

## Datenfernverarbeitung mit Hilfe von PanelWhiz

Für Nutzer des Betriebspanels bzw. des LIAB, die eine PanelWhiz Lizenz haben, gibt es die Möglichkeit die mit PanelWhiz erstellten Programmcodes einzuschicken und mit in die Datenfernverarbeitung einzubinden.

PanelWhiz ist eine Datenbank, die es ermöglicht in kurzer Zeit ohne großen Aufwand Paneldatensätze erstellen zu können (Haisken-DeNew/Hahn, 2006). Das FDZ bietet seinen Nutzern an, die mit PanelWhiz entstandenen Programme mit in die Datenfernverarbeitung einzubauen. Um diesen Service nutzen zu können, muss man allerdings die Lizenz für PanelWhiz besitzen ([www.PanelWhiz.eu](http://www.PanelWhiz.eu)). Im Folgenden soll kurz erläutert werden, wie PanelWhiz funktioniert und was bei der Datenfernverarbeitung beachtet werden muss.

### **Generierung des Panels:**

Nachdem PanelWhiz über Stata geöffnet und das Betriebspanel ausgewählt wurde stehen für die Erstellung eines Paneldatensatzes zwei Vorgehensweisen zur Variablenauswahl zur Verfügung.

#### 1. Auswahl über die Wellen:

Auf der Projektseite unter dem Menüpunkt „*select a wave*“ können die Variablen jeweils nach Wellen ausgewählt werden. Jede Welle ist mit einem Buchstaben versehen, entsprechend der Angaben im Datensatz. Die Welle 1993 ist dem Buchstaben „a“ zugeordnet, entsprechend aller Variablen dieser Welle. Die Welle 1994 dem Buchstaben „b“, usw. Wenn die benötigten Informationen in mehreren Wellen auftreten, ist die Aufnahme der Variable nur einmal notwendig und muss nicht in jeder Welle erfolgen. Bei welcher Welle die Aufnahme der Variablen erfolgt ist irrelevant.

#### 2. Auswahl über Schlüsselbegriffe:

Alternativ steht die direkte Auswahl von Variablen über den „*Keywordindex*“ zur Verfügung. Hinter den angebotenen Buchstaben finden sich dabei Schlüsselwörter (z.B. Personalsuche), zu denen PanelWhiz bestimmte Items aus allen Wellen auflistet. Durch das Anklicken der Variable wird diese dem Paneldatensatz zugeführt. Die Verwendung der Keywords eignet sich vor allem dafür, bereits erstellte Datensätze um einzelne Variablen zu ergänzen oder um schnell an Detailinformationen über Zusatzvariablen wie z.B. die Variable, die einen Querschnitt kennzeichnet zu gelangen.

Die Wellen 1994 und 1995 verfügen über Zusatzvariablen für neue Betriebe im IAB-Betriebspanel. Sie sind in den Daten mit einem „z“ nach dem Wellenbuchstaben gekennzeichnet (z.B. bz076) und somit auch in PanelWhiz identifizierbar.

Nach Auswahl der Items wird der zu erstellende Datensatz in einem Projekt abgelegt.

## Zusätzliche Hilfen beim Erstellen des Datensatzes

Bei der Auswahl über die Wellen bietet PanelWhiz den Nutzerinnen und Nutzern folgende hilfreiche Informationen zu den einzelnen Items an.

- S** Variablenkurzsummary (N, Mean, Std, Min. und Max.)
- K** Kerneldichtefunktionen (Verteilung der Variablen)
- N** Metadaten (item labels, item-correspondence, Kategorie)
- w, p** Plugininformationen
- i** illustriert die Entwicklung der Variablen im Verlauf des Panels
- X** Löschen von bereits herausgenommenen Variablen

## Abrufen/Speichern des Erstellten Datensatzes:

Über den Button „*Project Page*“ gelangt man zur Hauptmaske von PanelWhiz und kann die ausgewählten Items über [*VIEW OPTIONS*] noch einmal überprüfen und das Projekt speichern. Das Hinzufügen von weiteren Items ist somit auch nachträglich möglich.

Über den Befehl „*Plugins*“ [*RemoveALL/IncludeALL*] können die Nutzer die von PanelWhiz durchgeführte Datenkorrespondenz für ihre Forschungszwecke übernehmen bzw. verändern. Plugins berücksichtigen den Umstand, dass einzelne Items über die Wellen des IAB-Betriebspanels Veränderungen unterliegen (z.B. Anzahl der Antwortkategorien). Sie können über die w, p-Kommandos (siehe Tabelle) überprüft werden.

Der Zeitraum für den Paneldatensatz kann über den Befehl „*Preferences*“ modifiziert werden. Ebenso können hier Spracheinstellungen (deutsch/englisch) vorgenommen werden. Zusätzlich kann hier angegeben werden ob die Aufbereitung der Daten für das Betriebspanel oder den LIAB erfolgt.

Wenn die Datenerstellung endgültig beendet ist kann über den Menüpunkt „*[Do retrieval now!]*“ der definierte Paneldatensatz erstellt werden.

## Retrieved Data Page

Abschließend erscheint die „*Retrieved Data Page*“. Hier sind alle erstellten Dokumente und der Datensatz aufgelistet. Unter „*Data Files*“ ist der Paneldatensatz gespeichert und kann bei Bedarf in andere Verarbeitungsprogramme exportiert werden. Im Falle des IAB Betriebspanels handelt es sich um die Testdaten, die in PanelWhiz integriert sind. Mit dem über PanelWhiz erstellen Paneldatensatz sind deshalb keine inhaltlichen Auswertungen möglich. PanelWhiz speichert allerdings den Stata dofile mit Hilfe dessen der Paneldatensatz erstellt wurde unter „*Documentation Files*“ ab. Dieses dofile muss ans FDZ geschickt werden, damit der Paneldatensatz mit den Originaldaten erstellt werden kann. Der am eigenen PC erstellte Datensatz dient dem Forscher als Testdatensatz für weitere Analysen. Unter den „*Documentation Files*“ befindet sich auch das entstandene logfile und weitere Hinweise. Besonders ist auf das „*acknowledge.smcl-File*“ hinzuweisen, das die Zitation von PanelWhiz beinhaltet.

## Anpassungen für die Datenfernverarbeitung des FDZ

Bevor die durch PanelWhiz entstandenen dofiles an das FDZ geschickt werden können, sind noch einige Anpassungen nötig, damit ein reibungsloser Ablauf der Datenfernverarbeitung gewährleistet wird.

Die Namen der dofiles die mit Hilfe von PanelWhiz erstellt werden, müssen nach den FDZ Vorgaben zur Datenfernverarbeitung bezeichnet werden (Bsp.:1\_Panelerstellung.do, 1a\_plugins.do).

Anpassung der Pfadangaben: Damit im FDZ die Datenverarbeitung möglichst automatisiert ablaufen kann, müssen die Pfadangaben bzw. die globals im dofile folgendermaßen angepasst werden. Zuerst muss der Ort angepasst werden, aus dem PanelWhiz seine Informationen bezieht. Dies ist der orig Ordner. Bitte beachten Sie als nächstes, dass Sie sowohl das Projekt als auch der Enddatensatz im data Ordner speichern. Der logfile soll analog im log Ordner abgespeichert werden.

### Beispiel für ein angepasstes dofile: 1\_Panelerstellung.do

```
global          iabbp "orig";
global          proj  "data";
global          fstem "test";

/* -----( system settings and globals )----- */;

global          temp  "${proj}/tmp";
capture         mkdir "$proj", public;
capture         mkdir "$temp", public;

log             using "log/1_Panelerstellung", text replace;
```

Mit Hilfe des globals fstem wird festgelegt, ob die Aufbereitung der Daten für das Betriebspanel oder den LIAB erfolgt. Im Falle des Betriebspanels muss hier „test“, für die Arbeit mit dem LIAB „LIAB\_2007\_estab\_“ angegeben sein. Die Einstellungen können auch direkt in PanelWhiz unter dem Menüpunkt „*Preferences*“ vorgenommen werden (siehe Seite 2). Bitte denken Sie daran am Ende des dofiles auch den Speicherort für den entstandenen Paneldatensatz zu ändern und den Namen des Plugin-files anzupassen.

```
/* -----( Creating RAW Long File )----- */;

sort           idnum   year;
save           "data/Panel", replace;

/* -----( Running Plugins )----- */;

run            "prog\1a_plugins.do";

/* -----( Creating Long File )----- */;

sort           idnum   year;
save           "data/Panel", replace;
desc;
```

Das angepasste dofile zur Panelerstellung, muss wie die anderen dofiles auch in die Master Datei eingebunden werden. Das dofile, das PanelWhiz für die gewählten Plugins erstellt, muss nicht in die Master Datei eingebunden werden, da es im dofile zur Panelerstellung aufgerufen wird. Bei der Gestaltung des Masterfiles muss außerdem auf die Höhe des gesetzten Arbeitsspeichers geachtet werden. PanelWhiz benötigt 300m. Auch muss ein zusätzlicher adopath X:\datasetup“ aufgeführt werden. Bei Arbeiten während eines Gastaufenthaltes bei uns im Haus und anschließender Datenfernverarbeitung muss dieser adopath nicht in die Master Datei übernommen werden.

#### Beispiel für die master Datei:

```
clear
set mem 300m
set more off

adopath+ prog
adopath+ "X:\datasetup"

cd ..

**Erstellen des Panels:
do prog\1_Panelerstellung.do"
```

#### Anpassung der Spracheinstellung:

Die vorgenommene Spracheinstellung in PanelWhiz lässt sich an dem Befehl label lang EN/ label lang DE nach dem Einlesen der jeweiligen Welle erkennen. Für die Datenfernverarbeitung muss dieser durch lang en/ lang de ersetzt werden.

#### Hinweis:

Bei der Erstellung des Panels wird ausdrücklich daraufhin gewiesen, zusätzlich die Original Fragebögen als Hilfsmittel zu verwenden. Die Fragebögen können auf unserer Homepage bezogen werden unter

[http://fdz.iab.de/de/FDZ\\_Establishment\\_Data/IAB\\_Establishment\\_Panel/IAB\\_Establishment\\_Panel\\_Working\\_Tools.aspx](http://fdz.iab.de/de/FDZ_Establishment_Data/IAB_Establishment_Panel/IAB_Establishment_Panel_Working_Tools.aspx)

#### Weiterführende Informationen:

Haisken-DeNew, John P. und Markus Hahn (2006): "PanelWhiz: A Flexible Modularized Stata Interface for Accessing Large Scale Panel Data Sets"

[www.PanelWhiz.eu](http://www.PanelWhiz.eu)

Quelle: [http://doku.iab.de/fdz/iabb/Anleitung\\_PanelWhiz](http://doku.iab.de/fdz/iabb/Anleitung_PanelWhiz)