To: Wage Records Workgroup
From: Tony Glover
Senior Analyst
Wyoming Department of Employment
Research and Planning
Date: 7/16/2002
Subject: Description and Application of the Probability Program

This is the first of a series of applications to be released by Research and Planning (R&P) based on our turnover methodology. The purpose of this initial release is to establish a foundation and basic understanding of the process involved in forthcoming R&P applications. The output table of the attached program allows us to determine the duration individuals are retained by employers once hired. The forthcoming applications will allow us to differentiate the retention of individuals by industry, employer, demographics and any combination thereof. Understanding of the dynamics involved in employee-employer interactions is the initial step in applying the turnover methodology to answer practical questions.

The attached fxp (p_base.fxp) file is a compiled Foxpro script. It was written to work directly with the table output (micro_turn.dbf) of the original turnover program I sent the Interstate Turnover Group few months ago. As the methodology is in its early stages, and Tom wishes me to document it, I am not releasing a text version of the script at this time.

If you have not run (or do not wish to run) the earlier script, you will need to create a table in Foxpro with the following structure and this table must be named micro_turn.dbf.

Field Name	Type	length	Nulls
SSN	Numeric	9	No
YEAR	Numeric	4	No
QTR	Numeric	1	No
UI	Character	10	No
WAGES	Numeric	10	No
SIC	Character	4	No
OWN	Character	1	No
PERIOD	Numeric	3	No
FORWARD	Numeric	4	No

BACKWARD	Numeric	4	No
TURN	Character	3	No

To use the Foxpro script:

1) Place the attached file p_base.fxp in the same directory as micro_turn.dbf
2) Either view the directory using Windows Explorer or My Computer and double click on the file p_base.fxp or open Foxpro and select Program > Do from the menu options selecting p_base.fxp.
3) Wait, Wait, Wait, the program runs numerous queries and takes a while. I recommend running the program on a local hard drive rather than your server.

Output files

1) base1.dbf – The foundation file for this and subsequent probability programs. I have already created several other scripts to run similar tables for Industry, UI account, Sex and Age Groups and various combinations of these characteristics (I have included draft outputs of these other tables).
2) p_all.dbf – The output table given on the fourth page of this document.

Methodology

As stated earlier this program uses the turnover methodology developed at R&P over the past few years. The program runs through the micro_turn.dbf and determines the first time an SSN occurred with a UI account (SSN/UI) during the analysis period (for this combined state project 1996q1 to 2001q2). At that first quarter of occurrence the individual must be either an Entry Newhire (E-N) or a Both Newhire (B-N). For example, on the attached flow chart in the row designated "2 Quarters" all of the individuals listed in the W (Working) box were E-N and all those in the NW (Not Working) box were B-N in the first quarter. Therefore, we know how many maintained employment (E-N, 415,298) in the second quarter and how many left employment sometime during the first quarter (B-N, 238,918). Reviewing the row labeled "3 Quarters" the individuals who are still employed (W, 252,026) were continuous (C) in the second quarter of employment and the individuals designated no longer with the employer (NW, 163,275) were exits (X) in the second quarter. At the first exit (or in the

case of those that were hired and only worked for the first quarter B-N) the individual no longer appears in the subsequent quarters (see practical example on page 4).

The p_all.dbf is the aggregate result of data from 1996q1 to 1999q4. The end year and quarter were selected to insure a full eight quarter follow-up (please see table below). For example, if an individual was hired by an employer for the first time in 1999q4, the eight quarter follow-up period would include 1999q4 to 2001q2. While this follow-up time span only includes 7 quarters, it is important to remember that the individual's turnover category in the seventh quarter determines the individual's status, Working (W) or Not Working (NW), in the eighth. For example, if the individual was continuous in the seventh quarter, then they were working for the employer in the eighth; if an exit in the seventh, then they were no longer working for the employer in the eighth.

I have included a couple of demonstration tables and examples of interpretation. The flow chart and first table were created using the attached Foxpro script; the additional tables Industry, Industry X Sex and UI account were created using Wyoming's Wage Records (WR) data back to 1992.

Year	Q1	Q2	Q3	Q4
1996	E-N or B-N	E-N or B-N	E-N or B-N	E-N or B-N
1997	E-N or B-N	E-N or B-N	E-N or B-N	E-N or B-N
1998	E-N or B-N	E-N or B-N	E-N or B-N	E-N or B-N
1999	E-N or B-N	E-N or B-N	E-N or B-N	E-N or B-N
2000	Follow-Up	Follow-Up	Follow-Up	Follow-Up
2001	Follow-Up	Follow-Up		

Practical Example (Example of Output Table base1.dbf)

		1996				1997		
SSN	UI	Q1	Q2	Q3	Q4	Q1	Q2	Q3
01	A	E-N	C	C	C	C	C	C
02	B	E-N	X		E-R	C	X	
03	C	B-N		E-R	C	X		
04	D	B-N		B-R				
05	E	E-N	C	C	C	C	X	
06	F	E-N	C	X				
07	G	E-N	C	X				
08	H	B-N						
09	I	E-N	C	C	X			
10	J	B-N						

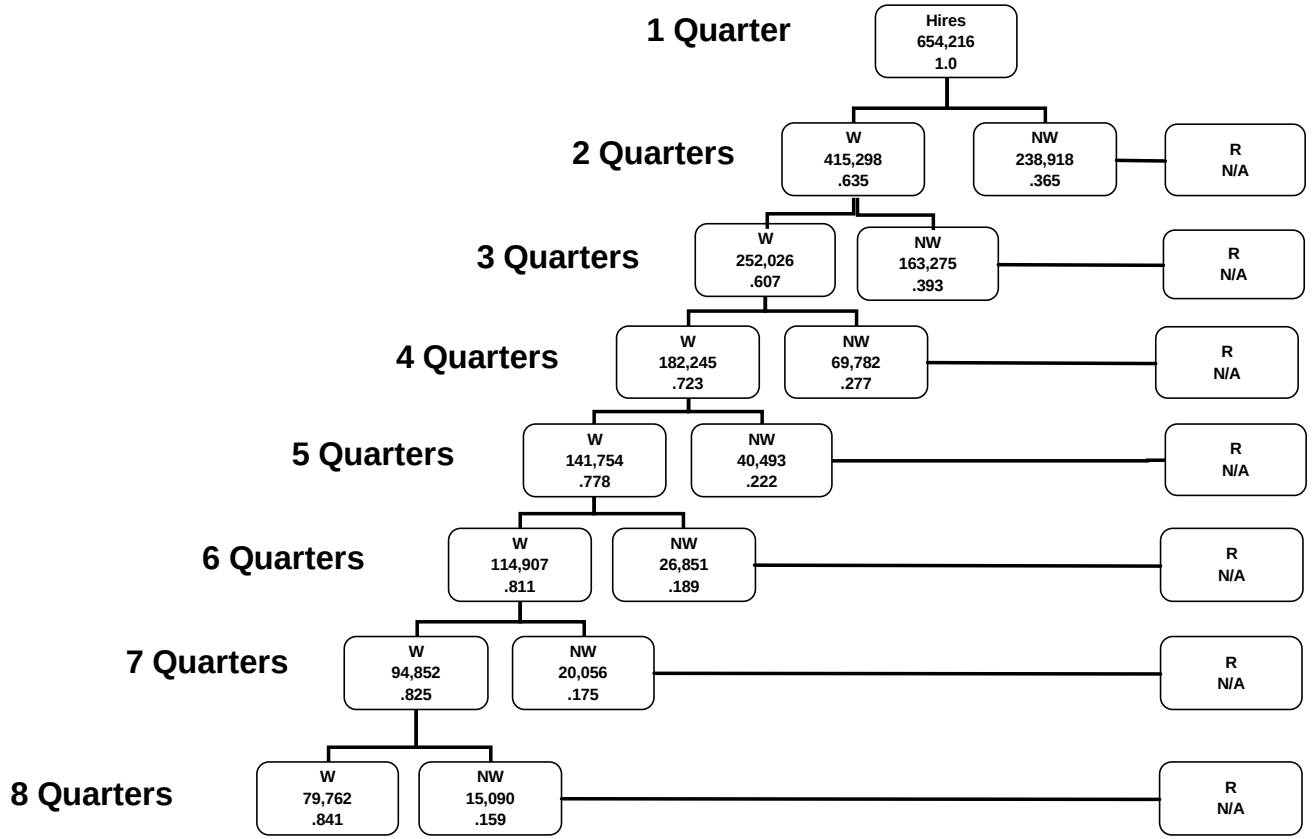
In the example table above you will note, based on earlier text that only the shaded cells are used in this analysis. They meet the definitions provided. In the first quarter of the SSN / UI account occurrence, the individual had to be either an Entry Newhire or a Both Newhire. Subsequent to the first Exit the record is no longer used. For example SSN 02 is considered Working in Q1 and Q2 but not included in the analysis beyond Q2. At a later date we will explore the cases involving rehires, but for the time being we are only concerned with the first time hiring point and the duration the individual is retained by the employer without a break in employment. The resulting N and probabilities for the eight quarters subsequent to hire for the above examples are given in the table below.

Example of Output Table p_all.dbf

Status	P1 (in This Case 1996q1)	P2	P3	P4	P5	P6	P7	P8
Working	10	6	5	3	2	2	1	1
P_W	1.000	.600	.833	.600	.666	1.000	.500	1.000
Not Working	0	4	1	2	1	0	1	0
P_NW	.000	.400	.167	.400	.333	.000	.500	.000
Total Remaining	10	6	5	3	2	2	1	1

In the example of p_all.dbf, we know that all individuals had to have worked for their specific UI accounts in the first quarter (P1). Therefore, the associated probability of working for the employer in P1 is 1.000. Examining the second quarter we find that only 6 of our initial 10 remain or 6/10 = .600. By P3, 5 of the 6 that were working in P2 remain or .833 and etc. P8 is determined by the individual's turnover category in Q7. In the above example SSN 01 had a continuous relationship with the employer in 1997q3 and we know that this means the individual had to work for the employer in 1997q4.

Wyoming's Probability Tree of Remaining with the Same Employer (all industries with four quarter restriction)



Analysis Period - I have adapted the script to use the output file entitled micro_turn.dbf from the program I sent out a few months ago to calculate turnover rates. Therefore, the total analysis period is from 1996q1 to 2001q2. To obtain the 8 qtr followup all hires (entries) at level p1 are retricted to 1996q1 to 1999q4.

R - The original development of this script used all of Wyoming's Wage Record turnover data and with the longer time spectrum I included R and p_r which were the number and probability of being rehired within four quarters by the same employer.

Four Quarter Restriction - All data represent only persons who worked in the state for any four quarters during the entire analysis period (1996q1 to 2001q2).

Tony Glover, Wyoming Department of Employment, Research and Planning. 7/23/2002

Output Table: p_all.dbf

level	w (Working)	nw (Not Working)	p_w	p_nw
p2	415,298	238,918	0.635	0.365
p3	252,026	163,275	0.607	0.393
p4	182,245	69,782	0.723	0.277
p5	141,754	40,493	0.778	0.222
p6	114,907	26,851	0.811	0.189
p7	94,852	20,056	0.825	0.175
p8	79,762	15,090	0.841	0.159

w - Number who continue to work for same employer in designated quarter

nw - Number who no longer work for the same employer in designated quarter

p_w - Probability of working for the same employer in designated quarter.

p_nw - Probability of exiting from the mployer in designated quarter.

level - Number represents the quarters relative to hire (entry)

p1 - which is not displayed in the output table w = p2(w + nw) as all individuals had to work for the employer in p1. Therefore, p_w = 1.000 and p_nw = .000

Example 1a: What's the probability of working for the same employer for four consecutive quarters, once hired?

p(4Q) = (p1(p_w)) (p2(p_w)) (p3(p_w)) (p4(p_w))

p(4Q) = (1.000) (.635) (.607) (.723) = .279

Example 1b: What's the probability of working for the same employer for two additional quarters?

p(2A beyond the 4) = (p5(p_w)) (p6(p_w))

p(2A beyond the 4) = (.778) (.811) = .631

industry	sub_indus	level	w	x	r	p_w	p_x	p_r
02 Mining	03 Oil & Gas Extraction	p2	22,193	15,947	1,595	0.582	0.418	0.100
02 Mining	03 Oil & Gas Extraction	p3	13,097	9,096	1,394	0.590	0.410	0.153
02 Mining	03 Oil & Gas Extraction	p4	9,656	3,441	713	0.737	0.263	0.207
02 Mining	03 Oil & Gas Extraction	p5	7,657	1,999	379	0.793	0.207	0.190
02 Mining	03 Oil & Gas Extraction	p6	6,223	1,434	277	0.813	0.187	0.193
02 Mining	03 Oil & Gas Extraction	p7	5,179	1,044	241	0.832	0.168	0.231
02 Mining	03 Oil & Gas Extraction	p8	4,380	799	211	0.846	0.154	0.264
02 Mining	03 Oil & Gas Extraction	p9	3,816	564	156	0.871	0.129	0.277

industry	sub_indus	level	w	x	r	p_w	p_x	p_r
02 Mining	02 Coal Mining	p2	4,487	1,124	570	0.800	0.200	0.507
02 Mining	02 Coal Mining	p3	3,473	1,014	298	0.774	0.226	0.294
02 Mining	02 Coal Mining	p4	3,054	419	86	0.879	0.121	0.205
02 Mining	02 Coal Mining	p5	2,789	265	35	0.913	0.087	0.132
02 Mining	02 Coal Mining	p6	2,555	234	86	0.916	0.084	0.368
02 Mining	02 Coal Mining	p7	2,310	245	76	0.904	0.096	0.310
02 Mining	02 Coal Mining	p8	2,140	170	60	0.926	0.074	0.353
02 Mining	02 Coal Mining	p9	1,991	149	47	0.930	0.070	0.315

Example 1a: What's the probability of an individual hired by the Oil and Gas Industry working for the same employer for four consecutive quarters, once hired?

p(O&G for 4Q) = (p1(p_w)) (p2(p_w)) (p3(p_w)) (p4(p_w))

p(O&G for 4Q) = (1.000) (.582) (.590) (.737) = .253

Example 1a: What's the probability of an individual hired by the Coal Mining Industry working for the same employer for four consecutive quarters, once hired?

p(Coal for 4Q) = (p1(p_w)) (p2(p_w)) (p3(p_w)) (p4(p_w))

p(Coal for 4Q) = (1.000) (.800) (.774) (.879) = .544

industry	sub_indus	sex	level	w	p_w	x	p_x	r	p_r
09 Services	08 Health Services	F	p2	25,782	0.779	7,323	0.221	849	0.116
09 Services	08 Health Services	F	p3	19,541	0.758	6,241	0.242	857	0.137
09 Services	08 Health Services	F	p4	16,001	0.819	3,540	0.181	570	0.161
09 Services	08 Health Services	F	p5	13,311	0.832	2,690	0.168	390	0.145
09 Services	08 Health Services	F	p6	11,294	0.848	2,017	0.152	327	0.162
09 Services	08 Health Services	F	p7	9,780	0.866	1,514	0.134	302	0.199
09 Services	08 Health Services	F	p8	8,559	0.875	1,221	0.125	255	0.209
09 Services	08 Health Services	F	p9	7,581	0.886	978	0.114	200	0.204

industry	sub_indus	sex	level	w	p_w	x	p_x	r	p_r
09 Services	08 Health Services	M	p2	5,200	0.703	2,201	0.297	214	0.097
09 Services	08 Health Services	M	p3	3,711	0.714	1,489	0.286	174	0.117
09 Services	08 Health Services	M	p4	2,983	0.804	728	0.196	97	0.133
09 Services	08 Health Services	M	p5	2,506	0.840	477	0.160	57	0.119
09 Services	08 Health Services	M	p6	2,154	0.860	352	0.140	64	0.182
09 Services	08 Health Services	M	p7	1,879	0.872	275	0.128	50	0.182
09 Services	08 Health Services	M	p8	1,685	0.897	194	0.103	40	0.206
09 Services	08 Health Services	M	p9	1,512	0.897	173	0.103	43	0.249

Example 1a: What's the probability of a female hired by the Health Services Industry working for the same employer for four consecutive quarters, once hired?

p(F Health for 4Q) = (p1(p_w)) (p2(p_w)) (p3(p_w)) (p4(p_w))

p(F Health for 4Q) = (1.000) (.779) (.758) (.819) = .484

Example 1a: What's the probability of a male hired by the Health Services Industry working for the same employer for four consecutive quarters,

p(M Health for 4Q) = (p1(p_w)) (p2(p_w)) (p3(p_w)) (p4(p_w))

p(M Health for 4Q) = (1.000) (.703) (.714) (.804) = .404

UI Account	level	w	x	r	p_w	p_x	p_r
XXXXXXXX1	p2	545	253	153	0.683	0.317	0.605
XXXXXXXX1	p3	110	435	302	0.202	0.798	0.694
XXXXXXXX1	p4	1	109	85	0.009	0.991	0.780
XXXXXXXX1	p5	0	1	0	0.000	1.000	0.000
XXXXXXXX2	p2	548	23	6	0.960	0.040	0.261
XXXXXXXX2	p3	486	62	20	0.887	0.113	0.323
XXXXXXXX2	p4	463	23	2	0.953	0.047	0.087
XXXXXXXX2	p5	425	38	7	0.918	0.082	0.184
XXXXXXXX2	p6	416	9	7	0.979	0.021	0.778
XXXXXXXX2	p7	399	17	2	0.959	0.041	0.118
XXXXXXXX2	p8	382	17	7	0.957	0.043	0.412
XXXXXXXX2	p9	380	2	0	0.995	0.005	0.000
XXXXXXXX3	p2	344	866	59	0.284	0.716	0.068
XXXXXXXX3	p3	101	243	10	0.294	0.706	0.041
XXXXXXXX3	p4	36	65	5	0.356	0.644	0.077
XXXXXXXX3	p5	9	27	8	0.250	0.750	0.296
XXXXXXXX3	p6	3	6	0	0.333	0.667	0.000
XXXXXXXX3	p7	0	3	0	0.000	1.000	0.000