formate im allgemeinen,
SAS Formate mit Ranges im speziellen
- SPSS → SAS Umwandlung von SPSS Wertelabels in SAS Formate.

2.5 Datenwerte

Im SAS System können numerische Variablen Werte von 2 bis 8 Bytes Länge speichern, Stringvariablen von 1 bis 32767 Bytes. Im SPSS System hingegen sind Stringvariablen auf 255 Zeichen Länge beschränkt.

Problematisch können bei der Übertragung daher werden:

- SAS → SPSS SAS Stringvariablen mit einer Länge größer als 255.

2.6 Fehlende Werte

Die beiden Programmpakete unterscheiden sich wie folgt in der Definition fehlender Werte: Im SAS System ist für alphanumerische-Variablen das Leerzeichen („ “) bzw. kein Zeichen („“) der fehlende Wert, für numerische Variablen hingegen besteht neben dem Punkt („.“) noch die Möglichkeit, zwischen verschiedenen Arten von fehlenden Werten („.A“, „.B“ . . .) zu unterscheiden. Im SPSS-System hingegen hat der Anwender die Möglichkeit, bestimmte Werte als „fehlend“ auszuzeichnen. So kann erreicht werden, dass bspw. alle Beobachtungen mit dem Wert „999“ in einer bestimmten Variable für nachfolgende Analysen als fehlend behandelt werden. Diese Beobachtungen werden somit ausgeschlossen, aber getrennt von den System-definierten fehlenden Werten ausgewiesen.

Problematisch können bei der Übertragung daher werden:

- SPSS → SAS Abbildung der SPSS-Benutzerdefinierten fehlenden Werte

2.7 Fazit

Somit werden in den nächsten beiden Abschnitten folgende Themen beleuchtet:

- *Vorgehen:* Hier wird kurz beschrieben, welche Schritte konkret für den Datenaustausch durchgeführt werden müssen.
- *Variablennamen:* Welche Auswirkungen haben die unterschiedlichen Einschränkungen für Variablennamen in SAS und SPSS und wie werden problematische Variablennamen umbenannt.

- *Datenwerte:* Werden die in den Variablen enthaltenen Werte richtig übertragen.
- *Formate/Werte-Labels:* Welche SAS-Formate und SPSS-Werte-Labels werden beim Datenaustausch zwischen den beiden Systemen übertragen. Inwiefern werden systemeigene und benutzerdefinierte Formate richtig zugewiesen.
- *Fehlende Werte:* Wie wird mit fehlenden Werten umgegangen.

3 Datenübertragung von SAS 8 nach SPSS 10

Die SPSS Software bietet drei verschiedene Möglichkeiten, SAS Dateien in das SPSS eigene Format zu überführen: Das Datei-Öffnen-Menü, der Befehl GET SAS und die Open Database Connectivity (ODBC), wobei auf letztere Möglichkeit nicht eingegangen wird. Außerdem können die beiden kommerziellen Programme DBMSCOPY und STAT/TRANSFER für die Datenumwandlung genutzt werden.

Für die Untersuchung der verschiedenen Möglichkeiten wurde eine SAS Datei mit 8 Variablen und 5 Beobachtungen künstlich erzeugt. Das zugehörige Programm finden Sie im Anhang.

3.1 Menü Datei - Öffnen

Vorgehen: Über das Menü Datei → Öffnen → Daten lassen sich durch Einstellung des entsprechenden Dateityps SAS V8 und SAS V6 Dateien für Windows (*.sd7, *.sas7bdat, *.sd2) öffnen, V6 SAS/UNIX Dateien (*.sd2) und SAS Transportdateien (*.tpt).

Variablenamen: Lange Variablenamen, wie etwa Geburtsdatum oder Geschlecht werden nach dem achten Zeichen abgeschnitten. Würden nach dem Abschneiden allerdings zwei Variablen gleichen Namens vorliegen (wenn es vorher etwa Geschlecht und Geschlecht2 gegeben hätte), wird der doppelte Variablenname ersetzt durch V+Nummer der Variablen, also z.B. V7, wenn Geschlecht2 die 7. Variable in der Datei ist.

Datenwerte: Stringvariablen werden, ohne Warnung, auf 255 Zeichen gekürzt.

Formate/Wertelabels: SAS Datumsvariablen werden auch im SPSS System korrekt als Datumsvariablen erkannt. Benutzer-eigene Formate jedoch, z.B. für das Geschlecht, werden nicht übertragen.

Anmerkung: Beim Zugriff auf SAS V6 Dateien für Windows (*.sd2) wird u.U. der SPSS-Assistent für Textimport geöffnet. Durch Umstellungen in der Systemsteuerung → Ländereinstellung (Punkt als Dezimaltrennzeichen, Komma

für die Zifferngruppierung) werden auch diese Dateien korrekt importiert. (Dieses Problem tritt mit dem Befehl GET SAS – siehe Abschnitt 3.2 – nicht auf.)

3.2 Befehl GET SAS

Vorgehen: Im Syntaxfenster können mit dem Befehl GET SAS die gleichen Dateitypen wie über das Dateimenü geöffnet werden. Der Umgang mit langen und doppelten Variablennamen sowie langen Stringvariablen ist völlig analog zum Menü. Ein Unterschied besteht darin, dass der Befehl GET SAS die Option FORMATS kennt, mit der auch SAS-Formate übertragen werden können. Dazu muss aber zunächst im SAS System der entsprechende Formatkatalog in eine SAS Systemdatei übertragen werden, was mit der Prozedur FORMAT und der Option CNTLOUT leicht zu bewerkstelligen ist:

```
LIBNAME library 'c:\fmt'; * Pfad zum Formatkatalog';
LIBNAME daten 'c:\daten'; * Pfad zu den Daten';
PROC FORMAT LIB=library CNTLOUT=daten.fmts;
RUN;
```

Im SPSS System kann dann folgender Befehl ausgeführt werden:

```
GET SAS DATA='c:\daten\daten.sas7bdat' /
      FORMATS='c:\daten\fmts.sas7bdat' .
EXECUTE.
```

Variablennamen: Analog zu 3.1 Menü Datei - Öffnen.

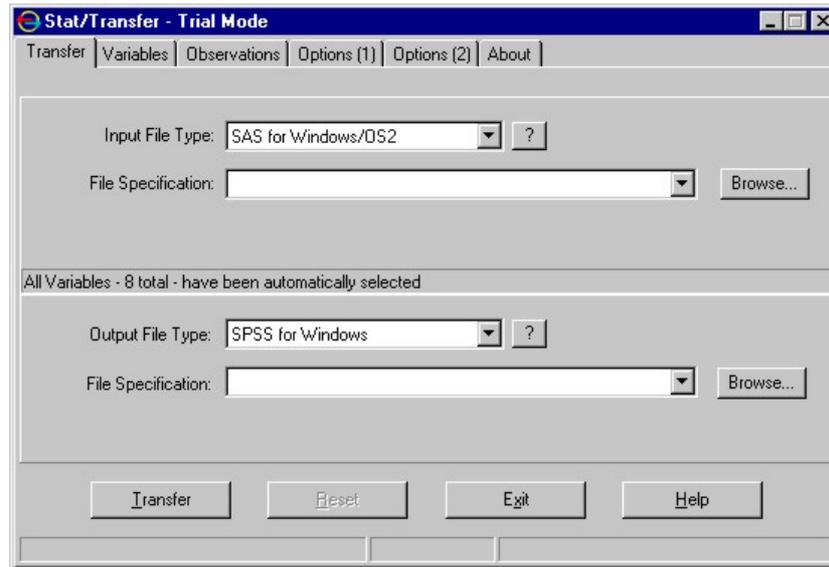
Datenwerte: Analog zu 3.1 Menü Datei - Öffnen.

Formate/Wertelabels: SAS Datumsvariablen werden korrekt als Datumsvariablen erkannt. Benutzerdefinierte Formate können wie soeben beschrieben übertragen werden. Allerdings, und hier kommt die Einschränkung, die im SPSS Syntax Guide dokumentiert ist, werden nur numerische Formate übertragen sowie Formate ohne Bereichdefinition, d.h. LOW, HIGH und OTHER werden ignoriert. Und: Bei SAS-Transportfiles wird die FORMATS-Option von SPSS ignoriert.

3.3 STAT/TRANSFER

Die kommerzielle Software STAT/TRANSFER unterstützt folgende SAS Dateitypen: V8 und V6 für Windows sowie Transportfiles. Im Unterschied zu den SPSS-eigenen Varianten werden jedoch keine SAS/UNIX-Files unterstützt.

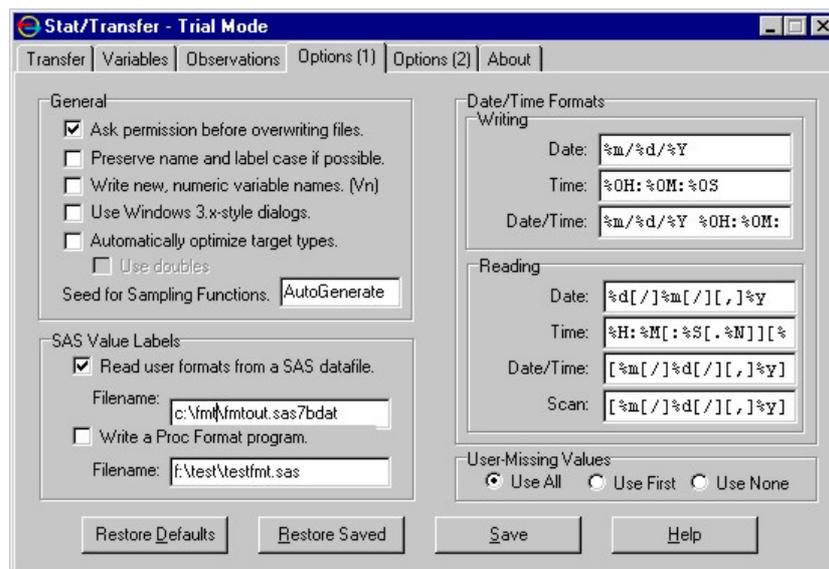
Vorgehen: Wie in folgendem Screenshot wiedergegeben werden der Typ und der Name der einzulesenden Datei und der zu erzeugenden Datei angegeben und dann der Transfer-Knopf angeklickt.



Variablenamen: Lange Variablenamen werden auf 8 Zeichen gekürzt. Ungültige Variablenamen werden um ein vorangestelltes @ ergänzt, etwa @_tgeld für den mit dem Unterstrich beginnenden Variablenamen _tgeld.

Datenwerte: Lange Stringvariablen werden auf 255 Zeichen gekürzt. Entstehen dadurch doppelte Variablenamen, so wird an den Namensstamm eine laufende Ziffer angehängt, etwa Geschle1.

Formate/Werte-Labels: Benutzereigene SAS Formate werden übertragen, sowohl numerische als auch Formate für Textvariablen. Dazu muss im Options(1) Fenster (siehe nächster Screenshot) der Name der SAS Datei genannt werden, welche die Formate enthält. (Evtl. ist vorher wie bei GET SAS ein PROC FORMAT Schritt notwendig.) Ranges können nicht übertragen werden.



3.4 DBMS/COPY

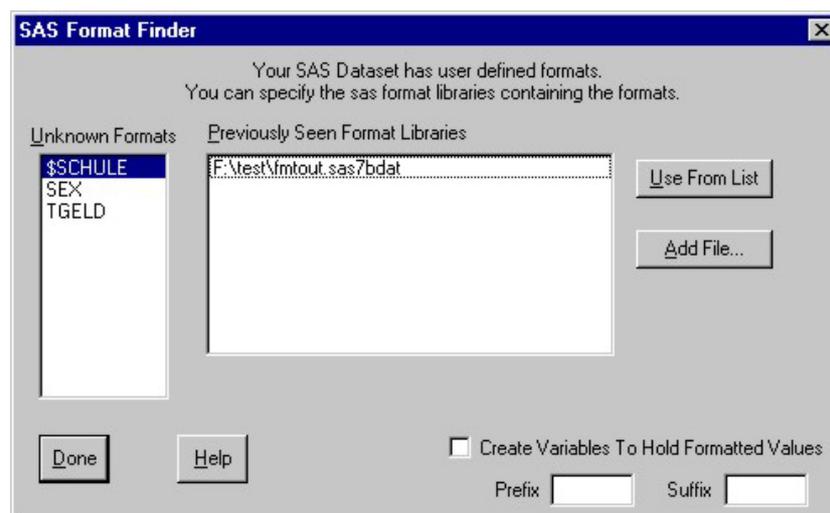
Das kommerzielle Programm DBMS/COPY unterstützt folgende SAS Dateitypen: V8 und V6 für Windows, V8 und 6 für UNIX sowie Transportfiles, auch komprimierte Transportfiles.

Vorgehen: Über einen Dialog (oder auch im Batchmodus) werden der Typ und der Name der einzulesenden Datei und der zu erzeugenden Datei abgefragt.

Variablenamen: Lange Variablenamen werden auf 8 Zeichen gekürzt. Entstehen dadurch doppelte Variablenamen, so wird an den Namensstamm eine laufende Ziffer angehängt, etwa Geschle1. Beginnt der Variablenname mit einem `_`, wird dieser durch ein `@` ersetzt (`@tgeld`).

Datenwerte: Lange Stringvariablen werden auf 255 Zeichen gekürzt.

Formate/Werte-Labels: Benutzereigene SAS Formate werden übertragen, sowohl numerische als auch Formate für Textvariablen. Nur Ranges können nicht übertragen werden. Im ‚SAS Format Finder‘-Fenster (siehe nächsten Screenshot) werden die aufgefundenen unbekanntenen Formate bzw. die Formatbibliothek, die allerdings eine SAS-Datei sein muss (siehe dazu auch 3.3), angezeigt.



3.5 SAS 6/8 nach SPSS 7.5/8/9

Die Übertragung von SAS Dateien in eine ältere SPSS Version erfordert die Erzeugung von SAS-Transportfiles mit Hilfe der Engine XPORT und dem anschließenden SPSS-Befehl GET SAS. Im SAS-Anwenderhandbuch im Netz (<http://www.urz.uni-heidelberg.de/statistik/sas-ah/>) finden Sie hierzu ein nützliches SAS-Makro von Andreas Christmann, weshalb an dieser Stelle nicht weiter auf die Konvertierung eingegangen werden soll.

4 Datenübertragung von SPSS 10 nach SAS 8

Auch für die Überführung von SPSS-Datensätzen nach SAS gibt es verschiedene Möglichkeiten, von denen hier folgende erläutert werden: Die SPSS Engine, die Prozedur CONVERT, das Makro %SPSS2SAS.F sowie die kommerzielle Software DBMS/COPY und STAT/TRANSFER.

4.1 SPSS-Engine und PROC CONVERT

Vorgehen: Mit Hilfe der SPSS-Engine kann von SAS lesend auf SPSS-Dateien zugegriffen werden. Dabei muss die SPSS-Datei im Transportformat (portableformat, *.por) abgespeichert sein.

Zunächst wird in einer LIBNAME-Anweisung auf die entsprechende Datei verwiesen, die dann mit dem zweistelligen Namen `bibliothek._first_` referenzierbar ist. Dabei ist `_first_` eine beliebige Zeichenkette.

```
LIBNAME spfile SPSS "c:/testda.por";
* 1. Prozedur PRINT greift direkt auf SPSS-Datei zu;
PROC PRINT DATA = spfile._first_;
RUN;
* 2. SPSS-Datei wird in eine SAS-Datei kopiert;
DATA testda;
  SET spfile._first_;
RUN;
PROC PRINT DATA =testda;
RUN;
```

Mit der Prozedur CONVERT lassen sich analog SPSS-Dateien in SAS-Dateien konvertieren. Intern verwendet die Prozedur CONVERT die SPSS-Engine und erzielt nahezu (s.u.) identische Ergebnisse.

```
FILENAME spfile "c:/testda.por";
PROC CONVERT SPSS=spfile OUT=testda2;
RUN;
PROC PRINT DATA=testda2;
RUN;
```

Variablennamen: Bei der Übersetzung von SPSS-Variablennamen werden alle SAS-ungültigen Sonderzeichen und Umlaute in den Unterstrich „_“ übersetzt. Werden mehrere Variablen dabei in den gleichen Namen umbenannt, so enthält die SAS-Datei nur eine Variable diesen Namens. Bei Verwendung der SPSS-Engine plus Datenschnitt mit SET ist dies die letzte Variable, die diesen Namen erhält, bei PROC CONVERT ist es die erste mit diesem Namen.

Kurioserweise können Prozeduren, die auf die referenzierte Datei direkt zugreifen, alle Variablen, auch doppelte, anzeigen (z.B. die Prozedur PRINT in obigem Programm im Schritt 1).

Formate/Wertelabels: SPSS-eigene Formate werden sowohl mit der SPSS Engine als auch mit der Prozedur CONVERT übertragen. Benutzerdefinierte Wertelabels und Wertelabels für Datumsformate werden jedoch nicht konvertiert.

Fehlende Werte: Systemeigene und benutzerdefinierte fehlende Werte werden gleichermaßen als fehlend, i.e. als „.“ übertragen. Möchte man Unterscheidungen zwischen verschiedenen Arten fehlender Werte auch im SAS treffen, so kann man in SPSS im Variablenfenster unter „Fehlende Werte“ „keine fehlenden Werte“ ankreuzen. Somit werden die benutzerdefinierten fehlenden Werte unverändert übertragen. Damit kann nachträglich eine erneute Auszeichnung dieser Werte als fehlend vorgenommen werden.

4.2 DBMS/COPY

Vorgehen: DBMS/COPY bietet einen interaktiven, sowie einen Batchmodus zur Steuerung der Datenkonvertierung. In beiden Fällen können benutzerdefinierte Wertelabels miteingelesen (oder ignoriert) werden. Für die Erzeugung der entsprechenden SAS-Formate wird von DBMS/COPY ein SAS-Programm mit einem PROC FORMAT-Schritt erzeugt, welches vor Verwendung der Formate in SAS ausgeführt werden muss.

Variablennamen: Bei Konflikten für Variablennamen werden diese in Name_1, Name_2 . . . umbenannt. Somit bleiben alle Variablen der ursprünglichen Datei erhalten.

Formate/Wertelabels: Bei der Konvertierung werden in SPSS definierte Wertelabels für Datumsvariablen (bspw. 24.12.2002 → Weihnachten) nicht übertragen. Im Falle von Konflikten bzgl. Variablennamen wird mehrmals das gleiche Format angelegt, das dann allen umbenannten Variablen gleichen Namensstamms zugeordnet wird. Hierbei können ungewollte Zuordnungen von Formaten zu Variablen erfolgen. Außerdem stimmen die mit PROC FORMAT erzeugten Formatnamen für alphanumerische Variablen nicht mit den zugeordneten Formatnamen überein. Letzteres Problem lässt sich aber dadurch lösen, dass die Namen der entsprechenden Formate im von DBMS/COPY erzeugten Programm geändert werden. Hierzu wird das zweite Zeichen der Formatnamen gelöscht (z.B. \$String → \$string).

Fehlende Werte: Von DBMS/COPY werden benutzerdefinierte und systemeigene fehlende Werte gleichermaßen als „fehlend“ übertragen. Um benutzerdefinierte fehlende Werte von systemeigenen unterscheiden zu können, muss, wie oben geschildert, in der SPSS-Datei die Option „keine fehlenden Werte“ gesetzt werden.

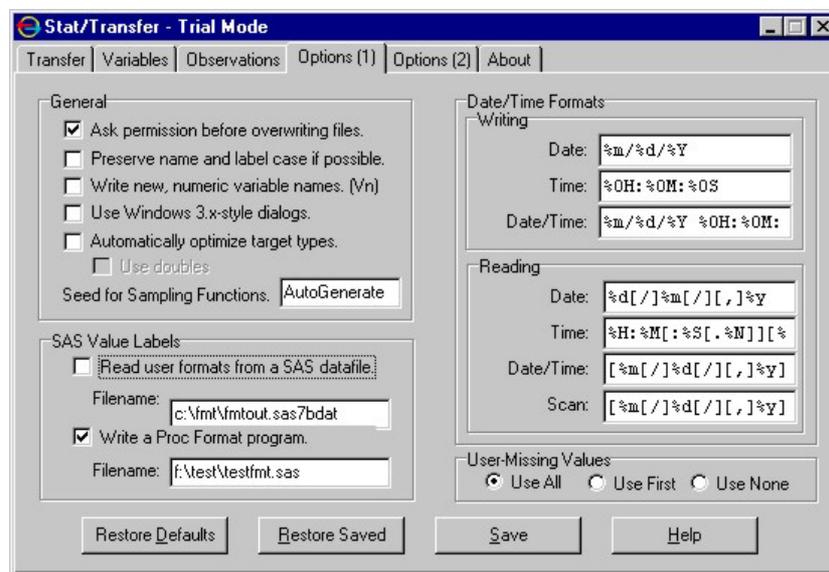
Hinweis: Über die Diskussionsliste SAS-L (www.sas-1.com) wurde am 15. Januar 2002 von Dale McLerran ein Makro zur Verfügung gestellt, das einige

der Probleme behebt, die bei der Übertragung von SPSS-Dateien nach SAS mittels DBMS/COPY auftreten.

4.3 STAT/TRANSFER

Vorgehen: STAT/TRANSFER bietet eine interaktive Datenkonvertierung an. Auch hier wird ein Programm erzeugt, das mit PROC FORMAT die Formate generiert und diese anschließend mit der Prozedur DATASETS zuweist. Dieses Programm muss vor Verwendung der Formate ausgeführt werden.

Variablenamen: Bei Konflikten für Variablenamen werden diese in Name1, Name2 ... umbenannt. Somit bleiben alle Variablen der ursprünglichen Datei erhalten.



Formate/Wertelabels: Mit STAT/TRANSFER lassen sich benutzerdefinierte Wertelabels übertragen. Hierbei werden aber in SPSS definierte Wertelabels für Datumsvariablen (bspw. 24.12.2002 → Weihnachten) nicht übertragen. Die simultane Verwendung von eigenen und internen Formaten ist nicht möglich. Die Namen der im erzeugten Programm mit PROC DATASETS zugewiesenen Formate für alphanumerische Variablen stimmen nicht mit den Namen der erzeugten Formate überein. Dieses Problem kann gelöst werden, indem die Namen alphanumerischer Formate im erzeugten Programm geändert werden. Hierzu wird ein \$ vor den Namen dieser Formate gesetzt (String → \$String).

Fehlende Werte: STAT/TRANSFER bietet auf der Registerkarte Options(1) die Möglichkeit, zwischen benutzerdefinierten und systemeigenen fehlenden Werten zu unterscheiden. So ist es bei Auswahl von „Use None“ nicht mehr notwendig, in der SPSS-Datei die Option „keine fehlenden Werte“ zu setzen.

4.4 %SPSS2SAS_F, ein SAS-Makro

Das Makro %SPSS2SAS_F liest eine SPSS-Datei im Transportformat (*.por) ein, liest Format-Informationen aus einer Textdatei ein, erzeugt entsprechende SAS-Formate und weist diese der eingelesenen SPSS-Datei zu.

Die benötigte Textdatei mit den Formatinformationen lässt sich in SPSS erstellen, in dem man (a) unter dem Menüpunkt „Datei → Info über Datendatei“ die entsprechende Datei auswählt (oder das Kommando DISPLAY ausführt) und (b) die so im SPSS-Ausgabefenster erzeugte Ausgabe mit copy and paste in einer Textdatei (bspw. INFOFILE.TXT) speichert.

Das Makro ist im SAS-Anwenderhandbuch im Netz (<http://www.urz.uni-heidelberg.de/stastistik/sas-ah>) abgelegt und frei verfügbar.

```
%spss2sas_f(infile=c:\testdainfo.txt,  
            porfile=c:\testda.por,  
            outdata=work.mak,  
            store=t)  
proc print data=work.mak;  
run;
```

Vorgehen: Das Makro wird von SAS aus aufgerufen und erzeugt sowohl die Formate als auch eine SAS-Datei in der die Formate den entsprechenden Variablen zugeordnet sind. Es kann zwischen temporärer und permanenter Speicherung der Formate gewählt werden.

Variablenamen: Namenskonflikte werden durch Umbenennung der Variablen in Name1, Name2 ... gelöst.

Formate/Wertelabels: Mit dem Makro werden alle benutzerdefinierten Wertelabels eingelesen und zugewiesen. Datumsformate werden jedoch nicht zugewiesen. Außerdem werden bei gleichzeitiger Verwendung von Benutzer- und Systemdefinierten Formaten in SAS nur die Benutzerdefinierten Formate berücksichtigt. Die Formate werden automatisch zugewiesen. Dabei werden auch Formate für umbenannte Variablen richtig zugewiesen.

Fehlende Werte: Das Makro überträgt benutzerdefinierte und systemeigene fehlende Werte gleichermaßen als „fehlend“. Um benutzerdefinierte fehlende Werte von systemeigenen unterscheiden zu können, muss, wie oben geschildert, in der SPSS-Datei die Option „keine fehlenden Werte“ gesetzt werden.

5 Zusammenfassung

Zusammenfassend lässt sich feststellen, dass wie zu erwarten war nur die Labels (Variablenlabels) völlig problemlos übertragen werden. Sowohl bei den Variab-

lennamen als auch bei Formaten/Wertelabels und den Datenwerten kann es zu Verlusten kommen. Änderungen werden bis auf eine Ausnahme (unbekannte Formate bei DBMS/COPY) ohne Warnung vorgenommen.

Die beiden folgenden Tabellen stellen die Ergebnisse komprimiert zusammen:

SAS → SPSS	Datei-Menü	GET SAS	DBMS/ COPY	STAT/ TRANSFER
Variablenamen	○	○	○	○
(Variablen-) Labels	●	●	●	●
Formate (Wertelabels)	—	*	*	*
Datenwerte	*	*	*	*

SPSS → SAS	Engine SPSS	Prozedur CONVERT	Makro SPSS2SAS.F	DBMS/ COPY	STAT/ TRANSFER
Variablen- namen	*	*	○	○	○
(Variablen-) Labels	●	●	●	●	●
Formate* (Wertelabels)	—/*	—/*	●/*	*/*	*/*
fehlende Werte	*	*	○	○	○

* eigene Formate / Systemformate

Legende:

- vollständige Konvertierung
- Variablenamen u.U. gekürzt und umbenannt
- * Konvertierung u.U. unvollständig
- Konvertierung nicht möglich

Da sich Probleme bei der Übertragung von Daten zwischen zwei Systemen wohl nie ganz vermeiden lassen, wäre es für den Anwender hilfreich, wenn die Konvertierungsprogramme Änderungen protokollieren würden. Anhand eines Protokolls kann der Anwender beurteilen, ob und wenn ja wo es zu Verlusten gekommen sein könnte.

Literatur

- [1] SPSS Inc.: GET SAS. SPSS Syntax Guide, PDF-File Spssbase.pdf im SPSS-Systemverzeichnis der Version 10.0.7, S. 441-444.
- [2] SAS-Anwenderhandbuch im Netz:
www.urz.uni-heidelberg.de/statistik/sas-ah/
- [3] SAS-L – Diskussionsliste. www.sas-l.com

Anhang

Programm zur Erzeugung der SAS Testdatei

```
LIBNAME daten 'f:\test';
LIBNAME library 'f:\test';
PROC FORMAT LIB=library;
  VALUE sex 1=männlich
          2=weiblich;
  VALUE $schule h=Hauptschule
               r=Realschule
               g=Gymnasium;
  VALUE tgeld 0=kein
             0-10=wenig
             10-50=mittel
             50-HIGH=viel;
RUN;
DATA daten.sas2spss;
  input nr vorname $ nachname $:300. geschlecht
        geburtsdatum ddmmyy10. schule $ _tgeld;
  infile cards missover;
  label nr = "Laufende Nummer"
        vorname = "Vorname (bis 8 Zeichen)"
        nachname= "Name (bis 300 Zeichen)"
        geschlecht = "Geschlecht (1=m, 2=w)"
        geburtsdatum = "Geburtsdatum (ddmmyy10.)"
        schule = "Schulart"
        _tgeld = Taschengeld;
  format geschlecht sex.
        geburtsdatum ddmmyy10. nr commax5.2
        schule $schule.
        _tgeld tgeld.;
  geschlecht2=geschlecht;
  if nr=3 then nachname=repeat("Abc",100);
  cards;
1.2 Hans Huttner 1 01.10.1981 g 0
2.3 Frieda Frodilein 2 03.11.1975 r 11
3 Langname Abc
4 Franz Fröhlich 1 04.04.1980 h 120
10 Joseph Hintermeier 1 19.12.1969 g 30
12 Anita Krössel-Mittermaier 2 01.02.1988 g 5
;
proc print data=daten.sas2spss;
run;
proc format lib=library cntlout=daten.fmtout;
run;
```

Die zur Testung verwendete SPSS-Datei findet sich im
SAS-Anwenderhandbuch im Netz

<http://www.urz.uni-heidelberg.de/statistik/sas-ah/>

im Kapitel %SPSS2SAS_F.

