Estimating Labour Market Transitions and Continuations using Repeated Cross Sectional Data

Pierre Brochu
University of Ottawa

June 1, 2009

Abstract

This paper proposes a population cohort approach for estimating labour market continuations (or transitions) using repeated cross sectional data. I show that the continuation probability can be written as a function of two unconditional means that do not condition on past labour market status. As such, I can construct a consistent set of standard errors that account for the full variability of cross sectional data. Using Current Population Survey data, I show that existing methods tend to systematically underestimate the true standard errors which can lead the researcher to incorrectly conclude that job stability has decreased.

JEL Classification: C41, J64

Key Words: repeated cross section data, duration analysis, employment, job stability

*I would like to thank David Green, Thomas Lemieux, Louis-Philippe Morin and Joris Pinkse for their helpful comments. Contact info: Pierre Brochu, Department of Economics, University of Ottawa, 55 Laurier Avenue East, Desmarais Building, Room 10105, Ottawa (Ontario) K1N 6N5, Canada. Phone: (613) 562-5800 ext. 1424. Fax: (613) 562-5999. Email: pbrochu@uottawa.ca*
1 Introduction

There is a long tradition of exploring labour market transitions in economics. Although the unemployment-employment transition has been the most frequently explored, other transitions or continuations have also been examined, such as the transition out of the labour force (e.g. Jones and Riddell (1999)) and the continuation of a job (job stability, e.g. Brochu (2008); Heisz (2005); Neumark, Polsky, and Hansen (1999)).

While using panel data to estimate these labour market transitions is generally the preferred approach, there are circumstances where that approach is problematic. For example, limited historical coverage (Canadian panels) and data limitations (U.S. panels) make it difficult to differentiate between cyclical and secular changes in job stability. With the absence of this differentiation, one cannot address the real question of interest in the job stability literature: how and why has job stability changed? In such instances, repeated cross sectional data sets offer a valid alternative.

In this paper, I propose a population cohort approach for estimating the continuation (or transition) probability when using repeated cross-section data. The proposed non-parametric approach is empirically tractable, and its identifying assumptions are relatively mild and easy to interpret. Using the proposed population cohort framework, I also re-examine the non-parametric estimator used in the job stability literature. I propose a consistent estimator for its standard errors—one that accounts for the full variability of cross sectional data.

Finally, I use Current Population Survey (CPS) data to show that the existing approaches tend to underestimate the true standard errors. This can lead the researcher to (incorrectly) conclude that job stability has changed.

2 Existing Approach

Following the existing cross sectional literature (e.g. Neumark, Polsky, and Hansen (1999); Heisz (2005)), one can present the retention rate simply as the fraction of “at-risk” individuals in the population that remains with the same employer in the next period

\[ R_{t}^{s,c} = \frac{N_{t+1}^{s+1,c}}{N_{t}^{s,c}} \]

where \( N_{t}^{s,c} \) is the number of people in the population that have time-invariant characteristics \( c \) who have been unemployed for \( s \) periods at time \( t \).

Researchers (e.g. Baker (1992); Neumark, Polsky, and Hansen (1999)) take advantage of the fact that base weights of representative cross sections, like the Current Population Survey (U.S.), sum up to their respective populations. The existing

---

1 See Brochu (2006) for details.
2 For the remainder of this paper, the at-risk group consists of individuals that have characteristics \( c \) and have been with the same employer for \( s \) as of time \( t \).
Estimator is

\[ \hat{Q}^{s,c}_t = \frac{n_{t+1}^{s+1,c}}{n_t} \]  

(2)

where \( n_{t}^{s,c} \) is the sum of the base weights of all individuals with characteristics \( c \) who have been employed \( s \) periods with the same employer as of period \( t \). By using weights as counts, the denominator (numerator) of Equation (2) directly estimates the denominator (numerator) of Equation (1).

Estimating population counts is very intuitive, and it is reasonable to think that the accuracy of the estimator will improve with larger samples; yet, this cannot be proven in any statistical sense. In addition, one cannot lay bare all underlying identifying assumptions without such a proof. Most importantly, the lack of precision carries over to the inference stage. Given the functional form of the estimator, there is no standard way to construct standard errors. In Section 5, I show that these approaches tend to systematically underestimate the true standard errors.

### 3 Proposed Approach

I start by assuming a population cohort. Having a population cohort simply means that there is more than one period of information for each individual in the population. For repeated cross sections, it requires that each sample (cross section) be drawn from the same population, but at different moments in time.

Let \( X_{it} \) be a vector of characteristics of individual \( i \) in period \( t \); the characteristics are, for now, assumed to be time-invariant. Further, let \( TEN_{it} \) represent the length of tenure, i.e. the number of periods the worker has been employed with the same employer as of period \( t \). The retention rate for the “at-risk” group, \( R^{s,c}_t \), can be written as

\[ R^{s,c}_t = \frac{E(1[TEN_{it+1} = s+1, X_{it+1} = c])}{E(1[TEN_{it} = s, X_{it} = c])} \]  

(3)

where \( 1[\cdot] \) is an indicator function that equals 1 if the conditions inside the bracket hold, and zero otherwise.\(^4\) One can estimate this retention rate using two repeated cross sections, i.e.

\[ \hat{R}^{s,c}_t = \frac{\sum_{i=1}^{n_{t+1}} 1[TEN_{it+1} = s+1, X_{it+1} = c]/n_{t+1}}{\sum_{i=1}^{n_t} 1[TEN_{it} = s, X_{it} = c]/n_t} \]  

(4)

where \( n_t \) is the sample size in year \( t \).

Equation (3) is a key insight of this paper. It is conducive to cross sectional analysis because the numerator does not condition on period \( t \) events. This holds true because an individual who

\^4\ Other researchers (e.g. Deaton (1985), Moffitt (1993)) who have estimated dynamic models using repeated cross sections have also relied on this assumption.

\^4\ The proof can be found in Appendix A.1.
has been with the same employer for \( s + 1 \) periods as of time \( t + 1 \), had to have been with the same employer in the previous period (and have one less period of tenure).

Conditioning on only time-invariant characteristics is a sufficient but not a necessary condition for this result to hold. One only needs to be able to infer - from a period \( t + 1 \) cross section - whether an individual who remained with the same employer would have been part of the at-risk group in period \( t \). Said differently, one needs to identify whose indicator function is a “1” in period \( t + 1 \). One can, therefore, estimate a broad range of retention rates. One can not only condition on gender, race, education, but also on age, industry and occupation. I elaborate on the latter three categories below.

The above method can easily deal with ageing when \([t, t + 1]\) spans one or more years. For the 1-year (4-year) rate, identifying a “1” in \( t + 1 \) is straightforward: the worker is simply one (four) year older than he was in period \( t \). Industry affiliation is job related, and as such, will change over time. Yet, one must be able to identify the industry affiliation only if he stayed with the same employer. This is possible if we assume that job tenure (i.e. the employer-employer relationship) ends when the individual switches industry. It is a similar story for occupation. This is a relatively mild assumption as long as the occupations/industries are not too narrowly defined.

Given the simple functional form of the proposed retention rate estimator, i.e. \( \hat{R}^{s,c}_{t} \), one can easily generate consistent standard errors. In Appendix A.2 I show how to do so by first deriving the asymptotic properties of \( \hat{R}^{s,c}_{t} \). Finally, applying Equation (4) to survey data where the probability of being selected is not the same across observations is straightforward. One replaces the sample means with weighted ones.\(^5\)

### 4 Links Between Methods

The American job stability literature (e.g. Swinnerton and Wial (1995); Neumark, Polsky, and Hansen (1999)) estimated 4-year retention rates.\(^6\) It would be untenable to assume that the two cross sections - drawn 4-years apart - come from the same underlying population; the working-age population will have changed due to deaths, emigration and immigration. Fortunately, the population-cohort framework can be extended to deal with such compositional changes.

Assuming that compositional changes break the tenure spell, one can write the retention rate as a function of two population means

\[
R^{s,c}_{t} = \frac{\text{adj}_t E(1[TEN_{it+1} = s + 1, X_{it+1} = c])}{E(1[TEN_{it} = s, X_{it} = c])} \tag{5}
\]

where \( \text{adj}_t \) is the population growth (or adjustment) factor. The population growth factor would be 1.2 if, for example, the population size increased by 20%. An intuitive proof is left to Appendix \(^5\)

\(^5\)Where the weights are normalized to sum up to 1 in each sample period.

\(^6\)The job tenure question was not part of the regular CPS question, but only included in select supplements.
A death easily meets the identifying assumption. For changes due to immigration and emigration, one requires that the migrant changes employer upon arrival in his new country. This empirical strategy would be appropriate if job transfers (where workers stay with same employer) are not the driving force behind migration patterns.

The existing approach, i.e. Equation (2), is in fact an estimator of $R_{s,c}^{t}$ as presented above. This become apparent if one rewrites Equation (2) as

$$Q_{s,c}^{t} = \frac{1}{1} \left( \sum_{i=1}^{n_{t}+1} nw_{it} \right)$$

where $nw_{it}$ is the normalized base weight of individual $i$ in year $t$ and $\hat{adj}_{t} = \frac{1}{\sum_{i=1}^{n_{t}+1} bw_{it}}$ (with $bw_{it}$ representing the base weight).

Given that the sum of the base weights add up to the target population in the CPS, $\hat{adj}_{t}$ is an estimate of the population growth. The second term of Equation (6) is simply the weighted version of $\hat{R}_{s,c}^{t}$ (which was discussed at the end of Section 3).

By rewriting the existing estimator as a function of the proposed one, I can identify its underlying assumptions - namely that changes in population must break the tenure spell. Second, I can also easily construct consistent standard errors. They will be similar to those of $\hat{R}_{s,c}^{t}$, but with an adjustment made for the population change.

5 Empirical Example

Within a retention rate approach, testing for differences in job stability across time or groups is straightforward—only a single restriction needs to be tested. I focus on time differences; the arguments are similar when testing across groups. The null hypothesis is $H_{0} : R_{j}^{s,c} - R_{1}^{s,c} = 0$, where $R_{j}^{s,c} - R_{1}^{s,c}$ is the difference in retention rate over a $j - 1$ period. The t-statistic, $t_{n}$, is

$$t_{n} = \frac{\hat{R}_{j}^{s,c} - \hat{R}_{1}^{s,c}}{\sqrt{\hat{V}_{R_{j}^{s,c} - R_{1}^{s,c}}/n}}$$

where $\hat{V}_{R_{j}^{s,c} - R_{1}^{s,c}}$ is the estimator of $\text{Avar}(\hat{R}_{j}^{s,c} - \hat{R}_{1}^{s,c})$.

The literature has used two approaches to estimating the standard errors. The first approach (e.g. Swinnerton and Wial (1995); Diebold, Neumark, and Polsky (1997)) applies a panel estimator.

---

7 In Appendix A.4., I show that Equation (2) and (6) are numerically equivalent.
8 Where the base weights are normalized to sum up to 1 in each sample period.
9 If one treats $\hat{adj}_{t}$ as a constant, then $se(Q_{s,c}^{t}) = \hat{adj}_{t} \cdot se(R_{s,c}^{t})$
10 To simplify the presentation I assume that the cross sections are all of size $n$. The asymptotic properties of the retention rate differential are left to Appendix A.5.
to the cross sectional data. The second approach which was first proposed by Neumark, Polsky, and Hansen (1999), and subsequently applied to Canadian data by Heisz (2005), treats the numerator of \( \hat{Q}_{s,c}^2 \) in Equation (2) as a random variable, but the denominator as constant. As such, these approaches do not account for the full variability of cross sectional data.

Not only should the standard errors account for the full variability of the cross sections, they must also be able to account for the possible correlation between \( \hat{R}_{s,c}^2 \) and \( \hat{R}_{s,c}^1 \). More precisely, \( \hat{R}_{s,c}^2 \) and \( \hat{R}_{s,c}^1 \) may be correlated since both the denominator of \( \hat{R}_{s,c}^2 \) and the numerator of \( \hat{R}_{s,c}^1 \) are functions of the same (year 2) observations. By allowing both the numerator and denominator in Equation (4) to have sampling distributions, the proposed approach can easily generate the necessary covariance term. This is not the case for existing methods.

I use CPS data to illustrate how the choice of standard errors estimator can matter at the inference stage. I rely on 4-year retention rates—as was previously done in the American job stability literature. Finally, I use the \( \hat{Q}_{s,c}^t \) estimator (instead of \( \hat{R}_{s,c}^t \)); assuming no change in the underlying population would be too restrictive an assumption in this case.

Table 1 examines changes in the 4-year male retention rates from 1996 to 2000. Standard errors are calculated for the NPH method, the DNP method, the proposed method, and the proposed method with no covariance term.

A systematic pattern emerges with the proposed method generating standard errors consistently larger than either DNP or NPH methods. This pattern was found to be robust for other time periods and other sub-populations. The proposed method generates standard errors that are up to 173.5% larger than the DNP estimates and up to 52.3% larger than the NPH estimates. From Table 1 one can observe that the gap is larger for longer tenured groups—groups with higher job stability. In general, the extent to which the NPH method underestimates the correct standard errors will be correlated with the size of the retention rate. Table 1 also indicates that accounting for the covariance term can increase or decrease the standard errors.

As a result, the DNP and NPH approaches to estimating standard errors may lead the researcher to falsely reject the null hypothesis of no change in job stability. Calculating t-statistics for males with 12+ years of tenure illustrates this point. Using either the DNP or NPH methods, one strongly rejects the null hypothesis at the 5% significance level. In fact, my method suggests that the null hypothesis should not be rejected, not even at the 10% level.

---

11 A similar difficulty occurs when testing across groups.

12 From Equation (14) and (21), one can see that conditioning on a sampling distribution for \( D_{1,t} \), a larger \( E(N_{t+1}) \) is associated with a larger first variance term; a term not accounted for by the NPH method.
<table>
<thead>
<tr>
<th>Tenure Group</th>
<th>1996</th>
<th>2000</th>
<th>Difference</th>
<th>Standard Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Method</td>
</tr>
<tr>
<td>0-2</td>
<td>0.4775</td>
<td>0.4999</td>
<td>0.0224</td>
<td>DNP (0.0076)**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NPH (0.0106)**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>proposed (0.0116)**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>proposed (0.0118) *</td>
</tr>
<tr>
<td>3-6</td>
<td>0.4522</td>
<td>0.4658</td>
<td>0.0136</td>
<td>DNP (0.0092)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NPH (0.0133)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>proposed (0.0148)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>proposed (0.0151)</td>
</tr>
<tr>
<td>7-11</td>
<td>0.7069</td>
<td>0.6554</td>
<td>-0.0515</td>
<td>DNP (0.0111)**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NPH (0.0197)**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>proposed (0.0250)**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>proposed (0.0244)**</td>
</tr>
<tr>
<td>12+</td>
<td>0.7288</td>
<td>0.6927</td>
<td>-0.0362</td>
<td>DNP (0.0083)**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NPH (0.0149)**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>proposed (0.0227)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>proposed (0.0194)*</td>
</tr>
</tbody>
</table>

** The estimated difference is significant at the 5% level
* The estimated difference is significant at the 10% level

References


A Appendix

A.1

**Proposition 1** The retention rate of a worker with time-invariant characteristics \(c\) who has been employed for \(s\) periods at time \(t\) can be expressed as

\[
R_{t}^{s,c} = \frac{E(1[\text{TEN}_{it+1} = s+1, X_{it+1} = c])}{E(1[\text{TEN}_{it} = s, X_{it} = c])}.
\]

**proof:**

\[
R_{t}^{s,c} = \frac{\text{Prob}(\text{TEN}_{it+1} = s+1 | \text{TEN}_{it} = s, X_{it} = c)}{\text{Prob}(\text{TEN}_{it} = s, X_{it} = c)} \quad (8)
\]

and since \(\text{TEN}_{it+1} = s+1\) implies \(\text{TEN}_{it} = s\), one can rewrite \(R_{t}^{s,c}\) as

\[
R_{t}^{s,c} = \frac{\text{Prob}(\text{TEN}_{it+1} = s+1, X_{it} = c)}{\text{Prob}(\text{TEN}_{it} = s, X_{it} = c)} \quad (9)
\]

\[
R_{t}^{s,c} = \frac{E(1[\text{TEN}_{it+1} = s+1, X_{it+1} = c])}{E(1[\text{TEN}_{it} = s, X_{it} = c])} \quad (10)
\]

\[
R_{t}^{s,c} = \frac{E(1[\text{TEN}_{it+1} = s+1, X_{it+1} = c])}{E(1[\text{TEN}_{it} = s, X_{it} = c])} \quad (11)
\]

\[
R_{t}^{s,c} = \frac{E(1[\text{TEN}_{it+1} = s+1, X_{it+1} = c])}{E(1[\text{TEN}_{it} = s, X_{it} = c])} \quad (12)
\]

A.2

For ease of exposition, define \(N_{it+1} = 1[\text{TEN}_{it+1} = s+1, X_{it+1} = c]\) and \(D_{it} = 1[\text{TEN}_{it} = s+1, X_{it} = c]\).

**Proposition 2** Assuming iid samples for each year that are drawn from a population cohort, independence across years, and \(\lim_{n_t,n_{t+1} \rightarrow \infty} \frac{n_t}{n_{t+1}} = 1\), then \(\sqrt{n_t}(\hat{R}_{t}^{s,c} - R_{t}^{s,c}) \overset{d}{\rightarrow} N(0, V)\) where \(V\) is

\[
V = \phi_1^2 V(D_{it}) + \phi_2^2 V(N_{it+1}) \quad (13)
\]

with

\[
\phi_1 = \frac{E(N_{it+1})}{[E(D_{it})]^2}, \quad \phi_2 = \frac{1}{E(D_{it})} \quad (14)
\]

**proof:**

a) (consistency)

Apply the Lindberg-Levy Central Limit Theorem

\[
\sum_{i=1}^{n_{t+1}} N_{it+1}/n_{t+1} \overset{p}{\rightarrow} E(N_{it+1}) \quad (15)
\]

\[
\sum_{i=1}^{n_t} D_{it}/n_t \overset{p}{\rightarrow} E(D_{it}) \quad (16)
\]

and use the result of Proposition 1

b) (Asymptotic normality)
Let \( \hat{N}_{t+1} = n_{t+1}^{-1} \sum_{i=1}^{n_{t+1}} N_{it+1} \), \( N_{t+1} = E(N_{it+1}) \) and \( V_{N_{t+1}} = V(N_{it+1}) \), and define \( \hat{D}_t, D_t \) and \( V_{D_t} \) in a similar fashion.

\[
\sqrt{n_t} (\hat{R}_{s,c}^t - R_{s,c}^t) = \sqrt{n_t} \left( \frac{\hat{N}_{t+1} - N_{t+1}}{D_t} - (\hat{D}_t - D_t)N_{t+1} \right)
\]

(17)

\[
= \sqrt{n_t} \left( \frac{\hat{N}_{t+1} - N_{t+1}}{D_t} - \frac{(\hat{D}_t - D_t)N_{t+1}}{D_t} \right) + o_p(1)
\]

(18)

\[
= -\phi_1 \sqrt{n_t} (\hat{D}_t - D_t) + \phi_2 \sqrt{n_t} (\hat{N}_{t+1} - N_{t+1}) + o_p(1)
\]

(19)

\[
d \rightarrow N \left( 0, \frac{\phi_1^2 V_{D_t}}{N} + \frac{\phi_2^2 V_{N_{t+1}}}{N} \right)
\]

(20)

Replacing the population moments in Equation (21) with the corresponding sample analogs generates a consistent estimator for the asymptotic variance. Taking the square root of the estimated variance will generate the standard errors.

A.3

**Proposition 3** Assume that the composition of a country’s population changes from period \( t \) to period \( t+1 \). Further assume that these compositional changes break (or interrupt) the job tenure spell. The retention rate can be expressed as

\[
R_{s,c}^t = \text{adj}_t \frac{E(1[TEN_{it+1} = s + 1, X_{it+1} = c])}{E(1[TEN_{it} = s, X_{it} = c])}
\]

(22)

where \( \text{adj}_t \) is the population growth factor.

**Proof:** To ease the presentation, I assume that the change in population is due to the arrival of one new immigrant in year \( t+1 \). Similar arguments would hold true for other population changes. Without loss of generality, assume a population of size \( N \) in year \( t \), and \( N + 1 \) in year \( t+1 \). Order the year \( t+1 \) population so that the new immigrant is last. By Proposition 1, the retention is

\[
R_{s,c}^t = \frac{\sum_{i=1}^{N} 1[TEN_{it+1} = s + 1, X_{it+1} = c]}{\sum_{i=1}^{N} 1[TEN_{it} = s, X_{it} = c]} / N
\]

(23)

By assuming that the change in population results in a break in job tenure, one can conclude that \( 1[TEN_{N+1,t+1} = s + 1, X_{N+1,t+1} = c] = 0 \). As a result, \( R_{s,c}^t \) can be rewritten as

\[
= \left( \frac{N + 1}{N} \right) \frac{\sum_{i=1}^{N+1} 1[TEN_{it+1} = s + 1, X_{it+1} = c]/N + 1}{\sum_{i=1}^{N} 1[TEN_{it} = s, X_{it} = c]/N}
\]

(24)

\[
\equiv \text{adj}_t \frac{E(1[TEN_{it+1} = s + 1, X_{it+1} = c])}{E(1[TEN_{it} = s, X_{it} = c])}
\]

(25)
A.4

**Proposition 4** Given repeated cross sections where the base weights sum up to the target population, then Equation (2) can be written as

$$\hat{Q}_t^{x,c} = \text{adj}_t \left( \frac{\sum_{i=1}^{n_t} \text{bw}_{it+1} [\text{TEN}_{it+1} = s + 1, X_{it+1} = c]}{\sum_{i=1}^{n_t} \text{bw}_{it} [\text{TEN}_{it} = s, X_{it} = c]} \right)$$  \hspace{1cm} (26)

**proof:**
For ease of notation let $\hat{\mu}$ also part of the population from which $D_{it}$ is drawn. For ease of exposition, define $N_{it+1} = 1[\text{TEN}_{it+1} = s + 1, X_{it+1} = c]$ and $D_{it} = 1[\text{TEN}_{it} = s + 1, X_{it} = c]$. One can rewrite this estimator as

$$\hat{Q}_t^{x,c} = \frac{\sum_{i=1}^{n_t} \text{bw}_{it+1} N_{it+1}}{\sum_{i=1}^{n_t} \text{bw}_{it} D_{it}}$$  \hspace{1cm} (27)

$$= \frac{\sum_{i=1}^{n_t} \text{bw}_{it+1}}{\sum_{i=1}^{n_t} \text{bw}_{it}} \cdot \frac{\sum_{i=1}^{n_t} \text{bw}_{it+1} N_{it+1}/n_t}{\sum_{i=1}^{n_t} \text{bw}_{it} D_{it}/n_t}$$  \hspace{1cm} (28)

$$= \frac{\sum_{i=1}^{n_t} \text{bw}_{it+1}}{\sum_{i=1}^{n_t} \text{bw}_{it}} \cdot \frac{\sum_{i=1}^{n_t} \text{bw}_{it+1} N_{it+1}/n_t}{\sum_{i=1}^{n_t} \text{bw}_{it} D_{it}/n_t}$$  \hspace{1cm} (29)

$$= \text{adj}_t \frac{\sum_{i=1}^{n_t} \text{bw}_{it+1} N_{it+1}/n_t+1}{\sum_{i=1}^{n_t} \text{bw}_{it} D_{it}/n_t}$$  \hspace{1cm} (30)

A.5

**Proposition 5** Assuming iid samples for each year, samples of equal size, independence across years, and no change in population, then $\sqrt{n}((\hat{R}_j - \bar{R}_1) - (R_j - R_1)) \xrightarrow{d} N(0,V)$ where $V$ depends on $j$, an integer greater than or equal to 2. Case 1: $j = 2$

$$V = \phi_1^2 V(D_{11}) + \phi_2^2 V(N_{i2}) + \phi_3^2 V(D_{i2}) + \phi_4^2 V(N_{i3}) + 2\phi_2\phi_3 \mu \text{Cov}(D_{i2}, N_{i2})$$  \hspace{1cm} (31)

**Case 2: $j \geq 3$**

$$V = \phi_1^2 V(D_{11}) + \phi_2^2 V(N_{i2}) + \phi_3^2 V(D_{ij}) + \phi_4^2 V(N_{ij+1})$$  \hspace{1cm} (32)

with

$$\phi_1 = \frac{E(N_{i2})}{E(D_{11})^2}, \hspace{0.5cm} \phi_2 = \frac{1}{E(D_{11})}, \hspace{0.5cm} \phi_3 = \frac{E(N_{ij+1})}{E(D_{ij})^2}, \hspace{0.5cm} \phi_4 = \frac{1}{E(D_{ij})}$$  \hspace{1cm} (33)

and $\mu$ is the probability that a random chosen person in the population from which $D_{it}$ is drawn, is also part of the population from which $N_{it}$ is drawn.

**proof:** For ease of notation let $\hat{N}_j = n_j^{-1} \sum_{i=1}^{n_j} N_{ij}$, $N_j = E(N_{ij})$ and $V_{N_j} = V(N_{ij})$, and define $\hat{D}_{j}$, $D_{j}$ and $V_{D_{j}}$ in a similar fashion. Finally, let $C_2 = \text{Cov}(D_{i2}, N_{i2})$
Case 1: $j = 2$

$$
\sqrt{n}((\hat{R}_2 - \hat{R}_1) - (R_2 - R_1))
= \sqrt{n}\left(\left(\frac{\hat{N}_3}{D_2} - \frac{N_3}{D_2}\right) - \left(\frac{\hat{N}_2}{D_1} - \frac{N_2}{D_1}\right)\right)
= \sqrt{n}(\frac{\hat{N}_3 - N_3}{D_2}\frac{D_2}{D_2} - (\hat{D}_2 - D_2)\frac{N_3}{D_2})
- \sqrt{n}(\frac{\hat{N}_2 - N_2}{D_1}\frac{D_1}{D_1} - (\hat{D}_1 - D_1)\frac{N_2}{D_1})
= \sqrt{n}(\frac{\hat{N}_3 - N_3}{D_2}\frac{D_2}{D^2} - (\hat{D}_2 - D_2)\frac{N_3}{D_2})
- \sqrt{n}(\frac{\hat{N}_2 - N_2}{D_1}\frac{D_1}{D^2} - (\hat{D}_1 - D_1)\frac{N_2}{D_1}) + o_p(1)
= \phi_1\sqrt{n}(\hat{D}_1 - D_1) - \phi_2\sqrt{n}(\hat{N}_2 - N_2) - \phi_3\sqrt{n}(\hat{D}_2 - D_2) + \phi_4\sqrt{n}(\hat{N}_3 - N_3) + o_p(1)
\xrightarrow{d} N\left(0, \phi_1^2V_{D_1} + \phi_2^2V_{N_2} + \phi_3^2V_{D_2} + \phi_4^2V_{N_3} + 2\phi_2\phi_3\mu C_2\right)
$$

(34)
(35)
(36)
(37)
(38)

Case 2: $j \geq 3$. The proof is similar to Case 1, with one exception. Since the four components of the test statistics, i.e. $\hat{N}_{j+1}, \hat{D}_j, \hat{N}_2$ and $\hat{D}_1$ are functions of different yearly samples when $j \geq 3$, the covariance term is zero. 
