Turnover Analyses; Definitions, Process, and Quantification.
Wyoming Department of Employment, Research and Planning
Tony Glover 12/11/2001

Turnover is the general term used to describe the churning that occurs in the labor market. In more specific terms, turnover analysis assigns a category representing an individual's relationship with an employer at a point time (i.e. were they just hired or on their way out). While there are some variations in the collection of Wage Records (WR) from state to state, this database is particularly well suited for tracking individual-employer interactions across time. Knowing an individual received wages from an employer in a year and quarter has value, but only offers a snapshot of the individual's interactions with the employer. By combining several years of quarterly WR data to create a time series, it is possible to assign a value describing the individual-employer interaction relative to the quarters prior and subsequent. For example, we know that John Doe worked for Employer A in year and quarter Y and was compensated with wages Z. This snapshot doesn't tell us that this was the quarter that John Doe was hired by Employer A and that John Doe had previously worked for Employer B and doubled his quarterly wages as a result of changing employers.

First, in this paper we define the terminology needed to assign the turnover categories. Next, we will discuss the processes for arriving at a category determination representing the individual-employer interaction. Once the categories are assigned, several ways to quantify the data are discussed. Lastly, we introduce several applications employed by Wyoming's Research and Planning (R&P) and suggest future avenues of research.

Operational Definitions

A few preliminary definitions are needed at this point and used subsequently to help define the Turnover Categories. The definition of the term "employed" in this paper is that the individual received wages (\$>0) from the employer in a given year and quarter and that this transaction was recorded in WR. The capitol letter "A" denotes that the individual was employed by Employer A in the quarter specified and letter "O" means that the individual was not employed with Employer A. Lastly, the lower case "a" specifies a quarter in which the individual could have worked for Employer A.

Initial inquiry using the prior, subsequent and quarter of interest led to the development of four mutually exclusive individual-employer interactions in Table 1. The definitions for these interactions are each provided three (an attempt to appeal to various learning styles) ways following the table.

Table 1: Four Mutually Exclusive Individual / Employer Interactions in the Reference Quarter (shaded area).

Turnover Category	Quarter Prior (Q-1)	Reference Quarter (Q0)	Quarter Subsequent (Q+1)
Entry (Hire)	O	A	A
Both (Hire and Exit)	O	A	O
Exit	A	A	O
Continuous	A	A	A

Notes

Subscript – Q0 = Reference Quarter, Q-1 = quarter prior and Q+1 = quarter subsequent.

Throughout this paper the A represents a specific employer account.

Throughout this paper the O means that the individual was not employed with EMPLOYER A.

Entry (Hire)

1) An Entry is the situation in which and individual is hired by an employer in a specified quarter and maintains that employment in the subsequent quarter.

2) The individual was employed with Employer A in the reference quarter (Q0) and the quarter subsequent (Q+1), but was not employed with Employer A the quarter prior (Q-1).

3) Of the three quarters in Table 1.

Quarters Employed with Employer A	Quarters Not Employed with Employer A
Q0, Q+1	Q-1

Both (Hire and Exit)

1) The Both category is both a Hire and an Exit in a quarter and could be an individual that starts a job and quits or someone who is hired for a temporary position.

2) The individual was employed with Employer A in the reference quarter (Q0), but was not employed with Employer A the quarter prior (Q-1) or the quarter subsequent (Q+1).

3) Of the three quarters in Table 1.

Quarters Employed with Employer A	Quarters Not Employed with Employer A
Q0	Q-1, Q+1

Exit

1) An Exit is an individual that was employed with an employer in the specified and prior quarters but no longer works for the employer in the subsequent quarter.

2) The individual was employed with Employer A in the reference quarter (Q0) and the quarter prior (Q-1), but was not employed with Employer A the quarter subsequent (Q+1).

3) Of the three quarters in Table 1.

Quarters Employed with Employer A	Quarters Not Employed with Employer A
Q-1, Q0	Q+1

Continuous

1) The individual maintains employment with Employer A for all quarters of interest.

2) The individual was employed with Employer A in the reference quarter (Q0), the quarter prior (Q-1) and the quarter subsequent (Q+1).

3) Of the three quarters in Table 1.

Quarters Employed with Employer A	Quarters Not Employed with Employer A
Q-1, Q0, Q+1	None

All of the data published to date by R&P have used these categories, combining the Entry and Both categories to calculate total hires and the Exit and Both categories to determine the total exits. During subsequent research a discussion arose indicating two distinct categories of Entries, those that are Newhires and those that are Rehires. Alaska's research shop recently published an article[1] on Hires using the two groups. Alaska

---

[1] Alaska Citation

defined a Rehire as someone who was hired by an employer which the individual had previously worked in the past four quarters. Four quarters was determined as appropriate as it would capture turnover situations in which an individual-employer interaction is a result of seasonal employment. Table 2 and the definitions below describe the extended turnover categories.

Table 2: Six Mutually Exclusive Individual / Employer Interactions in the Reference Quarter (shaded area).

Variable	Turnover Category	Q-4	Q-3	Q-2	Q-1	Q0	Q+1
E-N	Entry Newhire	O	O	O	O	A	A
E-R	Entry Rehire	a	a	a	O	A	A
B-N	Both Newhire	O	O	O	O	A	O
B-R	Both Rehire	a	a	a	O	A	O
X	Exit	a	a	a	A	A	O
C	Continuous	a	a	a	A	A	A

Notes
In the table above for an individual to be considered an E-R or a B-R the individual would have had to work for Employer A in at least one of the quarters containing "a" but they were not employed with Employer A in the prior quarter (O).

In adding the additional disaggregates of Entry and Both it becomes necessary to use six rather than three quarters of WR.

Entry Newhire (E-N)

1) An Entry Newhire is the same as an Entry with the exception that the individual had not worked for the employer in the last year (four consecutive quarters).

2) The individual was employed with Employer A in the reference quarter (Q0) and the quarter following (Q+1), but was not employed with Employer A during any of the four quarters prior (Q-4 to Q-1).

3)

Quarters Employed with Employer A	Quarters Not Employed with Employer A
Q0 and Q+1	Q-4 and Q-3 and Q-2 and Q-1

Entry Rehire (E-R)

1) An Entry Rehire is the same as an Entry with the exception that the individual has worked for the employer for at least one quarter in the prior year, excluding the prior quarter.

2) The individual was employed with Employer A in the reference quarter (Q0) and the quarter following (Q+1), was not employed with Employer A in the quarter prior (Q-1), but was employed in at least one of the remaining four quarters prior (Q-4 to Q-2).

3)

Quarters Employed with Employer A	Quarters Not Employed with Employer A	Employed in At Least One of the Following Quarters
Q0, Q+1	Q-1	Q-4, Q-3, Q-2

Both Newhire (B-N)

1) A Both Newhire is the same as a Both with the exception that the individual had not worked for the employer in the last year (four consecutive quarters).

2) The individual was employed with Employer A in the reference quarter (Q0), not employed the quarter following (Q+1) and was not employed with Employer A during any of the four quarters prior (Q-4 to Q-1).

3)

Quarters Employed with Employer A	Quarters Not Employed with Employer A
Q0	Q-4 and Q-3 and Q-2 and Q-1, Q+1

Both Rehire (B-R)

1) A Both Rehire is the same as an Both with the exception that the individual has worked for the employer for at least one quarter in the prior year, excluding the prior quarter.

2) The individual was employed with Employer A in the reference quarter (Q0), not employed the quarter following (Q+1), not employed the quarter prior (Q-1) but was employed in at least one of the remaining four quarters prior (Q-4 to Q-2).

3)

Quarters Employed with Employer A	Quarters Not Employed with Employer A	Employed in At Least One of the Following Quarters
Q0	Q-1, Q+1	Q-4, Q-3, Q-2

The Process

As mentioned earlier there are variations in the WR databases from state to state. For example, South Dakota uses the term "serial" as the field name for the unique employer's account number and Wyoming uses "ui" (Unemployment Insurance Account). Regardless of these slight variations, WR in other states are maintained for the same purposes as those in Wyoming: Unemployment Insurance benefit calculation. Correspondingly, the variables needed (listed in Table 3) to assign turnover categories to the each record are available in all states.

Table 3: Data Elements Needed from WR

Variable	Description	Notes
ssn	Social Security Number	Unique Individual Identifier
year	Year	Year of the record
qtr	Quarter	Quarter of the Record
ui	Employer Account Number	Unique Employer Identifier
wages	Wages	Wages from employer to employee

Note: Wages are not needed for the Turnover Category determination and are not mentioned again until Section 3 of this document. However, it is recommended that wages be carried as a field throughout the process.

Each record represents an individual employed in the year and quarter by a specific employer. Referencing Table 2, it is apparent that a total of six quarters of data (four quarters prior, the reference quarter and one quarter subsequent) are needed to complete one quarter of turnover analyses. To complete the turnover calculations for the four quarters of calendar year 2000 you need the WR with the above listed fields from 1999q1 to 2001q1.

Once the data for the desired periods are downloaded, all years and quarters need to be appended. For example, if I wanted to fill in the appropriate turnover categories for the entire year of 2000, my data set would now include all records in WR from1999q1 to 2001q1. The entire file needs to be sorted on the following variables in the following order. An example of the resulting data for John Doe (SSN = 999-99-9998) appears in Table 4.

1) ssn /ascending
2) ui / ascending
3) year /ascending
4) qtr /ascending

Table 4: WR Data Subsequent to Sort

Rec#	ssn	year	qtr	ui
1	999-99-9997	****	*	**********
2	999-99-9998	2000	1	1111111111
3	999-99-9998	2000	2	1111111111
4	999-99-9998	2000	4	1111111111
5	999-99-9998	1999	4	1111111112
6	999-99-9998	2000	1	1111111112
7	999-99-9998	2000	2	1111111112
8	999-99-9998	2000	3	1111111112
9	999-99-9998	2000	4	1111111113
10	999-99-9998	2000	2	1111111114
11	999-99-9998	2000	3	1111111114
12	999-99-9998	2000	4	1111111114
13	999-99-9998	2001	1	1111111114
14	999-99-9998	1999	3	1111111115
15	999-99-9998	2001	1	1111111115
16	999-99-9999	****	*	**********

Notes
The "Rec #" field is simply a place holder to facilitate discussion and not an actual part of the data set.

All records for John Doe in WR from 1999q1 to 2001q1 appear between the bold triple lines in Table 4. The records one and sixteen are provided to demonstrate that Table 4 is only a snapshot of John Doe's records but in reality contains all the SSNs that occur during the analysis period. Within the triple lines there are several double lines which separate the different employers for which John Doe has worked.

The sorting process takes us one step closer to the goal of assigning the appropriate Turnover Category to each record. The order of the records allows us to group the individual-employer interactions in a linear time sequence. For example, referring to the concepts in Table 2, the appropriate turnover categories are assigned to the data in Table 4, as follows.

Rec# 2 - This was the first quarter (2000q1) in which John Doe was employed by employer "1111111111" (from this point forward the last digit of the employer number is used to designate an individual employer). A review of Rec# 3 reveals that John was still employed with employer "1" in 2000q2. Based on the concepts from Table 2 the Turnover Category assigned to Rec#2 is Entry Newhire (E-N).

Rec# 3 - We see that John was employed by employer "1" in 2000q2 and we know John was employed with employer "1" in 2000q1. Rec#4 shows that the next time John shows up employed with employer "1" was not until two quarters later. Once again based on Table 2 Rec#3 is assigned the Turnover Category of Exit (X).

Rec# 4 – John was employed by employer "1" in 2000q4 but not employed by employer "1" in 2001q1. Further, the last time John was employed by employer "1" was in 2000q2. For Rec# 4, John is assigned the value of Both Rehire (B-R).

This process is fine and dandy if you have an infinite amount of time (Wyoming's WR from 1999q1 to 2001q1 approx 2,000,000 records), but the next few pages streamline the process into a formal algorithm. To do this we need to add a few variables, listed in Table 5 and defined below, to our WR file.

Table 5: Data Elements Added to WR

Variable	Field Size / Data Type	Notes
period	2 / Numeric	A numeric value assigned to a Year and Quarter
forward	2 / Numeric	Explained below
backward	3 / Numeric	Explained below
turn	3 / Character	Turnover Categories discussed earlier

Period – Is a count of quarters in chronological order from a point in time. R&P has chosen the beginning point in time as 1990q1. This means that the value for period assigned to all records with a Year and Quarter of 1990q1 is 1. Correspondingly, 1990q2 = 2, 1990q3 = 3, 1990q4 = 4, 1991q1 = 5, ****so on**** and 2001q1 = 45.

Forward – Forward is a calculated value representing the time transitioned between the current employee-employer interaction and the previous interaction with the same employer. For example, Table 6 repeats the data from Table 4 with the additional variables and values completed. Referring to Rec# 4 the value for forward is the difference between the period for Rec# 4 (44) and the period for Rec#3 (42). In other words the value of forward for Rec#4 is (44 – 42 = 2) two. One additional step must be completed, at the point where the individual or the employer changes (Rec#s 2, 5, 9, 10, 14, 16) the value of forward becomes zero (refer to appendix A, line#1 to line#14).

Backward – Same as forward with the exception that the value calculated represents the time transitioned between the current employee / employer interaction and the subsequent interaction with the same employer. Once again a value of zero is assigned (Rec#s 15, 13, 9, 8, 4, 1) when there is a change of individual or employer (refer to appendix A, line# 16 to line#31).

Table 6: WR Data of John Doe with Turnover Variables Completed

Rec#	ssn	year	qtr	ui	period	forward	backward	turn
1	999-99-9997	****	*	**********	**	**	0	***
2	999-99-9998	2000	1	1111111111	41	0	1	E-N
3	999-99-9998	2000	2	1111111111	42	1	2	X
4	999-99-9998	2000	4	1111111111	44	2	0	B-R
5	999-99-9998	1999	4	1111111112	40	0	1	?(E-N or E-R)
6	999-99-9998	2000	1	1111111112	41	1	1	C
7	999-99-9998	2000	2	1111111112	42	1	1	C
8	999-99-9998	2000	3	1111111112	43	1	0	X
9	999-99-9998	2000	4	1111111113	44	0	0	B-N
10	999-99-9998	2000	2	1111111114	42	0	1	E-N
11	999-99-9998	2000	3	1111111114	43	1	1	C
12	999-99-9998	2000	4	1111111114	44	1	1	C
13	999-99-9998	2001	1	1111111114	45	1	0	?(X or C)
14	999-99-9998	1999	3	1111111115	39	0	6	?(B-N or B-R)
15	999-99-9998	2001	1	1111111115	45	6	0	?(B-N or E-N)
16	999-99-9999	****	*	**********	**	0	**	***

Turn – The actual Turnover Category assigned to the record in the form of the variable names listed in Table 2. A few paragraphs ago we assigned the Turnover Categories to Rec#2 to Rec#4 by working our way through the concepts in Table 2. Table 7 and Line#

32 to Line# 54 in Appendix A introduce six logical statements that allow the assignment of the Turnover Category based on the calculated values of Forward and Backward.

Table 7: Six Mutually Exclusive Logical Statements for Turnover Category Assignment

Variable	Category	Logical Statement (Knowing the values of Forward and Backward)
E-N	Entry Newhire	(forward = 0 or forward > 4) and (backward = 1)
E-R	Entry Rehire	(forward > 1 and forward < 5) and (backward = 1)
B-N	Both Newhire	(forward = 0 or forward > 4) and (backward <> 1)
B-R	Both Rehire	(forward > 1 and forward < 5) and (backward <> 1)
X	Exit	(forward = 1) and (backward <> 1)
C	Continuous	(forward = 1) and (backward = 1)

Notes
The symbol <> stands for Not Equal

Table 7 is identical to Table 2 with the exception that whereas, Table 2 introduces the Turnover Categories symbolically, Table 7 places them in logical statements. A quick review of Table 6 demonstrates that the same Turnover Categories were assigned to the previously determined Rec#s 2, 3 and 4. Further, if you wished to use Table 2 and determine the correct Turnover Category for each of John Doe's records you would find that the statements hold true.

It is important to keep in mind that the period of analysis is the quarters that occur in the calendar year 2000. Using the statements in Table 7 or the code provided in Appendix A will result in incorrect assignment of Turnover Categories for Rec#s 5, 13, 14 and 15. This occurs as the data we used does not span enough quarters to determine whether the individual had worked for the employer previously (Rec#s 5, 14) or in the subsequent

quarter (Rec# 13 and 15) to the time frame of our analysis. To correct this error Rec#s 5, 13, 14 and 15 will be deleted. The procedures discussed in this section are in a flowchart in Appendix B.

Now that each record that occurs in our Wage Records data has been assigned a category related to the individual-employer interaction, the next section will offer several ways with which to summarize the data. First, we will offer operational definitions for various rates associated with Turnover (i.e. Turnover Rate, Continuous Rate, etc.). We will then demonstrate the difference between using covered employment versus individual-employer transactions using Wyoming's 2000 data.

Quantification
Data Merges and Grouping

Data Merges

Now that the Turnover Categories (E-N, E-R, B-N, B-R, C and X) have been assigned, the next step is to merge the ES-202 data with the Wage Records file on the UI account number. As research and planning acquires several data sets with demographics attached to SSNs we further enhance our data by merging the demographics along with the ES-202 data. Additionally we will need to add two fields for the Industry and Sub_ Industry names that are to be used in St Paul. Our Wage Records working file has the fields described in Table 8.

Table 8: Wage Records Working File Table Structure.

Variable	Description	Field Type and Size
Ssn	Social Security Number	Numeric / 9
Year	Year	Numeric / 4
Qtr	Quarter	Numeric / 1
Period	Numeric Year and Quarter	Numeric / 3
Ui	Employer Account Number	Character / 10
Wages	Wages	Numeric / 10
forward		Numeric / 3
backward		Numeric / 3
turn		Character / 3
Fields added by ES-202 Merge		
Sic	SIC	Character / 4
own	Ownership	Character / 1
Fields added by Drivers License Merge		
Dob	Date of Birth	Date / 8
gender	Gender	Character / 1
Fields added at this step		
Industry	Primary Industry Name	Character / 50
sub_indus	Sub Industry Name	Character / 50

The code that appears in Appendix C assigns the Industry and Sub Industry categories to the fields added in Table 8.

Grouping

The last step involved is the grouping of the data on Year, Qtr, Industry, Sub Industry and Turn.. Using the Foxpro Query builder take the following steps.

1) Click on File > New

2) Select the Radio Dial for "Query" and click on the "New File" Button

3) An Open pop up window appears. In this window locate your working file and highlight it. Click on the "OK" button. Another Open pop up window will appear. Click on the "Close" button.

4) This should leave your Query Designer window free to work with.

   a. Click on the "Fields" Tab and highlight "Year" in the "Available Fields" list. Click on the "Add" button to add this field to the "Selected Fields" list. Follow the same steps to add the following fields to the "Selected Fields" list; qtr, industry, sub_indus, turn.

   b. Remaining on the "Fields" tab click once in the "Functions and Expressions:" text box and type in the following statement: cnt(ssn). Click on the "Add" button to add this statement to the "Selected Fields" list. At this point your Query Designer window should look like Figure 1.
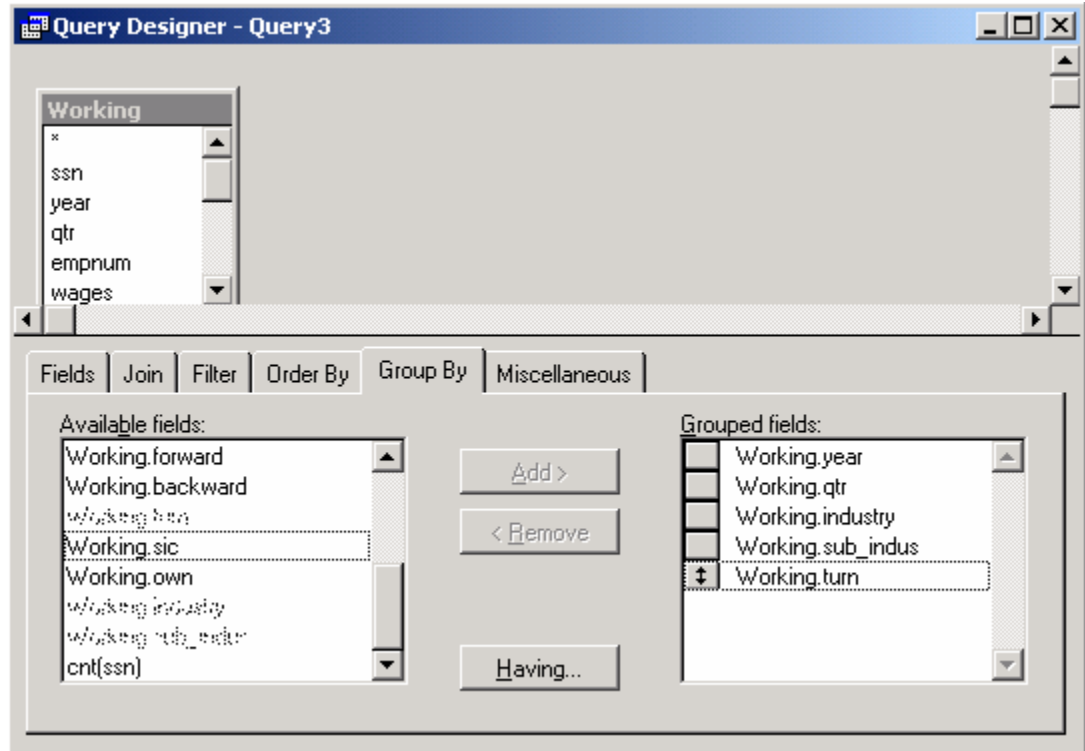
Figure 1
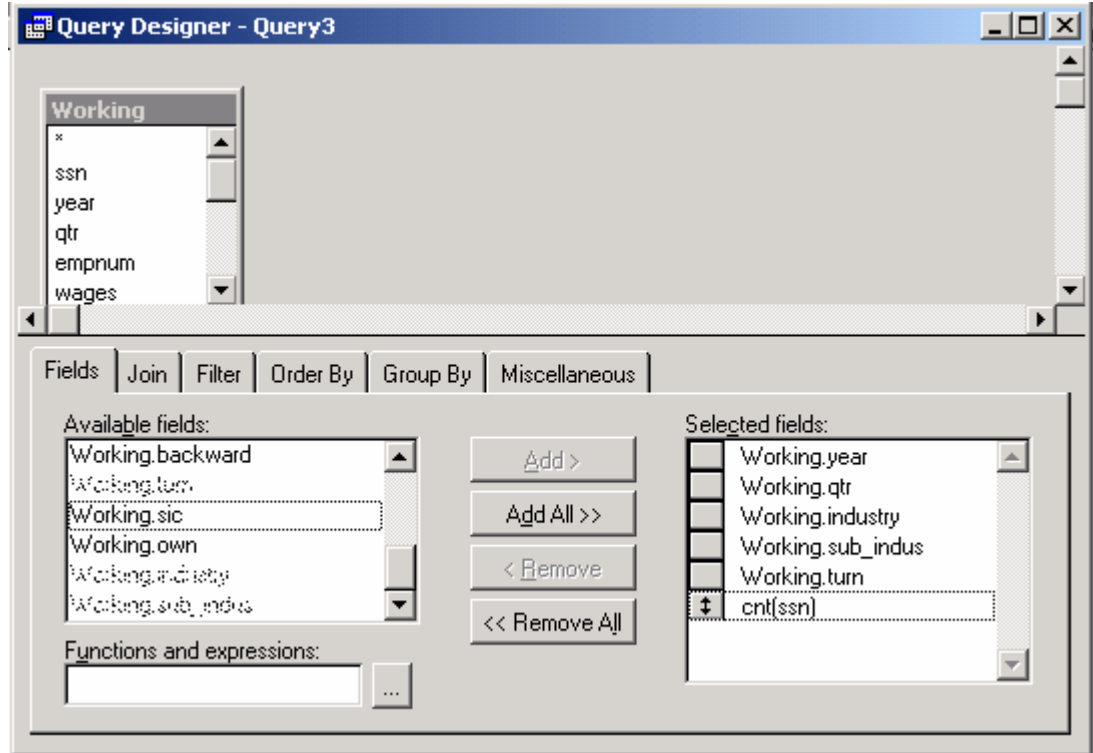
c. Click on the Group By tab and add the following fields from the "Available Fields:" to the "Grouped Fields:" list. The Query Designer Window should look like Figure 2.

Figure 2

d. Click on "Query" on the top file menu and select "Run Query" from the drop down menu, and wait. A little window should pop up with the name "Running Query…" and give you the progress of your query.

e. Your completed table should be displayed with one record per Year, Qtr, Industry, Sub Industry and Turnover Category with a total count of the SSNs in that group.

f. While viewing this table Click on "File" the "Export" to export your table to an Excel file for the remainder of the work. After clicking on File > Export a "Export" window will appear. Select "Microsoft Excel" from the "Type:" drop down menu in the "To:" window select the directory and type the file name which you wish to place the new Excel file.

g. Exit Foxpro

Definitions

Total Exits - The combination of Exit (X), Both Newhires (B-N) and Both Rehires (B-R).

Total Hires - The combination of Entry Newhires (E-N), Entry Rehires (E-R), Both Newhires (B-N) and Both Rehires (B-R).

Total Transactions - The combination of Entry (E), Both Newhires (B-N), Both Rehires (B-R), Exits (X) and Continuous (C).

Turnover Rate - Commonly defined as Total Exits divided by Total Transactions, expressed as a percentage.

Hire Rate - Commonly defined as Total Entries divided by Total Transactions, expressed as a percentage.

Continuous Rate - Defined as Total Continuous divided by Total Transactions, expressed as a percentage.

Flow Rate - Total Entries minus Total Exits divided by Total Transactions.

Appendix A: FoxPro Code for Turnover Category Assignment

line#	Code	Notes
1	go top	Goes to the first record of our sorted file
2	replace forward with 0	Replaces the value of forward with 0 for the first record
3	do while not eof(1)	Starts the do loopwhich will proceed through the file forward until it reaches the end of file (eof)
4	store ssn to s1	Stores the value of the ssn to a temporary variable s1
5	store ui to u1	Stores the value of the ui to a temporary variable u1
6	store period to p1	Stores the value of the period to a temporary variable p1
7	skip 1	Goes to the next record in the file
8	if ssn = s1 and ui = u1	Checks the value of the current records ssn (individual) and ui (employer)against the values of the last record if they are the same executes line# 9 if not goes to line# 11
9	replace forward with (period- p1)	If values of the above comparison are the same replaces forward with the time transitioned in quarters
10	Endif	Closes the if statement
11	if ssn <> s1 or ui <> u1	Basically a clause that executes line# 12 if line# 8 is not true
12	replace forward with 0	Fills the forward value with 0 if the individual changes or the individual changes employers
13	Endif	Closes statement
14	Enddo	As soon as the program reaches the last record closes the loop (line# 1 to line#14)
15		
16	go bottom	Goes to the end of the file
17	i = 2642645	i is the number of records in the file
18	replace backward with 0	Replaces the value of backward with 0 for the last record
19	do while i > 0	Starts the do loop which will proceed through the file backward until i = 0
20	store ssn to s1	Stores the value of the ssn to a temporary variable s1
21	store ui to u1	Stores the value of the ui to a temporary variable u1
22	store period to p1	Stores the value of the period to a temporary variable p1
23	skip -1	Goes to the previousrecord in the file
24	if ssn = s1 and ui = u1	Checks the value of the current records ssn (individual) and ui (employer) against the values of the last record if they are the same executes line# 9 if not goes to line# 11
25	replace backward with (p1– period)	If values of the above comparison are the same replaces forward with the time transitioned in quarters
26	Endif	Closes the if statement
27	if ssn <> s1 or ui <> u1	Basically a clause that executes line# 12 if line# 8 is not true
28	replace backward with 0	Fills the forward value with 0 if the individual changes or the individual changes employers
29	Endif	Closes statement
30	i = i– 1	Subtracts 1 from the value of i each time through loop
31	Enddo	As soon as the program reaches the last record closes the loop (line# 17 to line#30)

```
32
33        go top                                                  Refer to note on Line#1
34        do while not eof(1)                                      Refer to note on Line# 3
35             if (forward = 0 or forward > 4) and backward = 1
36                  replace turn with "EN"
37             Endif
38             if (forward > 1 and forward < 5) and backward = 1
39                  replace turn with "ER"
40             Endif
41             if (forward = 0 or forward > 4) and backward <> 1
42                  replace turn with "BN"
43             Endif                                               Refer to Table 6 for an explanation of the
44             if (forward > 1 and forward < 5) and backward <> 1  logical statements  for Line# 34 to Line# 51
45                  replace turn with "BR"
46             Endif
47             if forward = 1 and backward = 1
48                  replace turn with "C"
49             Endif
50             if forward = 1 and backward <> 1
51                  replace turn with "X"
52             Endif
53             skip 1                                              Refer to Line# 7
54        Enddo                                                    Refer to Line# 14
```

Appendix B: Flowchart of Section 2 Procedures

Step 1
Download WR data needed including the
following data fields.
ssn – Individual identifier
year – Year of record
qtr – Quarter of record
ui – Employer identifier

Step 2
Sort the file on the following criteria.
ssn – Ascending
ui – Ascending
year – Ascending
qtr – Ascending

Step 3
Add the following fields to the data.
period – Numeric 2 digits
forward – Numeric 2 digits
backward – Numeric 2 digits
turn – Character 3 digits

Step 4
Assign the values to period for each record in
the data.
1999Q4 = period 40
2000Q1 = period 41
Etc.
2001Q1 = period 45

Step 5
Assign the values to forward using code similar
to lines 1 to 14 in Appendix A.

Step 6a
Determine the number of records in your data
set and alter line 17 of Appendix A
appropriately.

Step 6b
Assign the values to backward using lines 17 to
31 of Appendix A.

Step 7
Assign Turnover Categories to turn using lines
33 to 54 of the code in Appendix A.

Step 8
Delete records that are outside the analysis
period

Appendix C: Industry and Sub_Industry Assign code

```
set talk off
go record 1
do while not eof(1)
            replace sub_indus with ""
            replace industry with ""

            If val(sic) > 0 and val(sic) < 1000
                  replace sub_indus with "Agriculture, Forestry, Fishing"
                  replace industry with "01 Agriculture, Forestry, Fishing"
            endif

            If val(sic) > 1000 and val(sic) < 1100
                  replace sub_indus with "Metal Mining"
                  replace industry with "02 Mining"
            endif
            If val(sic) > 1100 and val(sic) < 1300
                  replace sub_indus with "Coal Mining"
                  replace industry with "02 Mining"
            endif
            If val(sic) > 1300 and val(sic) < 1400
                  replace sub_indus with "Oil & Gas Extraction"
                  replace industry with "02 Mining"
            endif
            If val(sic) > 1400 and val(sic) < 1500
                  replace sub_indus with "Nonmetallic Minerals Mining"
                  replace industry with "02 Mining"
            endif

            If val(sic) > 1500 and val(sic) < 1600
                  replace sub_indus with "General Building Contractors"
                  replace industry with "03 Construction"
            endif
            If val(sic) > 1600 and val(sic) < 1700
                  replace sub_indus with "Heavy Construction"
                  replace industry with "03 Construction"
            endif
            If val(sic) > 1700 and val(sic) < 1800
                  replace sub_indus with "Special Trade Construction"
                  replace industry with "03 Construction"
            endif

            if val(sic) > 2000 and val(sic) < 2400
                  replace sub_indus with "Manufacturing – Nondurable Goods"
                  replace industry with "04 Manufacturing"
            endif
            if val(sic) > 2400 and val(sic) < 2600
                  replace sub_indus with "Manufacturing – Durable Goods"
                  replace industry with "04 Manufacturing"
            endif
            if val(sic) > 2600 and val(sic) < 3200
                  replace sub_indus with "Manufacturing – Nondurable Goods"
                  replace industry with "04 Manufacturing"
            endif
            if val(sic) > 3200 and val(sic) < 4000
                  replace sub_indus with "Manufacturing – Durable Goods"
                  replace industry with "04 Manufacturing"
            endif

            If val(sic) > 4000 and val(sic) < 4800
                  replace sub_indus with "Transportation"
                  replace industry with "05 TCPU"
            endif
            If val(sic) > 4800 and val(sic) < 5000
                  replace sub_indus with "Communications & Public Utilities"
                  replace industry with "05 TCPU"
            endif

            If val(sic) > 5000 and val(sic) < 5100
```

```
        replace sub_indus with "Wholesale - Durable Goods"
        replace industry with "06 Wholesale Trade"
endif
If val(sic) > 5100 and val(sic) < 5200
        replace sub_indus with "Wholesale - Nondurable Goods"
        replace industry with "06 Wholesale Trade"
endif

If val(sic) > 5200 and val(sic) < 5300
        replace sub_indus with "Building Materials & Garden Supplies"
        replace industry with "07 Retail Trade"
endif
If val(sic) > 5300 and val(sic) < 5400
        replace sub_indus with "General Merchandise Stores"
        replace industry with "07 Retail Trade"
endif
If val(sic) > 5400 and val(sic) < 5500
        replace sub_indus with "Food Stores"
        replace industry with "07 Retail Trade"
endif
If val(sic) > 5500 and val(sic) < 5600
        replace sub_indus with "Auto Dealers & Service Stations"
        replace industry with "07 Retail Trade"
endif
If val(sic) > 5600 and val(sic) < 5700
        replace sub_indus with "Apparels & Accessory Stores"
        replace industry with "07 Retail Trade"
endif
If val(sic) > 5700 and val(sic) < 5800
        replace sub_indus with "Furniture & Home Furnishings"
        replace industry with "07 Retail Trade"
endif
If val(sic) > 5800 and val(sic) < 5900
        replace sub_indus with "Eating & Drinking Places"
        replace industry with "07 Retail Trade"
endif
If val(sic) > 5900 and val(sic) < 6000
        replace sub_indus with "Miscellaneous Retail"
        replace industry with "07 Retail Trade"
endif

If val(sic) > 6000 and val(sic) < 6300
        replace sub_indus with "Finance"
        replace industry with "08 FIRE"
endif
If val(sic) > 6300 and val(sic) < 6500
        replace sub_indus with "Insurance"
        replace industry with "08 FIRE"
endif
If val(sic) > 6500 and val(sic) < 6600
        replace sub_indus with "Real Estate"
        replace industry with "08 FIRE"
endif
If val(sic) > 6600 and val(sic) < 6800
        replace sub_indus with "Finance"
        replace industry with "08 FIRE"
endif

If val(sic) > 7000 and val(sic) < 7200
        replace sub_indus with "Hotels & Other Lodging Places"
        replace industry with "09 Services"
endif
If val(sic) > 7200 and val(sic) < 7300
        replace sub_indus with "Personal Services"
        replace industry with "09 Services"
endif
If val(sic) > 7300 and val(sic) < 7500
        replace sub_indus with "Business Services"
        replace industry with "09 Services"
endif
If val(sic) > 7500 and val(sic) < 7600
```

```
                replace sub_indus with "Auto Repair, Services & Parking"
                replace industry with "09 Services"
        endif
        If val(sic) > 7600 and val(sic) < 7800
                replace sub_indus with "Miscellaneous Repair Service"
                replace industry with "09 Services"
        endif
        If val(sic) > 7800 and val(sic) < 7900
                replace sub_indus with "Motion Pictures"
                replace industry with "09 Services"
        endif
        If val(sic) > 7900 and val(sic) < 8000
                replace sub_indus with "Amusement & Recreation Services"
                replace industry with "09 Services"
        endif
        If val(sic) > 8000 and val(sic) < 8100
                replace sub_indus with "Health Services"
                replace industry with "09 Services"
        endif
        If val(sic) > 8100 and val(sic) < 8200
                replace sub_indus with "Legal Services"
                replace industry with "09 Services"
        endif
        If val(sic) > 8200 and val(sic) < 8300
                replace sub_indus with "Educational Services"
                replace industry with "09 Services"
        endif
        If val(sic) > 8300 and val(sic) < 8400
                replace sub_indus with "Social Services"
                replace industry with "09 Services"
        endif
        If val(sic) > 8400 and val(sic) < 8600
                replace sub_indus with "Museums & Botanical Gardens"
                replace industry with "09 Services"
        endif
        If val(sic) > 8600 and val(sic) < 8700
                replace sub_indus with "Membership Orginizations"
                replace industry with "09 Services"
        endif
        If val(sic) > 8700 and val(sic) < 8800
                replace sub_indus with "Engineering & Management Services"
                replace industry with "09 Services"
        endif
        If val(sic) > 8800 and val(sic) < 8900
                replace sub_indus with "Private Households"
                replace industry with "09 Services"
        endif
        If val(sic) > 8900 and val(sic) < 9000
                replace sub_indus with "Services, NEC"
                replace industry with "09 Services"
        endif

        If val(sic) > 9800
                replace sub_indus with "Not Classified"
                replace industry with "11 Not Classified"
        endif

        If val(sic) > 9000 and val(sic) < 9800
                replace sub_indus with "Public Administration"
                replace industry with "10 Public Administration"
        endif

        if val(sic) = 0
                replace sub_indus with "N/A"
                replace industry with "12 N/A"
        endif

        skip 1
enddo
```