Evaluation of a Modified Gross Flows Estimator for The Current Population Survey*

Stephen M. Miller[1], Connor Doherty[2],
Office of Survey Methods Research, U.S. Bureau of Labor Statistics[1]
Office of Employment and Unemployment Statistics, U.S. Bureau of Labor Statistics[2]

Abstract
We present a gross flows estimation approach which builds off the paper of Stasny and Fienberg (1985). Our method uses population weighted estimates from two consecutive months of matched data from the Current Population Survey (CPS) using the sampling weights from each of the two matched months to produce two sets of partial gross flows tables. We then use a modeling approach from Stasny and Fienberg to reconcile the two partial tables to produce an estimate of the population gross flows table. Closed form solutions are presented which require an optimization solution to determine Lagrange parameters. We use the method to produce estimated gross flows tables for CPS from 2003-2023 and estimate the variance of the estimates by replication.

Key Words: Monthly Labor Force Transitions

1. Introduction

The Current Population Survey (CPS) is a household survey conducted for the Bureau of Labor Statistics by the Unites States Census Bureau. It is the source of monthly estimates Unemployment and other labor force measures by demographic groups and regions of the United States. Primary interest is on the monthly stock estimates and changes in the those estimates over time. There is also an interest in the changes between labor force states over time which are called labor force gross flows. An example of such analyses is discussed in Ilg (2005).

The CPS is well suited to examine transitions between labor force states between month because it attempts to interview some of the same households in consecutive months. In particular, the CPS uses a 4-8-4 rotation sampling design, which means that once a household is sampled in their first month, they are in sample the next three months then exit the sample for 8 months and then interviewed again for 4 consecutive months before exiting the sample. Note that by design the household is interviewed in the same 4 months in two consecutive years. The design insures that approximating 75% of the sample overlaps from month-to-month. For more on the CPS and its design see Technical Paper 77 (2019).

1.1 Population Gross Flows Table

We begin by defining the population gross flows table for two consecutive months in Table 1. The table classifies all those people who are in the civilian noninstitutional population aged 16 and older (CNP) in either of two consecutive months. The columns and rows labeled E, U and N stand for those people classified as Employed, Unemployed or Not in Labor Force respectively. The column labeled F denotes those people who were in the CNP in the previous month $t-1$ but were not in the CNP in the current month t . We refer to those as outflows. The row labeled F denotes those people who were in the CNP in the current month t but were not in the CNP in the previous month $t-1$. We refer to those as inflows. Note that the intersection of the column and row labeled F is zero since the people in the table were in the CNP at least one of the two months. The column and row labeled T are the totals.

*Views expressed are those of the authors and do not necessarily reflect the views or policies of the U.S. Bureau of Labor Statistics.

Table 1: Population Gross Flows Table.

		time period t				
		E	U	N	F	T
time period $t-1$	E	T_{EE}	T_{EU}	T_{EN}	T_{EF}	T_{ET}
	U	T_{UE}	T_{UU}	T_{UN}	T_{UF}	T_{UT}
	N	T_{NE}	T_{NU}	T_{NN}	T_{NF}	T_{NT}
	F	T_{FE}	T_{FU}	T_{FN}	0	T_{FT}
	T	T_{TE}	T_{TU}	T_{TN}	T_{TF}	T_{TT}

If we denote the total CNP for time periods $t-1$ and t by CNP_{t-1} and CNP_t respectively, then the following relationships hold among the elements of Table 1.

$$\begin{aligned} CNP_{t-1} &= T_{ET} + T_{UT} + T_{NT} \\ CNP_t &= T_{TE} + T_{TU} + T_{TN} \\ T_{TT} &= CNP_{t-1} + T_{FT} = CNP_t + T_{TF} \\ CNP_t - CNP_{t-1} &= T_{FT} - T_{TF} \end{aligned} \tag{1.1}$$

where the last equation says the change in CNP from time period $t-1$ to t is the inflow minus the outflow. Our goal is to estimate the Population Gross Flows table each month in such a way that the row and column margins match the official monthly cross-sectional estimates.

1.2 Estimated Partial Gross Flows Tables

We begin with some notation. Let S_t denote the CPS sample for month t and let S_{t-1} denote the sample for month $t-1$. We will work with the merged sample $S_t \cup S_{t-1}$ which contains units which were collected in either or both months. Let $w_{t,i}$ denote the sample weight for unit i in month t and similarly $w_{t-1,i}$ for unit i in month $t-1$. In both cases we will assume we are using the CPS composite weights used to compute the official monthly CPS estimates, which implies that

$$\begin{aligned} CNP_t &= \sum_{i \in S_t \cup S_{t-1}} w_{t,i} \\ CNP_{t-1} &= \sum_{i \in S_t \cup S_{t-1}} w_{t-1,i} \end{aligned} \tag{1.2}$$

Also note that by definition

$$\begin{aligned} &\text{if } i \in S_t \text{ and } i \notin S_{t-1} \text{ then } w_{t-1,i} = 0 \\ &\text{if } i \notin S_t \text{ and } i \in S_{t-1} \text{ then } w_{t,i} = 0 \end{aligned} \tag{1.3}$$

Now define the following indicator variables at the unit level

$$\begin{aligned} D_{E*} = 1 &\quad \text{if unit } i \text{ is } E \text{ in month } t-1 \text{ and 0 otherwise} \\ D_{U*} = 1 &\quad \text{if unit } i \text{ is } U \text{ in month } t-1 \text{ and 0 otherwise} \\ D_{N*} = 1 &\quad \text{if unit } i \text{ is } N \text{ in month } t-1 \text{ and 0 otherwise} \\ D_{*E} = 1 &\quad \text{if unit } i \text{ is } E \text{ in month } t \text{ and 0 otherwise} \\ D_{*U} = 1 &\quad \text{if unit } i \text{ is } U \text{ in month } t \text{ and 0 otherwise} \\ D_{*N} = 1 &\quad \text{if unit } i \text{ is } N \text{ in month } t \text{ and 0 otherwise} \end{aligned} \tag{1.4}$$

Then

$$\begin{aligned}\hat{E}_{t-1} &= \sum_{i \in S_t \cup S_{t-1}} w_{t-1,i} D_{E*,i} \\ \hat{U}_{t-1} &= \sum_{i \in S_t \cup S_{t-1}} w_{t-1,i} D_{U*,i} \\ \hat{N}_{t-1} &= \sum_{i \in S_t \cup S_{t-1}} w_{t-1,i} D_{N*,i}\end{aligned} \tag{1.5}$$

are the official estimates of E, U and N for month $t-1$ while

$$\begin{aligned}\hat{E}_{t} &= \sum_{i \in S_t \cup S_{t-1}} w_{t,i} D_{*E,i} \\ \hat{U}_{t} &= \sum_{i \in S_t \cup S_{t-1}} w_{t,i} D_{*U,i} \\ \hat{N}_{t} &= \sum_{i \in S_t \cup S_{t-1}} w_{t,i} D_{*N,i}\end{aligned} \tag{1.6}$$

are the official estimates of E, U and N for month t . Also note that

$$\begin{aligned}CNP_t &= \hat{E}_t + \hat{U}_t + \hat{N}_t \\ CNP_{t-1} &= \hat{E}_{t-1} + \hat{U}_{t-1} + \hat{N}_{t-1}\end{aligned} \tag{1.7}$$

Next we define some indicator variables based on the classification in both months. We begin by defining

Row E

$$\begin{aligned}D_{EE,i} &= 1 \text{ if unit } i \text{ is } E \text{ at time } t-1 \text{ and } E \text{ at time } t \text{ and 0 otherwise} \\ D_{EU,i} &= 1 \text{ if unit } i \text{ is } E \text{ at time } t-1 \text{ and } U \text{ at time } t \text{ and 0 otherwise} \\ D_{EN,i} &= 1 \text{ if unit } i \text{ is } E \text{ at time } t-1 \text{ and } N \text{ at time } t \text{ and 0 otherwise} \\ D_{EF,i} &= 1 \text{ if unit } i \text{ is } E \text{ at time } t-1 \text{ and } F \text{ at time } t \text{ and 0 otherwise}\end{aligned} \tag{1.8}$$

Row U

$$\begin{aligned}D_{UE,i} &= 1 \text{ if unit } i \text{ is } U \text{ at time } t-1 \text{ and } E \text{ at time } t \text{ and 0 otherwise} \\ D_{UU,i} &= 1 \text{ if unit } i \text{ is } U \text{ at time } t-1 \text{ and } U \text{ at time } t \text{ and 0 otherwise} \\ D_{UN,i} &= 1 \text{ if unit } i \text{ is } U \text{ at time } t-1 \text{ and } N \text{ at time } t \text{ and 0 otherwise} \\ D_{UF,i} &= 1 \text{ if unit } i \text{ is } U \text{ at time } t-1 \text{ and } F \text{ at time } t \text{ and 0 otherwise}\end{aligned} \tag{1.9}$$

Row N

$$\begin{aligned}D_{NE,i} &= 1 \text{ if unit } i \text{ is } N \text{ at time } t-1 \text{ and } E \text{ at time } t \text{ and 0 otherwise} \\ D_{NU,i} &= 1 \text{ if unit } i \text{ is } N \text{ at time } t-1 \text{ and } U \text{ at time } t \text{ and 0 otherwise} \\ D_{NN,i} &= 1 \text{ if unit } i \text{ is } N \text{ at time } t-1 \text{ and } N \text{ at time } t \text{ and 0 otherwise} \\ D_{NF,i} &= 1 \text{ if unit } i \text{ is } N \text{ at time } t-1 \text{ and } F \text{ at time } t \text{ and 0 otherwise}\end{aligned} \tag{1.10}$$

Row F

$$\begin{aligned}D_{FE,i} &= 1 \text{ if unit } i \text{ is } F \text{ at time } t-1 \text{ and } E \text{ at time } t \text{ and 0 otherwise} \\ D_{FU,i} &= 1 \text{ if unit } i \text{ is } F \text{ at time } t-1 \text{ and } U \text{ at time } t \text{ and 0 otherwise} \\ D_{FN,i} &= 1 \text{ if unit } i \text{ is } F \text{ at time } t-1 \text{ and } N \text{ at time } t \text{ and 0 otherwise}\end{aligned} \tag{1.11}$$

Next we define two additional categories that are relevant for the sample estimates that were not needed for defining the population gross flows table. We let M denote if a unit is missing from the sample at either time $t-1$ or t. We will let R denote if a unit has been rotated in or out of the sample at either time $t-1$ or t. For example, a unit who was MIS4 or MIS8 at time $t-1$ would be R at time t, while a unit that was MIS1 or MIS5 at time t would be R at time $t-1$. It is important to distinguish those that are M from those that are R. Those that are M should have been eligible to be interviewed while those who are R could not have been interviewed by virtue of the 4-8-4 sample rotation design. This allows us to define the following additional indicator variables,

Row E

$$\begin{aligned} D_{EM,i} &= 1 \text{ if unit } i \text{ is } E \text{ at time } t-1 \text{ and } M \text{ at time } t \text{ and } 0 \text{ otherwise} \\ D_{ER,i} &= 1 \text{ if unit } i \text{ is } E \text{ at time } t-1 \text{ and } R \text{ at time } t \text{ and } 0 \text{ otherwise} \end{aligned} \tag{1.12}$$

Row U

$$\begin{aligned} D_{UM,i} &= 1 \text{ if unit } i \text{ is } U \text{ at time } t-1 \text{ and } M \text{ at time } t \text{ and } 0 \text{ otherwise} \\ D_{UR,i} &= 1 \text{ if unit } i \text{ is } U \text{ at time } t-1 \text{ and } R \text{ at time } t \text{ and } 0 \text{ otherwise} \end{aligned} \tag{1.13}$$

Row N

$$\begin{aligned} D_{NM,i} &= 1 \text{ if unit } i \text{ is } N \text{ at time } t-1 \text{ and } M \text{ at time } t \text{ and } 0 \text{ otherwise} \\ D_{NR,i} &= 1 \text{ if unit } i \text{ is } N \text{ at time } t-1 \text{ and } R \text{ at time } t \text{ and } 0 \text{ otherwise} \end{aligned} \tag{1.14}$$

Row M

$$\begin{aligned} D_{ME,i} &= 1 \text{ if unit } i \text{ is } M \text{ at time } t-1 \text{ and } E \text{ at time } t \text{ and } 0 \text{ otherwise} \\ D_{MU,i} &= 1 \text{ if unit } i \text{ is } M \text{ at time } t-1 \text{ and } U \text{ at time } t \text{ and } 0 \text{ otherwise} \\ D_{MN,i} &= 1 \text{ if unit } i \text{ is } M \text{ at time } t-1 \text{ and } N \text{ at time } t \text{ and } 0 \text{ otherwise} \end{aligned} \tag{1.15}$$

(1.16)

Row R

$$\begin{aligned} D_{RE,i} &= 1 \text{ if unit } i \text{ is } R \text{ at time } t-1 \text{ and } E \text{ at time } t \text{ and } 0 \text{ otherwise} \\ D_{RU,i} &= 1 \text{ if unit } i \text{ is } R \text{ at time } t-1 \text{ and } U \text{ at time } t \text{ and } 0 \text{ otherwise} \\ D_{RN,i} &= 1 \text{ if unit } i \text{ is } R \text{ at time } t-1 \text{ and } N \text{ at time } t \text{ and } 0 \text{ otherwise} \end{aligned} \tag{1.17}$$

(1.18)

We can construct weighted estimates based on these indicator variables. We illustrate this in Table 2 and Table 3 below.

Table 2: Sample Estimate Gross Flows Table A based upon weighting from time $t-1$.

		time period t					
		E	U	N	F	M	R
time period $t-1$	E	A_{EE}	A_{EU}	A_{EN}	A_{EF}	A_{EM}	A_{ER}
	U	A_{UE}	A_{UU}	A_{UN}	A_{UF}	A_{UM}	A_{UR}
	N	A_{NE}	A_{NU}	A_{NN}	A_{NF}	A_{NM}	A_{NR}

Table 3: Sample Estimate Gross Flows Table B based upon weighting from time t.

		time period t		
		E	U	N
time period $t-1$	E	B_{EE}	B_{EU}	B_{EN}
	U	B_{UE}	B_{UU}	B_{UN}
	N	B_{NE}	B_{NU}	B_{NN}
	F	B_{FE}	B_{FU}	B_{FN}
	M	B_{ME}	B_{MU}	B_{MN}
	R	B_{RE}	B_{RU}	B_{RN}

These tables can be constructed by month-in-sample group or by aggregating over those groups. For example the first five columns of A can be computed for $A^{(i)}$ for $i = 1, 2, 3, 5, 6, 7$ and the last column for $i = 4, 8$, and with the first five rows of $B^{(i)}$ for $i = 2, 3, 4, 6, 7, 8$ and the last row for $i = 1, 5$. We discuss modeling and estimation of both cases in the next section. The current BLS gross flows estimation method only uses the agregate form.

1.3 Current Official Gross Flows Estimates

BLS uses a method of raking described in Frazis, Robison, Evans and Duff (2005) to estimate the population gross flows table each month. In addition to the CPS their method also uses data on death rates from the National Center for Health Statistics (NCHS) to improve death estimates (flows out of the civilian non-institutional population) obtained directly from the CPS. In addition they specifically account for those persons who are inflows who turned 16 in the current month and whose labor force status was know the previous month. They use only a portion of the partial flow tables A and B discussed earlier. In particular the first four rows of B and the flow column of A is used. They also use the stock estimates for E,U and N for both the current and previous time period for raking.

Our approach uses all the elements of the A and B matrices and ties it to the population gross flows table with a statistical model similar to that described in Stasny and Feinberg (1985) and Stasny (1988). We do not incorporate death rate information from NCHS and rely entirely on data from the CPS. Stasny and Fienberg did not directly model the A and B matrices as we do but we still apply their modeling approach. This is discussed in the next section.

2. Modeling the Partial Gross Flows Tables

We will use an approach which is similar to models in Stasny (1988). A key aspect of the modeling is that we define parameters which control the flow into the missing category as well as parameters which account for differences by month-in-sample. We consider two cases. The first is modeling the aggregated tables, and the second for the tables constructed by month-in-sample.

2.1 Modeling the Aggregate Partial Gross Flows Tables

We next define the following expectations in terms of the rows of matrix A for each month-in-sample. We begin with row E.

Row E month-in-sample i = 1,2,3,5,6,7

$$\begin{aligned} E\{A^{(i)}_{EE}\} &= \theta_{AEi}(1 - P_{AE})T_{EE} \\ E\{A^{(i)}_{EU}\} &= \theta_{AEi}(1 - P_{AE})T_{EU} \\ E\{A^{(i)}_{EN}\} &= \theta_{AEi}(1 - P_{AE})T_{EN} \\ E\{A^{(i)}_{EF}\} &= \theta_{AEi}(1 - P_{AE})T_{EF} \\ E\{A^{(i)}_{EM}\} &= \theta_{AEi}P_{AE}T_{ET} \end{aligned} \tag{2.1}$$

Row E summed over i = 1,2,3,5,6,7

$$\begin{aligned} E\{A_{EE}\} &= (1 - \theta_{AE48})(1 - P_{AE})T_{EE} \\ E\{A_{EU}\} &= (1 - \theta_{AE48})(1 - P_{AE})T_{EU} \\ E\{A_{EN}\} &= (1 - \theta_{AE48})(1 - P_{AE})T_{EN} \\ E\{A_{EF}\} &= (1 - \theta_{AE48})(1 - P_{AE})T_{EF} \\ E\{A_{EM}\} &= (1 - \theta_{AE48})P_{AE}T_{ET} \end{aligned} \tag{2.2}$$

Row E month-in-sample i = 4,8

$$E\{A^{(i)}_{ER}\} = \theta_{AEi}T_{ET} \tag{2.3}$$

Row E summed over i = 4,8

$$E\{A_{ER}\} = \theta_{AE48} T_{ET} \tag{2.4}$$

The parameters θ_{AEi} for $i = 1, 2, 3, 5, 6, 7$ are the month-in-sample effects for Employment in month $t - 1$ where by definition

$$\sum_{i=1}^{8} \theta_{AEi} = 1 \tag{2.5}$$

We define $\theta_{AE48} = \theta_{AE4} + \theta_{AE4}$. The parameters P_{AE} represent what percent of the totals $T_{EE}, T_{EU}, T_{EN}, T_{EF}$ flow into the missing category from employed. When modeling the aggregate partial gross flows tables we do not let P_{AE} vary by month-in-sample, but we will allow for P_{AEi} $i = 1, 2, 3, 5, 6, 7$ to vary when we model the tables at the month-in-sample level. For convenience we define

$$T_{ET} = T_{EE} + T_{EU} + T_{EN} + T_{EF} \tag{2.6}$$

We similarly define row U.

Row U month-in-sample i = 1,2,3,5,6,7

$$\begin{aligned} E\{A_{UE}^{(i)}\} &= \theta_{AUi}(1 - P_{AU}) T_{UE} \\ E\{A_{UU}^{(i)}\} &= \theta_{AUi}(1 - P_{AU}) T_{UU} \\ E\{A_{UN}^{(i)}\} &= \theta_{AUi}(1 - P_{AU}) T_{UN} \\ E\{A_{UF}^{(i)}\} &= \theta_{AUi}(1 - P_{AU}) T_{UF} \\ E\{A_{UM}^{(i)}\} &= \theta_{AUi} P_{AU} T_{UT} \end{aligned} \tag{2.7}$$

Row U summed over i = 1,2,3,5,6,7

$$\begin{aligned} E\{A_{UE}\} &= (1 - \theta_{AU48})(1 - P_{AU}) T_{UE} \\ E\{A_{UU}\} &= (1 - \theta_{AU48})(1 - P_{AU}) T_{UU} \\ E\{A_{UN}\} &= (1 - \theta_{AU48})(1 - P_{AU}) T_{UN} \\ E\{A_{UF}\} &= (1 - \theta_{AU48})(1 - P_{AU}) T_{UF} \\ E\{A_{UM}\} &= (1 - \theta_{AU48}) P_{AU} T_{UT} \end{aligned} \tag{2.8}$$

Row U month-in-sample i = 4,8

$$E\{A_{UR}^{(i)}\} = \theta_{AUi} T_{UT} \tag{2.9}$$

Row U summed over i = 4,8

$$E\{A_{UR}\} = \theta_{AU48} T_{UT} \tag{2.10}$$

We similarly define row N.

Row N month-in-sample i = 1,2,3,5,6,7

$$\begin{aligned} E\{A_{NE}^{(i)}\} &= \theta_{ANi}(1 - P_{AN}) T_{NE} \\ E\{A_{NU}^{(i)}\} &= \theta_{ANi}(1 - P_{AN}) T_{NU} \\ E\{A_{NN}^{(i)}\} &= \theta_{ANi}(1 - P_{AN}) T_{NN} \\ E\{A_{NF}^{(i)}\} &= \theta_{ANi}(1 - P_{AN}) T_{NF} \\ E\{A_{NM}^{(i)}\} &= \theta_{ANi} P_{AN} T_{NT} \end{aligned} \tag{2.11}$$

Row N summed over i = 1,2,3,5,6,7

$$\begin{aligned} E\{A_{NE}\} &= (1 - \theta_{AN48})(1 - P_{AN}) T_{NE} \\ E\{A_{NU}\} &= (1 - \theta_{AN48})(1 - P_{AN}) T_{NU} \\ E\{A_{NN}\} &= (1 - \theta_{AN48})(1 - P_{AN}) T_{NN} \\ E\{A_{NF}\} &= (1 - \theta_{AN48})(1 - P_{AN}) T_{NF} \\ E\{A_{NM}\} &= (1 - \theta_{AN48}) P_{AN} T_{NT} \end{aligned} \tag{2.12}$$

Row N month-in-sample $i = 4,8$

$$E\{A_{NR}^{(i)}\} = \theta_{ANi}T_{NT} \tag{2.13}$$

Row N summed over $i = 4,8$

$$E\{A_{NR}\} = \theta_{AN48}T_{NT} \tag{2.14}$$

We next define the following expectations in terms of the columns of matrix B for each month-in-sample. We begin with column E.

Column E month-in-sample $i = 2,3,4,6,7,8$

$$\begin{aligned} E\{B_{EE}^{(i)}\} &= \theta_{BEi}(1 - P_{BE})T_{EE} \\ E\{B_{UE}^{(i)}\} &= \theta_{BEi}(1 - P_{BE})T_{UE} \\ E\{B_{NE}^{(i)}\} &= \theta_{BEi}(1 - P_{BE})T_{NE} \\ E\{B_{FE}^{(i)}\} &= \theta_{BEi}(1 - P_{BE})T_{FE} \\ E\{B_{ME}^{(i)}\} &= \theta_{BEi}P_{BE}T_{TE} \end{aligned} \tag{2.15}$$

Column E summed over $i = 2,3,4,6,7,8$

$$\begin{aligned} E\{B_{EE}\} &= (1 - \theta_{BE15})(1 - P_{BE})T_{EE} \\ E\{B_{UE}\} &= (1 - \theta_{BE15})(1 - P_{BE})T_{UE} \\ E\{B_{NE}\} &= (1 - \theta_{BE15})(1 - P_{BE})T_{NE} \\ E\{B_{FE}\} &= (1 - \theta_{BE15})(1 - P_{BE})T_{FE} \\ E\{B_{ME}\} &= (1 - \theta_{BE15})P_{BE}T_{TE} \end{aligned} \tag{2.16}$$

Column E month-in-sample $i = 1,5$

$$E\{B_{RE}^{(i)}\} = \theta_{BEi}T_{TE} \tag{2.17}$$

Column E summed over $i = 1,5$

$$E\{B_{RE}\} = \theta_{BE15}T_{TE} \tag{2.18}$$

We define $\theta_{BE15} = \theta_{BE1} + \theta_{BE5}$. We similarly define column U.

Column U month-in-sample $i = 2,3,4,6,7,8$

$$\begin{aligned} E\{B_{EU}^{(i)}\} &= \theta_{BUi}(1 - P_{BU})T_{EU} \\ E\{B_{UU}^{(i)}\} &= \theta_{BUi}(1 - P_{BU})T_{UU} \\ E\{B_{NU}^{(i)}\} &= \theta_{BUi}(1 - P_{BU})T_{NU} \\ E\{B_{FU}^{(i)}\} &= \theta_{BUi}(1 - P_{BU})T_{FU} \\ E\{B_{MU}^{(i)}\} &= \theta_{BUi}P_{BU}T_{TU} \end{aligned} \tag{2.19}$$

Column U summed over $i = 2,3,4,6,7,8$

$$\begin{aligned} E\{B_{EU}\} &= (1 - \theta_{BU15})(1 - P_{BU})T_{EU} \\ E\{B_{UU}\} &= (1 - \theta_{BU15})(1 - P_{BU})T_{UU} \\ E\{B_{NU}\} &= (1 - \theta_{BU15})(1 - P_{BU})T_{NU} \\ E\{B_{FU}\} &= (1 - \theta_{BU15})(1 - P_{BU})T_{FU} \\ E\{B_{MU}\} &= (1 - \theta_{BU15})P_{BU}T_{TU} \end{aligned} \tag{2.20}$$

Column U month-in-sample $i = 1,5$

$$E\{B_{RU}^{(i)}\} = \theta_{BUi}T_{TU} \tag{2.21}$$

Column U summed over i = 1,5

$$E\{B_{RU}\} = \theta_{BU15}T_{TU} \tag{2.22}$$

We similarly define column N.

Column N month-in-sample i = 2,3,4,6,7,8

$$\begin{aligned} E\{B_{EN}^{(i)}\} &= \theta_{BNi}(1-P_{BN})T_{EN} \\ E\{B_{UN}^{(i)}\} &= \theta_{BNi}(1-P_{BN})T_{UN} \\ E\{B_{NN}^{(i)}\} &= \theta_{BNi}(1-P_{BN})T_{NN} \\ E\{B_{FN}^{(i)}\} &= \theta_{BNi}(1-P_{BN})T_{FN} \\ E\{B_{MN}^{(i)}\} &= \theta_{BNi}P_{BN}T_{TN} \end{aligned} \tag{2.23}$$

Column N summed over i = 2,3,4,6,7,8

$$\begin{aligned} E\{B_{EN}\} &= (1-\theta_{BN15})(1-P_{BN})T_{EN} \\ E\{B_{UN}\} &= (1-\theta_{BN15})(1-P_{BN})T_{UN} \\ E\{B_{NN}\} &= (1-\theta_{BN15})(1-P_{BN})T_{NN} \\ E\{B_{FN}\} &= (1-\theta_{BN15})(1-P_{BU})T_{FN} \\ E\{B_{MN}\} &= (1-\theta_{BN15})P_{BN}T_{TN} \end{aligned} \tag{2.24}$$

Column N month-in-sample i = 1,5

$$E\{B_{RN}^{(i)}\} = \theta_{BNi}T_{TN} \tag{2.25}$$

Column N summed over i = 1,5

$$E\{B_{RN}\} = \theta_{BN15}T_{TU} \tag{2.26}$$

2.2 Modeling the Partial Gross Flows Tables by Month-in-Sample

We next define the following expectations in terms of the rows of matrix A for each month-in-sample. We begin with row E.

Row E month-in-sample i = 1,2,3,5,6,7

$$\begin{aligned} E\{A_{EE}^{(i)}\} &= \theta_{AEi}(1-P_{AEi})T_{EE} \\ E\{A_{EU}^{(i)}\} &= \theta_{AEi}(1-P_{AEi})T_{EU} \\ E\{A_{EN}^{(i)}\} &= \theta_{AEi}(1-P_{AEi})T_{EN} \\ E\{A_{EF}^{(i)}\} &= \theta_{AEi}(1-P_{AEi})T_{EF} \\ E\{A_{EM}^{(i)}\} &= \theta_{AEi}P_{AEi}T_{ET} \end{aligned} \tag{2.27}$$

Row E month-in-sample i = 4,8

$$E\{A_{ER}^{(i)}\} = \theta_{AEi}T_{ET} \tag{2.28}$$

The parameters θ_{AEi} for $i = 1, 2, 3, 5, 6, 7$ are the month-in-sample effects for Employment in month $t-1$ where by definition

$$\sum_{i=1}^{8} \theta_{AEi} = 1 \tag{2.29}$$

The parameters P_{AEi} represent what percent of the totals $T_{EE}, T_{EU}, T_{EN}, T_{EF}$ flow into the missing category for employed. For convenience we define

$$T_{ET} = T_{EE} + T_{EU} + T_{EN} + T_{EF} \tag{2.30}$$

We similarly define row U.

Row U month-in-sample i = 1,2,3,5,6,7

$$\begin{aligned} E\{A_{UE}^{(i)}\} &= \theta_{AUi}(1-P_{AUi})T_{UE} \\ E\{A_{UU}^{(i)}\} &= \theta_{AUi}(1-P_{AUi})T_{UU} \\ E\{A_{UN}^{(i)}\} &= \theta_{AUi}(1-P_{AUi})T_{UN} \\ E\{A_{UF}^{(i)}\} &= \theta_{AUi}(1-P_{AUi})T_{UF} \\ E\{A_{UM}^{(i)}\} &= \theta_{AUi}P_{AUi}T_{UT} \end{aligned} \quad (2.31)$$

Row U month-in-sample i = 4,8

$$E\{A_{UR}^{(i)}\} = \theta_{AUi}T_{UT} \quad (2.32)$$

Row N month-in-sample i = 1,2,3,5,6,7

$$\begin{aligned} E\{A_{NE}^{(i)}\} &= \theta_{ANi}(1-P_{ANi})T_{NE} \\ E\{A_{NU}^{(i)}\} &= \theta_{ANi}(1-P_{ANi})T_{NU} \\ E\{A_{NN}^{(i)}\} &= \theta_{ANi}(1-P_{ANi})T_{NN} \\ E\{A_{NF}^{(i)}\} &= \theta_{ANi}(1-P_{ANi})T_{NF} \\ E\{A_{NM}^{(i)}\} &= \theta_{ANi}P_{ANi}T_{NT} \end{aligned} \quad (2.33)$$

Row N month-in-sample i = 4,8

$$E\{A_{NR}^{(i)}\} = \theta_{ANi}T_{NT} \quad (2.34)$$

We next define the following expectations in terms of the columns of matrix B for each month-in-sample. We begin with column E.

Column E month-in-sample i = 2,3,4,6,7,8

$$\begin{aligned} E\{B_{EE}^{(i)}\} &= \theta_{BEi}(1-P_{BEi})T_{EE} \\ E\{B_{UE}^{(i)}\} &= \theta_{BEi}(1-P_{BEi})T_{UE} \\ E\{B_{NE}^{(i)}\} &= \theta_{BEi}(1-P_{BEi})T_{NE} \\ E\{B_{FE}^{(i)}\} &= \theta_{BEi}(1-P_{BEi})T_{FE} \\ E\{B_{ME}^{(i)}\} &= \theta_{BEi}P_{BEi}T_{TE} \end{aligned} \quad (2.35)$$

Column E month-in-sample i = 1,5

$$E\{B_{RE}^{(i)}\} = \theta_{BEi}T_{TE} \quad (2.36)$$

Column U month-in-sample i = 2,3,4,6,7,8

$$\begin{aligned} E\{B_{EU}^{(i)}\} &= \theta_{BUi}(1-P_{BUi})T_{EU} \\ E\{B_{UU}^{(i)}\} &= \theta_{BUi}(1-P_{BUi})T_{UU} \\ E\{B_{NU}^{(i)}\} &= \theta_{BUi}(1-P_{BUi})T_{NU} \\ E\{B_{FU}^{(i)}\} &= \theta_{BUi}(1-P_{BUi})T_{FU} \\ E\{B_{MU}^{(i)}\} &= \theta_{BUi}P_{BUi}T_{TU} \end{aligned} \quad (2.37)$$

Column U month-in-sample i = 1,5

$$E\{B_{RU}^{(i)}\} = \theta_{BUi}T_{TU} \quad (2.38)$$

Column N month-in-sample i = 2,3,4,6,7,8

$$\begin{aligned} E\{B_{EN}^{(i)}\} &= \theta_{BNi}(1-P_{BNi})T_{EN} \\ E\{B_{UN}^{(i)}\} &= \theta_{BNi}(1-P_{BNi})T_{UN} \\ E\{B_{NN}^{(i)}\} &= \theta_{BNi}(1-P_{BNi})T_{NN} \\ E\{B_{FN}^{(i)}\} &= \theta_{BNi}(1-P_{BNi})T_{FN} \\ E\{B_{MN}^{(i)}\} &= \theta_{BNi}P_{BNi}T_{TN} \end{aligned} \tag{2.39}$$

Column N month-in-sample i = 1,5

$$E\{B_{RN}^{(i)}\} = \theta_{BNi}T_{TN} \tag{2.40}$$

3. Estimating the Model Parameters

We define an objective function F which resembles a log-likelihood equation for a multinomial distribution as Stasny (1988) used. We do not assume a multinomial distribution but still employ the form as a useful objective function. We will obtain parameter estimates by maximizing the objective function under the constraints that the gross flow table parameters match the row and column stock totals. We define two distinct objective function: F for the aggregate partial gross flows tables and F^* for the partial gross flows tables by month-in-sample.

3.1 Defining the Objective Function for the Aggregated Partial Gross Flows Tables

Define F as

$$\begin{aligned} F = \; & A_{EE}\ln(E\{A_{EE}\}) + A_{EU}\ln(E\{A_{EU}\}) + A_{EN}\ln(E\{A_{EN}\}) + \\ & A_{EF}\ln(E\{A_{EF}\}) + A_{EM}\ln(E\{A_{EM}\}) + A_{ER}\ln(E\{A_{ER}\}) + \\ & A_{UE}\ln(E\{A_{UE}\}) + A_{UU}\ln(E\{A_{UU}\}) + A_{UN}\ln(E\{A_{UN}\}) + \\ & A_{UF}\ln(E\{A_{UF}\}) + A_{UM}\ln(E\{A_{UM}\}) + A_{UR}\ln(E\{A_{UR}\}) + \\ & A_{NE}\ln(E\{A_{NE}\}) + A_{NU}\ln(E\{A_{NU}\}) + A_{NN}\ln(E\{A_{NN}\}) + \\ & A_{NF}\ln(E\{A_{NF}\}) + A_{NM}\ln(E\{A_{NM}\}) + A_{NR}\ln(E\{A_{NR}\}) + \\ & B_{EE}\ln(E\{B_{EE}\}) + B_{UE}\ln(E\{B_{UE}\}) + B_{NE}\ln(E\{B_{NE}\}) + \\ & B_{FE}\ln(E\{B_{FE}\}) + B_{ME}\ln(E\{B_{ME}\}) + B_{RE}\ln(E\{B_{RE}\}) + \\ & B_{EU}\ln(E\{B_{EU}\}) + B_{UU}\ln(E\{B_{UU}\}) + B_{NU}\ln(E\{B_{NU}\}) + \\ & B_{FU}\ln(E\{B_{FU}\}) + B_{MU}\ln(E\{B_{MU}\}) + B_{RU}\ln(E\{B_{RU}\}) + \\ & B_{EN}\ln(E\{B_{EN}\}) + B_{UN}\ln(E\{B_{UN}\}) + B_{NN}\ln(E\{B_{NN}\}) + \\ & B_{FN}\ln(E\{B_{FN}\}) + B_{MN}\ln(E\{B_{MN}\}) + B_{RN}\ln(E\{B_{RN}\}) \\ & +\lambda_1(T_{EE}+T_{EU}+T_{EN}+T_{EF}-A_{ET}-A_{ER}) \\ & +\lambda_2(T_{UE}+T_{UU}+T_{UN}+T_{UF}-A_{UT}-A_{UR}) \\ & +\lambda_3(T_{NE}+T_{NU}+T_{NN}+T_{NF}-A_{NT}-A_{NR}) \\ & +\lambda_4(T_{EE}+T_{UE}+T_{NE}+T_{FE}-B_{TE}-B_{RE}) \\ & +\lambda_5(T_{EU}+T_{UU}+T_{NU}+T_{FU}-B_{TU}-B_{RU}) \\ & +\lambda_6(T_{EN}+T_{UN}+T_{NN}+T_{FN}-B_{TN}-B_{RN}) \end{aligned} \tag{3.1}$$

where

$$
\begin{aligned}
A_{ET} &= A_{EE} + A_{EU} + A_{EN} + A_{EF} + A_{EM} \\
A_{UT} &= A_{UE} + A_{UU} + A_{UN} + A_{UF} + A_{UM} \\
A_{NT} &= A_{NE} + A_{NU} + A_{NN} + A_{NF} + A_{NM} \\
B_{TE} &= B_{EE} + B_{UE} + B_{NE} + B_{FE} + B_{ME} \\
B_{TU} &= B_{EU} + B_{UU} + B_{NU} + B_{FU} + B_{MU} \\
B_{TN} &= B_{EN} + B_{UN} + B_{NN} + B_{FN} + B_{MN}
\end{aligned}
\tag{3.2}
$$

Note that F is a function of the parameters

$$
\begin{aligned}
& T_{EE}, T_{EU}, T_{EN}, T_{EF} \\
& T_{UE}, T_{UU}, T_{UN}, T_{UF} \\
& T_{NE}, T_{NU}, T_{NN}, T_{NF} \\
& T_{FE}, T_{FU}, T_{FN} \\
& \theta_{AE48}, \theta_{AU48}, \theta_{AN48}, \theta_{BE15}, \theta_{BU15}, \theta_{BN15} \\
& P_{AE}, P_{AU}, P_{AN}, P_{BE}, P_{BU}, P_{BN} \\
& \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6
\end{aligned}
\tag{3.3}
$$

where the λ_i are the Lagrange multipliers to ensure the gross flows table matches the row and column stock estimates. With some algebra we can decompose F into

$$
F = F_1 + F_2 + F_3 \tag{3.4}
$$

where

$$
\begin{aligned}
F_1 = {} & A_{ET} \ln(1 - \theta_{AE48}) + A_{ER} \ln(\theta_{AE48}) + \\
& A_{UT} \ln(1 - \theta_{AU48}) + A_{UR} \ln(\theta_{AU48}) + \\
& A_{NT} \ln(1 - \theta_{AN48}) + A_{NR} \ln(\theta_{AN48}) + \\
& B_{TE} \ln(1 - \theta_{BE15}) + B_{RE} \ln(\theta_{BE15}) + \\
& B_{TU} \ln(1 - \theta_{BU15}) + B_{RU} \ln(\theta_{BU15}) + \\
& B_{TN} \ln(1 - \theta_{BN15}) + B_{RN} \ln(\theta_{BN15})
\end{aligned}
\tag{3.5}
$$

$$
\begin{aligned}
F_2 = {} & (A_{EE} + A_{EU} + A_{EN} + A_{EF}) \ln(1 - P_{AE}) + A_{EM} \ln(P_{AE}) \\
& (A_{UE} + A_{UU} + A_{UN} + A_{UF}) \ln(1 - P_{AU}) + A_{UM} \ln(P_{AU}) \\
& (A_{NE} + A_{NU} + A_{NN} + A_{NF}) \ln(1 - P_{AN}) + A_{NM} \ln(P_{AN}) \\
& (B_{EE} + B_{UE} + B_{NE} + B_{FE}) \ln(1 - P_{BE}) + B_{ME} \ln(P_{BE}) \\
& (B_{EU} + B_{UU} + B_{NU} + B_{FU}) \ln(1 - P_{BU}) + B_{MU} \ln(P_{BU}) \\
& (B_{EN} + B_{UN} + B_{NN} + B_{FN}) \ln(1 - P_{BN}) + B_{MN} \ln(P_{BN})
\end{aligned}
\tag{3.6}
$$

$$
\begin{aligned}
F_3 \quad = \quad & (A_{EE} + B_{EE}) \ln(T_{EE}) + (A_{EU} + B_{EU}) \ln(T_{EU}) + \\
& (A_{EN} + B_{EN}) \ln(T_{EN}) + A_{EF} \ln(T_{EF}) + \\
& (A_{EM} + A_{ER}) \ln(T_{ET}) + \\
& (A_{UE} + B_{UE}) \ln(T_{UE}) + (A_{UU} + B_{UU}) \ln(T_{UU}) + \\
& (A_{UN} + B_{UN}) \ln(T_{UN}) + A_{UF}) \ln(T_{UF}) + \\
& (A_{UM} + A_{UR}) \ln(T_{UT}) + \\
& (A_{NE} + B_{NE}) \ln(T_{NE}) + (A_{NU} + B_{NU}) \ln(T_{NU}) + \\
& (A_{NN} + B_{NN}) \ln(T_{NN}) + A_{NF}) \ln(T_{NF}) + \\
& (A_{NM} + A_{NR}) \ln(T_{NT}) + \\
& B_{FE} \ln(T_{FE}) + (B_{ME} + B_{RE}) \ln(T_{TE}) + \\
& B_{FU} \ln(T_{FU}) + (B_{MU} + B_{RU}) \ln(T_{TU}) + \\
& B_{FN} \ln(T_{FN}) + (B_{MN} + B_{RE}) \ln(T_{TN}) + \\
& +\lambda_1(T_{EE} + T_{EU} + T_{EN} + T_{EF} - A_{ET} - A_{ER}) \\
& +\lambda_2(T_{UE} + T_{UU} + T_{UN} + T_{UF} - A_{UT} - A_{UR}) \\
& +\lambda_3(T_{NE} + T_{NU} + T_{NN} + T_{NF} - A_{NT} - A_{NR}) \\
& +\lambda_4(T_{EE} + T_{UE} + T_{NE} + T_{FE} - B_{TE} - B_{RE}) \\
& +\lambda_5(T_{EU} + T_{UU} + T_{NU} + T_{FU} - B_{TU} - B_{RU}) \\
& +\lambda_6(T_{EN} + T_{UN} + T_{NN} + T_{FN} - B_{TN} - B_{RN})
\end{aligned} \tag{3.7}
$$

3.2 Defining the Objective Function for the Partial Gross Flows Tables by Month-in-Sample

We can also write a second objective function F^*. First define

$$
\begin{aligned}
S_A &= \{1, 2, 3, 5, 6, 7\} \\
S_A^c &= \{4, 8\} \\
S_B &= \{2, 3, 4, 6, 7, 8\} \\
S_B^c &= \{1, 5\}
\end{aligned} \tag{3.8}
$$

Now define

$$
\begin{aligned}
F^* \quad = \quad & \sum_{i\in S_A} A_{EE}^{(i)} \ln(E\{A_{EE}^{(i)}\}) + \sum_{i\in S_A} A_{EU}^{(i)} \ln(E\{A_{EU}^{(i)}\}) + \\
& \sum_{i\in S_A} A_{EN}^{(i)} \ln(E\{A_{EN}^{(i)}\}) + \sum_{i\in S_A} A_{EF}^{(i)} \ln(E\{A_{EF}^{(i)}\}) + \\
& \sum_{i\in S_A} A_{EM}^{(i)} \ln(E\{A_{EM}^{(i)}\}) + \sum_{i\in S_A^c} A_{ER}^{(i)} \ln(E\{A_{ER}^{(i)}\}) + \\
& \sum_{i\in S_A} A_{UE}^{(i)} \ln(E\{A_{UE}^{(i)}\}) + \sum_{i\in S_A} A_{UU}^{(i)} \ln(E\{A_{UU}^{(i)}\}) + \\
& \sum_{i\in S_A} A_{UN}^{(i)} \ln(E\{A_{UN}^{(i)}\}) + \sum_{i\in S_A} A_{UF}^{(i)} \ln(E\{A_{UF}^{(i)}\}) + \\
& \sum_{i\in S_A} A_{UM}^{(i)} \ln(E\{A_{UM}^{(i)}\}) + \sum_{i\in S_A^c} A_{UR}^{(i)} \ln(E\{A_{UR}^{(i)}\}) + \\
& \sum_{i\in S_A} A_{NE}^{(i)} \ln(E\{A_{NE}^{(i)}\}) + \sum_{i\in S_A} A_{NU}^{(i)} \ln(E\{A_{NU}^{(i)}\}) + \\
& \sum_{i\in S_A} A_{NN}^{(i)} \ln(E\{A_{NN}^{(i)}\}) + \sum_{i\in S_A} A_{NF}^{(i)} \ln(E\{A_{NF}^{(i)}\}) + \\
& \sum_{i\in S_A} A_{NM}^{(i)} \ln(E\{A_{NM}^{(i)}\}) + \sum_{i\in S_A^c} A_{NR}^{(i)} \ln(E\{A_{NR}^{(i)}\}) + \\
& \sum_{i\in S_B} B_{EE}^{(i)} \ln(E\{B_{EE}^{(i)}\}) + \sum_{i\in S_B} B_{UE}^{(i)} \ln(E\{B_{UE}^{(i)}\}) + \\
& \sum_{i\in S_B} B_{NE}^{(i)} \ln(E\{B_{NE}^{(i)}\}) + \sum_{i\in S_B} B_{FE}^{(i)} \ln(E\{B_{FE}^{(i)}\}) + \\
& \sum_{i\in S_B} B_{ME}^{(i)} \ln(E\{B_{ME}^{(i)}\}) + \sum_{i\in S_B^c} B_{RE}^{(i)} \ln(E\{B_{RE}^{(i)}\}) + \\
& \sum_{i\in S_B} B_{EU}^{(i)} \ln(E\{B_{EU}^{(i)}\}) + \sum_{i\in S_B} B_{UU}^{(i)} \ln(E\{B_{UU}^{(i)}\}) + \\
& \sum_{i\in S_B} B_{NU}^{(i)} \ln(E\{B_{NU}^{(i)}\}) + \sum_{i\in S_B} B_{FU}^{(i)} \ln(E\{B_{FU}^{(i)}\}) + \\
& \sum_{i\in S_B} B_{MU}^{(i)} \ln(E\{B_{MU}^{(i)}\}) + \sum_{i\in S_B^c} B_{RU}^{(i)} \ln(E\{B_{RU}^{(i)}\}) + \\
& \sum_{i\in S_B} B_{EN}^{(i)} \ln(E\{B_{EN}^{(i)}\}) + \sum_{i\in S_B} B_{UN}^{(i)} \ln(E\{B_{UN}^{(i)}\}) + \\
& \sum_{i\in S_B} A_{NN}^{(i)} \ln(E\{B_{NN}^{(i)}\}) + \sum_{i\in S_B} B_{FN}^{(i)} \ln(E\{B_{FN}^{(i)}\}) + \\
& \sum_{i\in S_B} B_{MN}^{(i)} \ln(E\{B_{MN}^{(i)}\}) + \sum_{i\in S_{B^c}} B_{RN}^{(i)} \ln(E\{B_{RN}^{(i)}\}) \\
& +\lambda_1(T_{EE} + T_{EU} + T_{EN} + T_{EF} - A_{ET} - A_{ER}) \\
& +\lambda_2(T_{UE} + T_{UU} + T_{UN} + T_{UF} - A_{UT} - A_{UR}) \\
& +\lambda_3(T_{NE} + T_{NU} + T_{NN} + T_{NF} - A_{NT} - A_{NR}) \\
& +\lambda_4(T_{EE} + T_{UE} + T_{NE} + T_{FE} - B_{TE} - B_{RE}) \\
& +\lambda_5(T_{EU} + T_{UU} + T_{NU} + T_{FU} - B_{TU} - B_{RU}) \\
& +\lambda_6(T_{EN} + T_{UN} + T_{NN} + T_{FN} - B_{TN} - B_{RN})
\end{aligned} \tag{3.9}
$$

Note that F^* is a function of the parameters

$$\begin{aligned}
&T_{EE}, T_{EU}, T_{EN}, T_{EF} \\
&T_{UE}, T_{UU}, T_{UN}, T_{UF} \\
&T_{NE}, T_{NU}, T_{NN}, T_{NF} \\
&T_{FE}, T_{FU}, T_{FN} \\
&\theta_{AEi}, \theta_{AUi}, \theta_{ANi} \text{ for } i = 1,2,3,4,5,6,7,8 \\
&\theta_{BEi}, \theta_{BUi}, \theta_{BNi} \text{ for } i = 1,2,3,4,5,6,7,8 \\
&P_{AEi}, P_{AUi}, P_{ANi} \text{ for } i = 1,2,3,5,6,7 \\
&P_{BEi}, P_{BUi}, P_{BNi} \text{ for } i = 2,3,4,6,7,8 \\
&\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6
\end{aligned} \tag{3.10}$$

where the λ_i are the Lagrange multipliers to ensure the gross flows table matches the row and column stock estimates. F^* can be decomposed as

$$F^* = F_1^* + F_2^* + F_3 \tag{3.11}$$

where F_3 remains the same as the decomposition of F, while F_1^* and F_2^* are given by

$$\begin{aligned}
F_1^* \quad = \quad &\sum_{i \in S_A} A_{ET}^{(i)} \ln(\theta_{AEi}) + \sum_{i \in S_A^c} A_{ER}^{(i)} \ln(\theta_{AEi}) + \\
&\sum_{i \in S_A} A_{UT}^{(i)} \ln(\theta_{AUi}) + \sum_{i \in S_A^c} A_{UR}^{(i)} \ln(\theta_{AUi}) + \\
&\sum_{i \in S_A} A_{NT}^{(i)} \ln(\theta_{ANi}) + \sum_{i \in S_A^c} A_{NR}^{(i)} \ln(\theta_{ANi}) + \\
&\sum_{i \in S_B} B_{TE}^{(i)} \ln(\theta_{BEi}) + \sum_{i \in S_B^c} B_{RE}^{(i)} \ln(\theta_{BEi}) + \\
&\sum_{i \in S_B} B_{TU}^{(i)} \ln(\theta_{BUi}) + \sum_{i \in S_B^c} B_{RU}^{(i)} \ln(\theta_{BUi}) + \\
&\sum_{i \in S_B} B_{TN}^{(i)} \ln(\theta_{BNi}) + \sum_{i \in S_B^c} B_{RN}^{(i)} \ln(\theta_{BNi}) +
\end{aligned} \tag{3.12}$$

$$\begin{aligned}
F_2^* \quad = \quad &\sum_{i \in S_A} \Big[(A_{EE}^{(i)} + A_{EU}^{(i)} + A_{EN}^{(i)} + A_{EF}^{(i)}) \ln(1 - P_{AEi}) + A_{EM}^{(i)} \ln(P_{AEi})\Big] \\
&\sum_{i \in S_A} \Big[(A_{UE}^{(i)} + A_{UU}^{(i)} + A_{UN}^{(i)} + A_{UF}^{(i)}) \ln(1 - P_{AUi}) + A_{UM}^{(i)} \ln(P_{AUi})\Big] \\
&\sum_{i \in S_A} \Big[(A_{NE}^{(i)} + A_{NU}^{(i)} + A_{NN}^{(i)} + A_{NF}^{(i)}) \ln(1 - P_{ANi}) + A_{NM}^{(i)} \ln(P_{ANi})\Big] \\
&\sum_{i \in S_B} \Big[(B_{EE}^{(i)} + B_{UE}^{(i)} + B_{NE}^{(i)} + B_{FE}^{(i)}) \ln(1 - P_{BEi}) + B_{ME}^{(i)} \ln(P_{BEi})\Big] \\
&\sum_{i \in S_B} \Big[(B_{EU}^{(i)} + B_{UU}^{(i)} + B_{NU}^{(i)} + B_{FU}^{(i)}) \ln(1 - P_{BUi}) + B_{MU}^{(i)} \ln(P_{BUi})\Big] \\
&\sum_{i \in S_B} \Big[(B_{EN}^{(i)} + B_{UN}^{(i)} + B_{NN}^{(i)} + B_{FN}^{(i)}) \ln(1 - P_{BNi}) + B_{MN}^{(i)} \ln(P_{BNi})\Big]
\end{aligned} \tag{3.13}$$

3.3 Estimation of Month-in-Sample Effects

F_1 defined in 3.5 and F_1^* defined in 3.12 are the two functions we need to maximize to obtain the estimated month-in-sample effects. We begin with F_1 and the parameter θ_{AE48}.

$$\frac{\partial F_1}{\partial \theta_{AE48}} = -A_{ET}(1 - \theta_{AE48})^{-1} + A_{ER}\theta_{AE48}^{-1} \tag{3.14}$$

Setting that derivative equal to zero yields the estimator $\hat{\theta}_{AE48}$, and the others follow in a similar way.

$$\begin{aligned}
\hat{\theta}_{AE48} &= (A_{ET}+A_{ER})^{-1}A_{ER} \\
\hat{\theta}_{AU48} &= (A_{UT}+A_{UR})^{-1}A_{UR} \\
\hat{\theta}_{AN48} &= (A_{NT}+A_{NR})^{-1}A_{NR} \\
\hat{\theta}_{BE15} &= (B_{TE}+B_{RE})^{-1}B_{RE} \\
\hat{\theta}_{BU15} &= (B_{TU}+B_{RU})^{-1}B_{RU} \\
\hat{\theta}_{BN15} &= (B_{TN}+B_{RN})^{-1}B_{RN}
\end{aligned} \tag{3.15}$$

Similarly for F_1^* we obtain

$$\begin{aligned}
\hat{\theta}_{AEi} &= (A_{ET}+A_{ER})^{-1}A_{ET}^{(i)} \text{ for } i \in S_A \\
\hat{\theta}_{AEi} &= (A_{ET}+A_{ER})^{-1}A_{ER}^{(i)} \text{ for } i \in S_A^c \\
\hat{\theta}_{AUi} &= (A_{UT}+A_{UR})^{-1}A_{UT}^{(i)} \text{ for } i \in S_A \\
\hat{\theta}_{AUi} &= (A_{UT}+A_{UR})^{-1}A_{UR}^{(i)} \text{ for } i \in S_A^c \\
\hat{\theta}_{ANi} &= (A_{NT}+A_{NR})^{-1}A_{NT}^{(i)} \text{ for } i \in S_A \\
\hat{\theta}_{ANi} &= (A_{NT}+A_{NR})^{-1}A_{NR}^{(i)} \text{ for } i \in S_A^c \\
\hat{\theta}_{BEi} &= (B_{TE}+B_{RE})^{-1}B_{TE}^{(i)} \text{ for } i \in S_B \\
\hat{\theta}_{BEi} &= (B_{TE}+B_{RE})^{-1}B_{RE}^{(i)} \text{ for } i \in S_B^c \\
\hat{\theta}_{BUi} &= (B_{TU}+B_{RU})^{-1}B_{TU}^{(i)} \text{ for } i \in S_B \\
\hat{\theta}_{BUi} &= (B_{TU}+B_{RU})^{-1}B_{RU}^{(i)} \text{ for } i \in S_B^c \\
\hat{\theta}_{BNi} &= (B_{TN}+B_{RN})^{-1}B_{TN}^{(i)} \text{ for } i \in S_B \\
\hat{\theta}_{BNi} &= (B_{TN}+B_{RN})^{-1}B_{RN}^{(i)} \text{ for } i \in S_B^c
\end{aligned} \tag{3.16}$$

3.4 Estimation of Missing Rates

F_2 defined in 3.6 and F_2^* defined in 3.13 are the two functions we need to maximize to obtain the estimated missing rates. We begin with F_2 and the parameter P_{AE}.

$$\frac{\partial F_2}{\partial P_{AE}} = -(A_{EE}+A_{EU}+A_{EN}+A_{EF})(1-P_{AE})^{-1} + A_{EM}P_{AE}^{-1} \tag{3.17}$$

Setting that derivative equal to zero yields the estimator $\hat{P}_{AE}$, and the others follow in a similar way.

$$\begin{aligned}
\hat{P}_{AE} &= A_{ET}^{-1}A_{EM} \\
\hat{P}_{AU} &= A_{UT}^{-1}A_{UM} \\
\hat{P}_{AN} &= A_{NT}^{-1}A_{NM} \\
\hat{P}_{BE} &= B_{TE}^{-1}B_{ME} \\
\hat{P}_{BU} &= B_{TU}^{-1}B_{MU} \\
\hat{P}_{BN} &= B_{TN}^{-1}B_{MN}
\end{aligned} \tag{3.18}$$

Similarly for F_2^* we obtain

$$\begin{aligned}
\hat{P}_{AEi} &= (A_{ET}^{(i)})^{-1}A_{EM}^{(i)} \text{ for } i \in S_A \\
\hat{P}_{AUi} &= (A_{UT}^{(i)})^{-1}A_{UM}^{(i)} \text{ for } i \in S_A \\
\hat{P}_{ANi} &= (A_{NT}^{(i)})^{-1}A_{NM}^{(i)} \text{ for } i \in S_A \\
\hat{P}_{BEi} &= (B_{TE}^{(i)})^{-1}B_{ME}^{(i)} \text{ for } i \in S_B \\
\hat{P}_{BUi} &= (B_{TU}^{(i)})^{-1}B_{MU}^{(i)} \text{ for } i \in S_B \\
\hat{P}_{BNi} &= (B_{TN}^{(i)})^{-1}B_{MN}^{(i)} \text{ for } i \in S_B
\end{aligned} \tag{3.19}$$

3.5 Estimation of Population Flows

In this section we need to maximize the function F_3 in 3.7. We begin with the parameter T_{EE},

$$\frac{\partial F_3}{\partial T_{EE}} = T_{EE}^{-1}(A_{EE} + B_{EE}) + T_{ET}^{-1}(A_{EM} + A_{ER}) + T_{TE}^{-1}(B_{ME} + B_{RE}) - \lambda_1 - \lambda_4 \quad (3.20)$$

Setting this derivative equal to 0 gives

$$\hat{T}_{EE}^{-1}(A_{EE} + B_{EE}) + \hat{T}_{ET}^{-1}(A_{EM} + A_{ER}) + \hat{T}_{TE}^{-1}(B_{ME} + B_{RE}) = \hat{\lambda}_1 + \hat{\lambda}_4 \quad (3.21)$$

By constraint

$$\begin{aligned} \hat{T}_{ET} &= A_{ET} + A_{ER} \\ \hat{T}_{TE} &= B_{TE} + B_{RE} \end{aligned} \quad (3.22)$$

After some algebra we obtain

$$\hat{T}_{EE} = \frac{A_{EE} + B_{EE}}{\hat{\lambda}_1 - 1 + (1 - \hat{\theta}_{AE48})(1 - \hat{P}_{AE}) + \hat{\lambda}_4 - 1 + (1 - \hat{\theta}_{BE15})(1 - \hat{P}_{BE})} \quad (3.23)$$

Derivation of expressions for the other parameters follow in a similar way. The complete set of expressions is as follows.

$$\hat{T}_{EE} = \frac{A_{EE} + B_{EE}}{\hat{\lambda}_1 - 1 + (1 - \hat{\theta}_{AE48})(1 - \hat{P}_{AE}) + \hat{\lambda}_4 - 1 + (1 - \hat{\theta}_{BE15})(1 - \hat{P}_{BE})} \quad (3.24)$$

$$\hat{T}_{EU} = \frac{A_{EU} + B_{EU}}{\hat{\lambda}_1 - 1 + (1 - \hat{\theta}_{AE48})(1 - \hat{P}_{AE}) + \hat{\lambda}_5 - 1 + (1 - \hat{\theta}_{BU15})(1 - \hat{P}_{BU})}$$

$$\hat{T}_{EN} = \frac{A_{EN} + B_{EN}}{\hat{\lambda}_1 - 1 + (1 - \hat{\theta}_{AE48})(1 - \hat{P}_{AE}) + \hat{\lambda}_6 - 1 + (1 - \hat{\theta}_{BN15})(1 - \hat{P}_{BN})}$$

$$\hat{T}_{UE} = \frac{A_{UE} + B_{UE}}{\hat{\lambda}_2 - 1 + (1 - \hat{\theta}_{AU48})(1 - \hat{P}_{AU}) + \hat{\lambda}_4 - 1 + (1 - \hat{\theta}_{BE15})(1 - \hat{P}_{BE})}$$

$$\hat{T}_{UU} = \frac{A_{UU} + B_{UU}}{\hat{\lambda}_2 - 1 + (1 - \hat{\theta}_{AU48})(1 - \hat{P}_{AU}) + \hat{\lambda}_5 - 1 + (1 - \hat{\theta}_{BU15})(1 - \hat{P}_{BU})}$$

$$\hat{T}_{UN} = \frac{A_{UN} + B_{UN}}{\hat{\lambda}_2 - 1 + (1 - \hat{\theta}_{AU48})(1 - \hat{P}_{AU}) + \hat{\lambda}_6 - 1 + (1 - \hat{\theta}_{BN15})(1 - \hat{P}_{BN})}$$

$$\hat{T}_{NE} = \frac{A_{NE} + B_{NE}}{\hat{\lambda}_3 - 1 + (1 - \hat{\theta}_{AN48})(1 - \hat{P}_{AN}) + \hat{\lambda}_4 - 1 + (1 - \hat{\theta}_{BE15})(1 - \hat{P}_{BE})}$$

$$\hat{T}_{NU} = \frac{A_{NU} + B_{NU}}{\hat{\lambda}_3 - 1 + (1 - \hat{\theta}_{AN48})(1 - \hat{P}_{AN}) + \hat{\lambda}_5 - 1 + (1 - \hat{\theta}_{BU15})(1 - \hat{P}_{BU})}$$

$$\hat{T}_{NN} = \frac{A_{NN} + B_{NN}}{\hat{\lambda}_3 - 1 + (1 - \hat{\theta}_{AN48})(1 - \hat{P}_{AN}) + \hat{\lambda}_6 - 1 + (1 - \hat{\theta}_{BN15})(1 - \hat{P}_{BN})}$$

$$\hat{T}_{EF} = \frac{A_{EF}}{\hat{\lambda}_1 - 1 + (1 - \hat{\theta}_{AE48})(1 - \hat{P}_{AE})}$$

$$\hat{T}_{UF} = \frac{A_{UF}}{\hat{\lambda}_2 - 1 + (1 - \hat{\theta}_{AU48})(1 - \hat{P}_{AU})}$$

$$\hat{T}_{NF} = \frac{A_{NF}}{\hat{\lambda}_3 - 1 + (1 - \hat{\theta}_{AN48})(1 - \hat{P}_{AN})}$$

$$\hat{T}_{FE} = \frac{B_{FE}}{\hat{\lambda}_4 - 1 + (1 - \hat{\theta}_{BE15})(1 - \hat{P}_{BE})}$$

$$\hat{T}_{FU} = \frac{B_{FU}}{\hat{\lambda}_5 - 1 + (1 - \hat{\theta}_{BU15})(1 - \hat{P}_{BU})}$$

$$\hat{T}_{FN} = \frac{B_{FN}}{\hat{\lambda}_6 - 1 + (1 - \hat{\theta}_{BN15})(1 - \hat{P}_{BN})}$$

The form of the estimates in 3.24 suggests an alternative parameterization and estimation method. We can write

$$\begin{aligned}
\hat{T}_{EE} &= (\hat{\alpha}_E + \hat{\beta}_E)^{-1}(A_{EE} + B_{EE}) \\
\hat{T}_{EU} &= (\hat{\alpha}_E + \hat{\beta}_U)^{-1}(A_{EU} + B_{EU}) \\
\hat{T}_{EN} &= (\hat{\alpha}_E + \hat{\beta}_N)^{-1}(A_{EN} + B_{EN}) \\
\hat{T}_{UE} &= (\hat{\alpha}_U + \hat{\beta}_E)^{-1}(A_{UE} + B_{UE}) \\
\hat{T}_{UU} &= (\hat{\alpha}_U + \hat{\beta}_U)^{-1}(A_{UU} + B_{UU}) \\
\hat{T}_{UN} &= (\hat{\alpha}_U + \hat{\beta}_N)^{-1}(A_{UN} + B_{N}) \\
\hat{T}_{NE} &= (\hat{\alpha}_N + \hat{\beta}_E)^{-1}(A_{NE} + B_{NE}) \\
\hat{T}_{NU} &= (\hat{\alpha}_N + \hat{\beta}_U)^{-1}(A_{NU} + B_{NU}) \\
\hat{T}_{NN} &= (\hat{\alpha}_N + \hat{\beta}_N)^{-1}(A_{NN} + B_{NN}) \\
\hat{T}_{EF} &= (\hat{\alpha}_E)^{-1}(A_{EF}) \\
\hat{T}_{UF} &= (\hat{\alpha}_U)^{-1}(A_{UF}) \\
\hat{T}_{NF} &= (\hat{\alpha}_N)^{-1}(A_{NF}) \\
\hat{T}_{FE} &= (\hat{\beta}_E)^{-1}(B_{FE}) \\
\hat{T}_{FU} &= (\hat{\beta}_U)^{-1}(B_{FU}) \\
\hat{T}_{FN} &= (\hat{\beta}_N)^{-1}(B_{FN})
\end{aligned} \tag{3.25}$$

where $\hat{(}\alpha_E, \hat{\alpha}_U, \hat{\alpha}_U, \hat{\beta}_E, \hat{\beta}_U, \hat{\beta}_N)$ are derived by solving the equations

$$\begin{aligned}
A_{ET} + A_{ER} &= \frac{A_{EE} + B_{EE}}{\alpha_E + \beta_E} + \frac{A_{EU} + B_{EU}}{\alpha_E + \beta_U} + \frac{A_{EN} + B_{EN}}{\alpha_E + \beta_N} + \frac{A_{EF}}{\alpha_E} \\
A_{UT} + A_{UR} &= \frac{A_{UE} + B_{UE}}{\alpha_U + \beta_E} + \frac{A_{UU} + B_{UU}}{\alpha_U + \beta_U} + \frac{A_{UN} + B_{UN}}{\alpha_U + \beta_N} + \frac{A_{UF}}{\alpha_U} \\
A_{NT} + A_{NR} &= \frac{A_{NE} + B_{NE}}{\alpha_N + \beta_E} + \frac{A_{NU} + B_{NU}}{\alpha_N + \beta_U} + \frac{A_{NN} + B_{NN}}{\alpha_N + \beta_N} + \frac{A_{NF}}{\alpha_N} \\
B_{TE} + B_{RE} &= \frac{A_{EE} + B_{EE}}{\alpha_E + \beta_E} + \frac{A_{UE} + B_{UE}}{\alpha_U + \beta_E} + \frac{A_{NE} + B_{NE}}{\alpha_N + \beta_E} + \frac{B_{FE}}{\beta_E} \\
B_{TU} + B_{RU} &= \frac{A_{EU} + B_{EU}}{\alpha_E + \beta_U} + \frac{A_{UU} + B_{UU}}{\alpha_U + \beta_U} + \frac{A_{NU} + B_{NU}}{\alpha_N + \beta_U} + \frac{B_{FU}}{\beta_U} \\
B_{TN} + B_{RN} &= \frac{A_{EN} + B_{EN}}{\alpha_E + \beta_N} + \frac{A_{UN} + B_{UN}}{\alpha_U + \beta_N} + \frac{A_{NN} + B_{NN}}{\alpha_N + \beta_N} + \frac{B_{FN}}{\beta_N}
\end{aligned} \tag{3.26}$$

These could be solved by non-linear least squares for example.

3.6 Model Testing and Initial Starting Values

Under both models for the partial gross flows tables, we can show that

$$\begin{aligned}
E\left\{\frac{\partial F_3}{\partial T_{EE}}\right\} &= 2-\lambda_1-\lambda_4 \qquad (3.27)\\
E\left\{\frac{\partial F_3}{\partial T_{EU}}\right\} &= 2-\lambda_1-\lambda_5\\
E\left\{\frac{\partial F_3}{\partial T_{EN}}\right\} &= 2-\lambda_1-\lambda_6\\
E\left\{\frac{\partial F_3}{\partial T_{UE}}\right\} &= 2-\lambda_2-\lambda_4\\
E\left\{\frac{\partial F_3}{\partial T_{UU}}\right\} &= 2-\lambda_2-\lambda_5\\
E\left\{\frac{\partial F_3}{\partial T_{UN}}\right\} &= 2-\lambda_2-\lambda_6\\
E\left\{\frac{\partial F_3}{\partial T_{NE}}\right\} &= 2-\lambda_3-\lambda_4\\
E\left\{\frac{\partial F_3}{\partial T_{NU}}\right\} &= 2-\lambda_3-\lambda_5\\
E\left\{\frac{\partial F_3}{\partial T_{NN}}\right\} &= 2-\lambda_3-\lambda_6\\
E\left\{\frac{\partial F_3}{\partial T_{EF}}\right\} &= 1-\lambda_1\\
E\left\{\frac{\partial F_3}{\partial T_{UF}}\right\} &= 1-\lambda_2\\
E\left\{\frac{\partial F_3}{\partial T_{NF}}\right\} &= 1-\lambda_3\\
E\left\{\frac{\partial F_3}{\partial T_{FE}}\right\} &= 1-\lambda_4\\
E\left\{\frac{\partial F_3}{\partial T_{UE}}\right\} &= 1-\lambda_5\\
E\left\{\frac{\partial F_3}{\partial T_{FN}}\right\} &= 1-\lambda_6
\end{aligned}$$

Therefore a test of model adequacy would test the null hypothesis

$$H_0 : \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = \lambda_6 = 1 \qquad (3.28)$$

We will discuss this further in the Results section. Reasonable starting values for the nonlinear maximization of F_3 is to set

$$\hat{\lambda}_1^{(0)} = \hat{\lambda}_2^{(0)} = \hat{\lambda}_3^{(0)} = \hat{\lambda}_4^{(0)} = \hat{\lambda}_5^{(0)} = \hat{\lambda}_6^{(0)} = 1 \qquad (3.29)$$

which yields

$$\begin{aligned}
\hat{T}^{(0)}_{EE} &= \frac{A_{EE}+B_{EE}}{(1-\hat{\theta}_{AE48})(1-\hat{P}_{AE})+(1-\hat{\theta}_{BE15})(1-\hat{P}_{BE})} \\
\hat{T}^{(0)}_{EU} &= \frac{A_{EU}+B_{EU}}{(1-\hat{\theta}_{AE48})(1-\hat{P}_{AE})+(1-\hat{\theta}_{BU15})(1-\hat{P}_{BU})} \\
\hat{T}^{(0)}_{EN} &= \frac{A_{EN}+B_{EN}}{(1-\hat{\theta}_{AE48})(1-\hat{P}_{AE})+(1-\hat{\theta}_{BN15})(1-\hat{P}_{BN})} \\
\hat{T}^{(0)}_{UE} &= \frac{A_{UE}+B_{UE}}{(1-\hat{\theta}_{AU48})(1-\hat{P}_{AU})+(1-\hat{\theta}_{BE15})(1-\hat{P}_{BE})} \\
\hat{T}^{(0)}_{UU} &= \frac{A_{UU}+B_{UU}}{(1-\hat{\theta}_{AU48})(1-\hat{P}_{AU})+(1-\hat{\theta}_{BU15})(1-\hat{P}_{BU})} \\
\hat{T}^{(0)}_{UN} &= \frac{A_{UN}+B_{UN}}{(1-\hat{\theta}_{AU48})(1-\hat{P}_{AU})+(1-\hat{\theta}_{BN15})(1-\hat{P}_{BN})} \\
\hat{T}^{(0)}_{NE} &= \frac{A_{NE}+B_{NE}}{(1-\hat{\theta}_{AN48})(1-\hat{P}_{AN})+(1-\hat{\theta}_{BE15})(1-\hat{P}_{BE})} \\
\hat{T}^{(0)}_{NU} &= \frac{A_{NU}+B_{NU}}{(1-\hat{\theta}_{AN48})(1-\hat{P}_{AN})+(1-\hat{\theta}_{BU15})(1-\hat{P}_{BU})} \\
\hat{T}^{(0)}_{NN} &= \frac{A_{NN}+B_{NN}}{(1-\hat{\theta}_{AN48})(1-\hat{P}_{AN})+(1-\hat{\theta}_{BN15})(1-\hat{P}_{BN})} \\
\hat{T}^{(0)}_{EF} &= \frac{A_{EF}}{(1-\hat{\theta}_{AE48})(1-\hat{P}_{AE})} \\
\hat{T}^{(0)}_{UF} &= \frac{A_{UF}}{(1-\hat{\theta}_{AU48})(1-\hat{P}_{AU})} \\
\hat{T}^{(0)}_{NF} &= \frac{A_{NF}}{(1-\hat{\theta}_{AN48})(1-\hat{P}_{AN})} \\
\hat{T}^{(0)}_{FE} &= \frac{B_{FE}}{(1-\hat{\theta}_{BE15})(1-\hat{P}_{BE})} \\
\hat{T}^{(0)}_{FU} &= \frac{B_{FU}}{(1-\hat{\theta}_{BU15})(1-\hat{P}_{BU})} \\
\hat{T}^{(0)}_{FN} &= \frac{B_{FN}}{(1-\hat{\theta}_{BN15})(1-\hat{P}_{BN})}
\end{aligned} \tag{3.30}$$

For the alternative estimation method described in 3.27 we can use these starting values.

$$\begin{aligned}
\hat{\alpha}^{(0)}_{E} &= (1-\hat{\theta}_{AE48})(1-\hat{P}_{AE}) \\
\hat{\alpha}^{(0)}_{U} &= (1-\hat{\theta}_{AU48})(1-\hat{P}_{AU}) \\
\hat{\alpha}^{(0)}_{N} &= (1-\hat{\theta}_{AN48})(1-\hat{P}_{AN}) \\
\hat{\beta}^{(0)}_{E} &= (1-\hat{\theta}_{BE15})(1-\hat{P}_{BE}) \\
\hat{\beta}^{(0)}_{U} &= (1-\hat{\theta}_{BU15})(1-\hat{P}_{BU}) \\
\hat{\beta}^{(0)}_{N} &= (1-\hat{\theta}_{BN15})(1-\hat{P}_{BN})
\end{aligned} \tag{3.31}$$

3.7 Construction of Gross Flows Weights

Recall that

$$A_{EE}+B_{EE} = \sum_{i\in S_t\cup S_{t-1}} (w_{t-1,i}+w_{t,i})D_{EE,i} \tag{3.32}$$

Now we define the gross weight for the EE transition as

$$w_{EE,i} = (\hat{\alpha}_E + \hat{\beta}_E)^{-1}(w_{t-1,i} + w_{t,i})D_{EE,i} \text{ for } i \in S_t \cup S_{t-1} \tag{3.33}$$

which implies

$$\hat{T}_{EE} = \sum_{i \in S_t \cup S_{t-1}} w_{EE,i} \tag{3.34}$$

Similarly the entire set of gross flows weights are defined as

$$\begin{aligned}
w_{EE,i} &= (\hat{\alpha}_E + \hat{\beta}_E)^{-1}(w_{t-1,i} + w_{t,i})D_{EE,i} \text{ for } i \in S_t \cup S_{t-1} \\
w_{EU,i} &= (\hat{\alpha}_E + \hat{\beta}_U)^{-1}(w_{t-1,i} + w_{t,i})D_{EU,i} \text{ for } i \in S_t \cup S_{t-1} \\
w_{EN,i} &= (\hat{\alpha}_E + \hat{\beta}_N)^{-1}(w_{t-1,i} + w_{t,i})D_{EN,i} \text{ for } i \in S_t \cup S_{t-1} \\
w_{UE,i} &= (\hat{\alpha}_U + \hat{\beta}_E)^{-1}(w_{t-1,i} + w_{t,i})D_{UE,i} \text{ for } i \in S_t \cup S_{t-1} \\
w_{UU,i} &= (\hat{\alpha}_U + \hat{\beta}_U)^{-1}(w_{t-1,i} + w_{t,i})D_{UU,i} \text{ for } i \in S_t \cup S_{t-1} \\
w_{UN,i} &= (\hat{\alpha}_U + \hat{\beta}_N)^{-1}(w_{t-1,i} + w_{t,i})D_{UN,i} \text{ for } i \in S_t \cup S_{t-1} \\
w_{NE,i} &= (\hat{\alpha}_N + \hat{\beta}_E)^{-1}(w_{t-1,i} + w_{t,i})D_{NE,i} \text{ for } i \in S_t \cup S_{t-1} \\
w_{NU,i} &= (\hat{\alpha}_N + \hat{\beta}_U)^{-1}(w_{t-1,i} + w_{t,i})D_{NU,i} \text{ for } i \in S_t \cup S_{t-1} \\
w_{NN,i} &= (\hat{\alpha}_N + \hat{\beta}_N)^{-1}(w_{t-1,i} + w_{t,i})D_{NN,i} \text{ for } i \in S_t \cup S_{t-1} \\
w_{EF,i} &= (\hat{\alpha}_E)^{-1}(w_{t-1,i})D_{EF,i} \text{ for } i \in S_t \cup S_{t-1} \\
w_{UF,i} &= (\hat{\alpha}_U)^{-1}(w_{t-1,i})D_{UF,i} \text{ for } i \in S_t \cup S_{t-1} \\
w_{NF,i} &= (\hat{\alpha}_N)^{-1}(w_{t-1,i})D_{NF,i} \text{ for } i \in S_t \cup S_{t-1} \\
w_{FE,i} &= (\hat{\beta}_E)^{-1}(w_{t,i})D_{FE,i} \text{ for } i \in S_t \cup S_{t-1} \\
w_{FU,i} &= (\hat{\beta}_U)^{-1}(w_{t,i})D_{FU,i} \text{ for } i \in S_t \cup S_{t-1} \\
w_{FN,i} &= (\hat{\beta}_N)^{-1}(w_{t,i})D_{FN,i} \text{ for } i \in S_t \cup S_{t-1}
\end{aligned} \tag{3.35}$$

4. Results

In this section we present our empirical results. We examined Monthly data from December 2002 to December 2023.

4.1 Population Flows

In this section we examine the Gross flows Tables estimates. We first present a panel graph in Figure 1. This shows graphs of each of the cells in the gross flows table for the total CNP. We see that our estimates match the official estimates very closely except for the outflow column and inflow row.

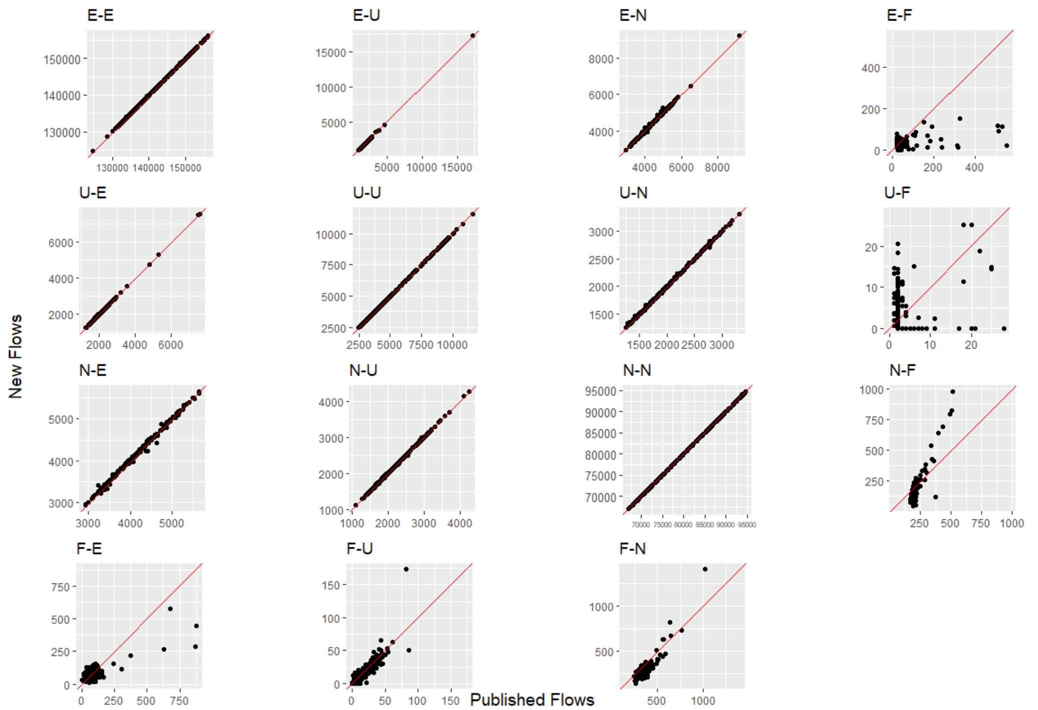

Figure 1: New Flows versus Official Flows for 2003 to 2023 for Total CNP.

We can examine this more carefully in Table 4 by computing the percent relative absolute differences between our estimates and the official estimates. We do this over the period January 2003 thru January 2023 broken out by Sex and for Total CNP. Again we see that our estimates match the official estimates very closely except for the outflows and the inflows. There appears to be greater difference in the outflows.

Finally in Table 5 we compute the average coefficient of variation for the year 2023 by Sex and Total CNP for each of the entries of the gross flows table. We do not have comparable estimates to present for the official estimates.

Table 4: % Relative Absolute Difference Compared to Official Estimates for the Monthly Gross Flows Table Estimates (averaged over 2003-2023).

Men						
		E	U	N	F	T
time period $t-1$	E	0.02%	0.49%	0.73%	62.84%	0.00%
	U	0.39%	0.13%	0.54%	142.07%	0.01%
	N	0.75%	0.49%	0.07%	34.17%	0.00%
	F	42.52%	30.20%	15.31%	–	13.67%
	T	0.00%	0.01%	0.00%	25.12%	–
Women						
		E	U	N	F	T
time period $t-1$	E	0.02%	0.47%	0.56%	68.26%	0.00%
	U	0.39%	0.16%	0.39%	139.95%	0.01%
	N	0.58%	0..40%	0.06%	38.50%	0.00%
	F	41.45%	31.07%	19.17%	–	18.07%
	T	0.00%	0.01%	0.00%	31.31%	–
Total						
		E	U	N	F	T
time period $t-1$	E	0.02%	0.48%	0.57%	58.23%	0.00%
	U	0.36%	0.10%	0.41%	119.59%	0.01%
	N	0.61%	0.43%	0.06%	35.01%	0.00%
	F	41.82%	23.38%	16.78%	–	15.70%
	T	0.00%	0.00%	0.00%	27.98%	–

Table 5: Average %CV for 2023 of Monthly Gross Flows Table Estimates.

Men						
		E	U	N	F	T
time period $t-1$	E	0.34%	9.53%	5.99%	68.19%	0.32%
	U	8.89%	5.66%	9.64%	40.36%	3.92%
	N	6.53%	9.55%	0.65%	26.14%	0.64%
	F	28.59%	77.41%	15.83%	–	10.77%
	T	0.32%	3.90%	0.64%	20.41%	–
Women						
		E	U	N	F	T
time period $t-1$	E	0.42%	10.44%	5.89%	74.33%	0.40%
	U	9.40%	6.59%	9.71%	28.28%	4.18%
	N	6.35%	9.51%	0.51%	29.46%	0.51%
	F	29.75%	67.26%	16.08%	–	11.48%
	T	0.41%	4.16%	0.52%	23.91%	–
Total						
		E	U	N	F	T
time period $t-1$	E	0.28%	7.24%	4.41%	46.60%	0.27%
	U	6.56%	4.47%	6.77%	53.37%	2.99%
	N	4.86%	6.70%	0.42%	19.71%	0.42%
	F	23.00%	65.94%	11.68%	–	8.45%
	T	0.27%	2.97%	0.43%	16.48%	–

4.2 Month-in-Sample Effects

In this section we examine the month-in-sample effects. Since month-in-sample indexes are a common way to examine month-in-sample effects we chose to look at them here. They are the same as the θ defined in 3.16 except they are multiplied by 8, which means that a value of 1.0 indicates the MIS effect is no different that the average of all 8 MIS groups. We present the MIS indexes for annual averages for the years 2003 thru 2023 in order to reduce the variability of the estimates. In Figure 2 we plot the MIS indexes for total employment. We see that the indexes are relatively stable over the 20 years except for the Covid-19 pandemic period. The pandemic effect is also apparent in Figure 4 for not-in-labor force, while less apparent but still there for Unemployed in Figure 3. Similar impacts have previously been discussed in McIllece (2020).

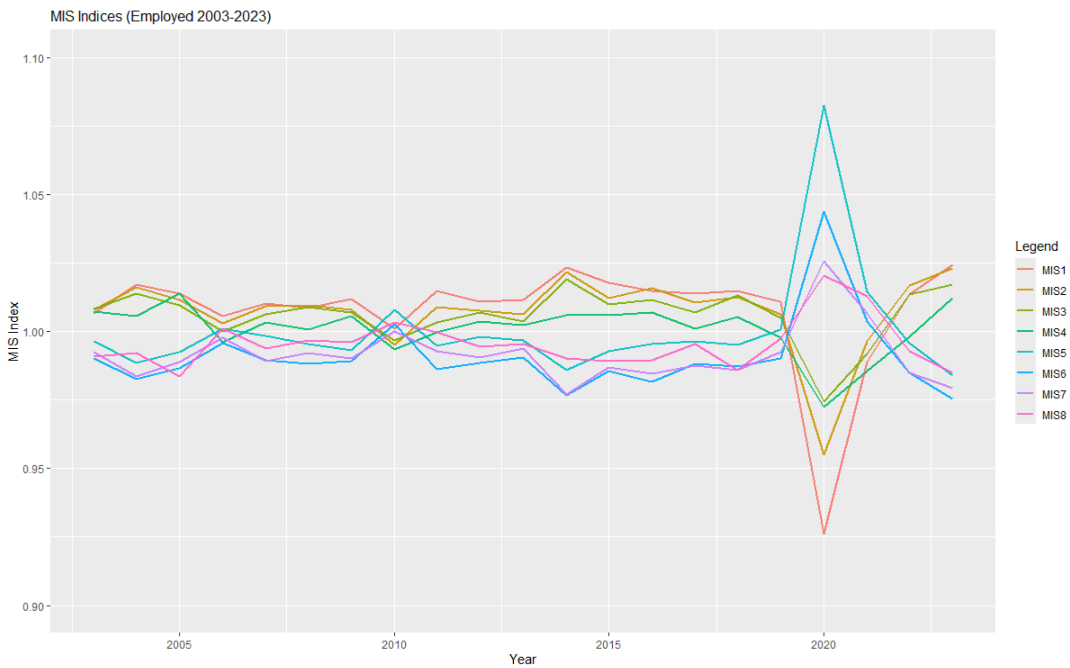

Figure 2: Annual Average Month-in-Sample Indexes for Total Employed 2003-2023.

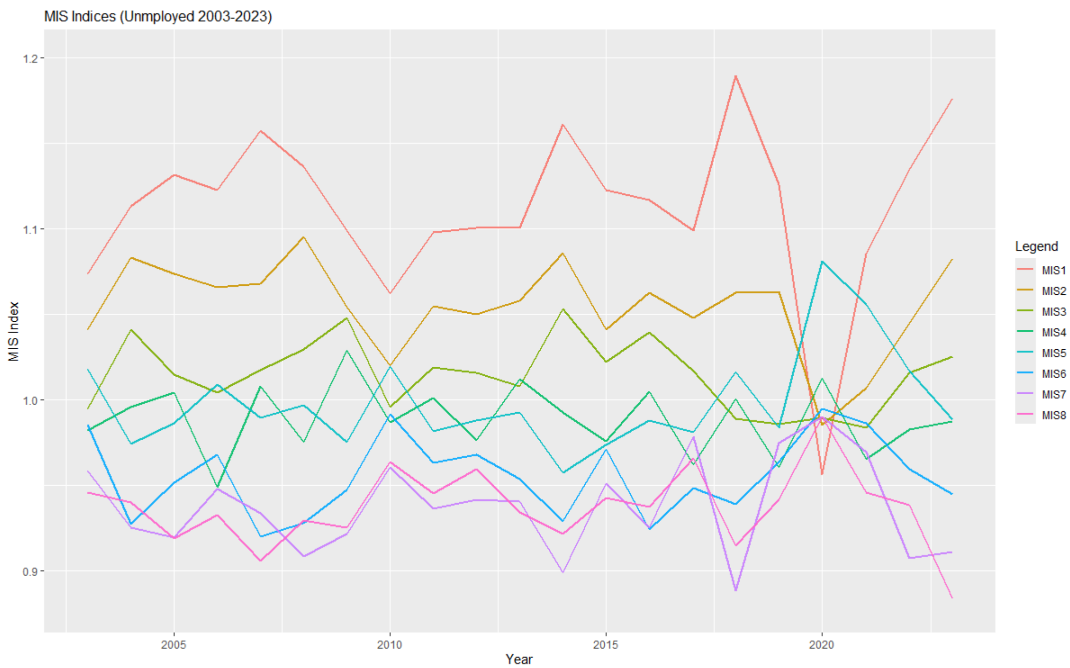

Figure 3: Annual Average Month-in-Sample Indexes for Total Unemployed 2003-2023.

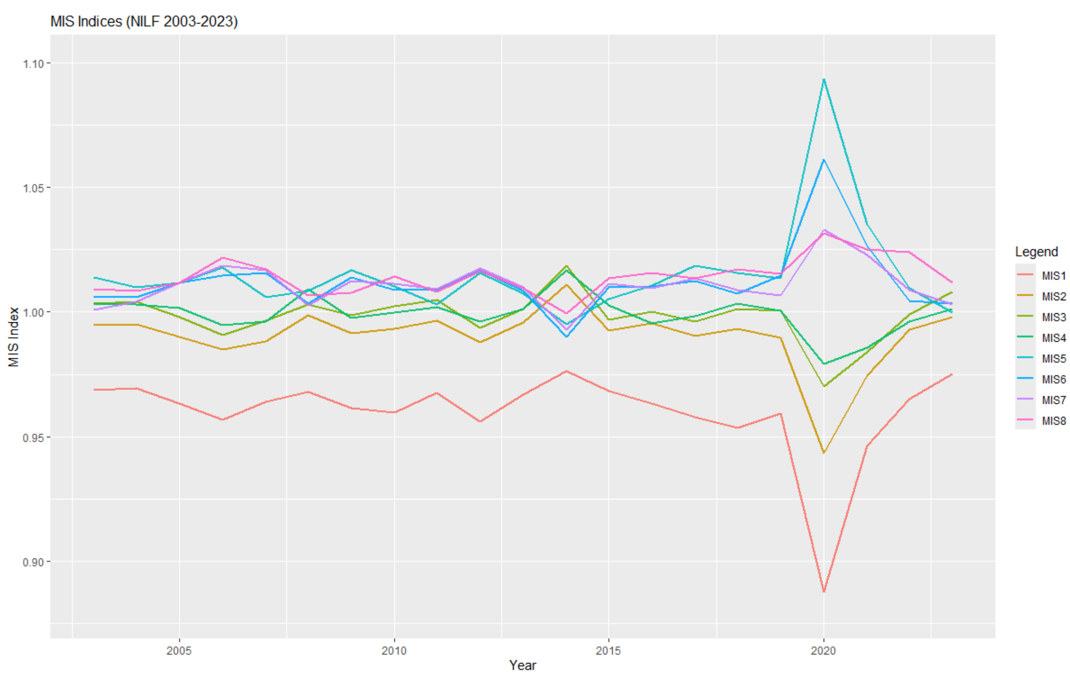

Figure 4: Annual Average Month-in-Sample Indexes for Total Not-in-Labor Force 2003-2023.

4.3 Missing Rates

In this section we examine the missing rates. We present estimates of the missing rate as defined in equation 3.19. We plot the missing rates for each of the six month-in-sample pairs and the graphs each are 12-month moving averages of the missing rates to smooth them out and reduce the variability. We begin by examining Figure 5. The monthly missing rates are plotted along with a solid red line indicating the 12-month moving average. If we look first at MIS 1-2 we see a steady increasing in the missing rate over time, which mimics the type-A non response rate increase we have seen in CPS overall. If we examine MIS 2-3 we see an decreasing response rate but overall the missing rate is lower than for MIS 1-2. To help see that there is a faint blue line also included in the MIS 2-3 graph which is the 12-month moving average for MIS 1-2, so you can see the red line for MIS 2-3 is shifted down relative to that of MIS 1-2. Likewise for the MIS 3-4 pair the faint blue line is the 12-month moving average for MIS 2-3 while the red is the 12-month moving average for MIS 3-4. Again we see a decrease in the overall missing rate as you go from MIS 2-3 to MIS 3-4. We treat MIS 5-6, MIS 6-7 and MIS 7-8 similarly. MIS 5-6 is presented as MIS 1-2 was (by itself) and then MIS 6-7 is shown relative to MIS 5-6, and MIS 7-8 is shown relative to MIS 6-7. Again we see an overall decrease in the missing rates once housing units have reentered the sample after an 8 month absence. We see similar patterns in the reaming 5 sets of panels graphs in Figures 6 thru Figure 10. Overall flows to missing greatest for Unemployed while the missingness for Employed is greater than for Non-in-Labor Force.

Figure 5: Missing Rates for Table *A* Employed.

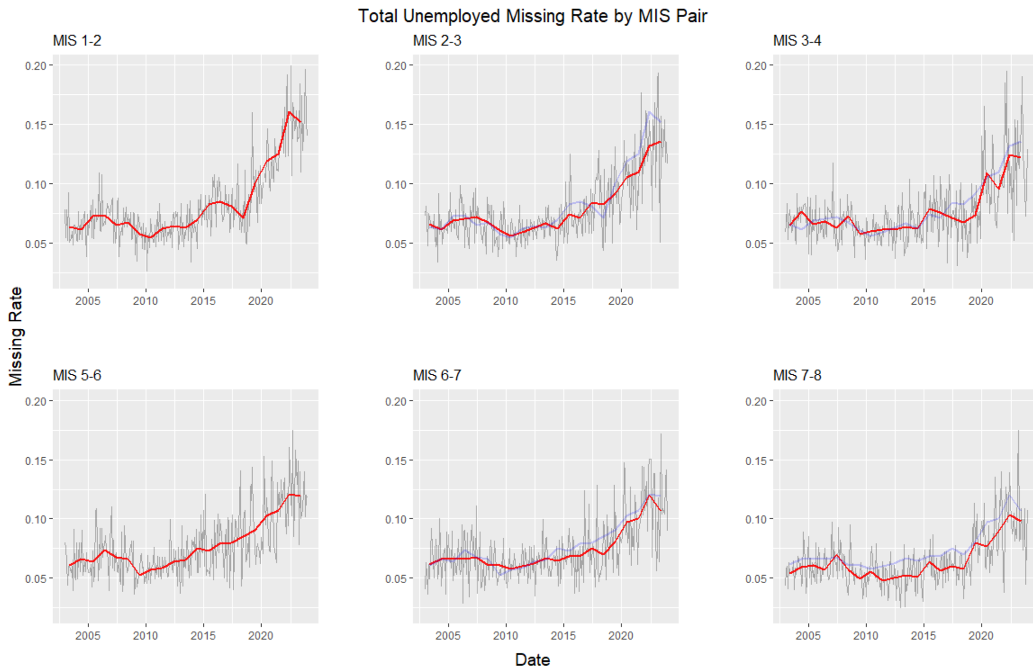

Figure 6: Missing Rates for Table *A* Unemployed.

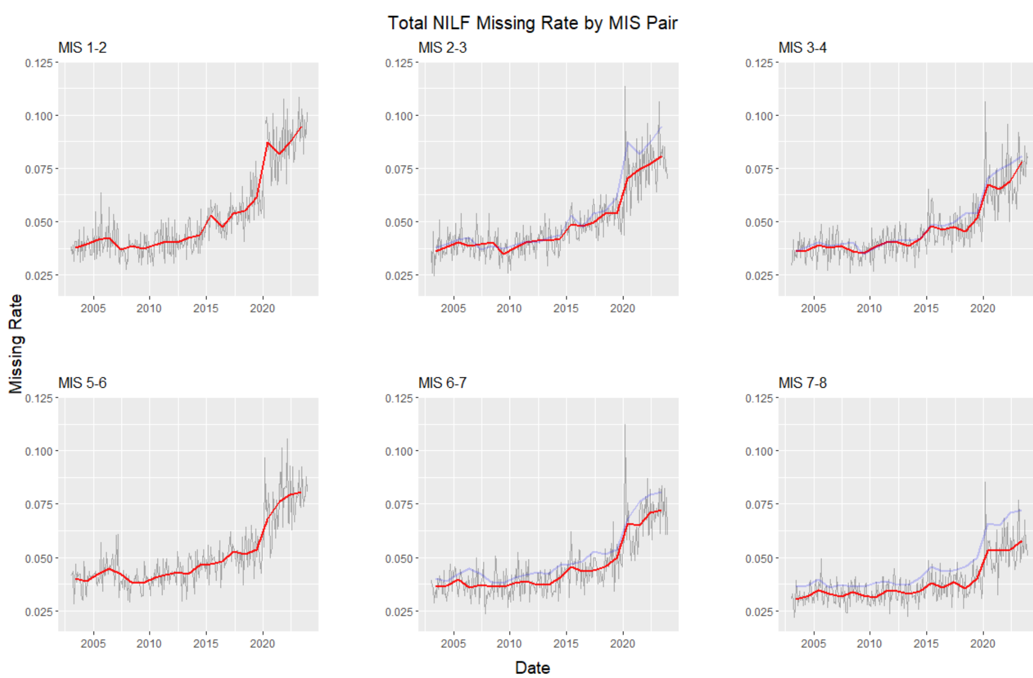

Figure 7: Missing Rates for Table *A* Not-in-Labor-Force.

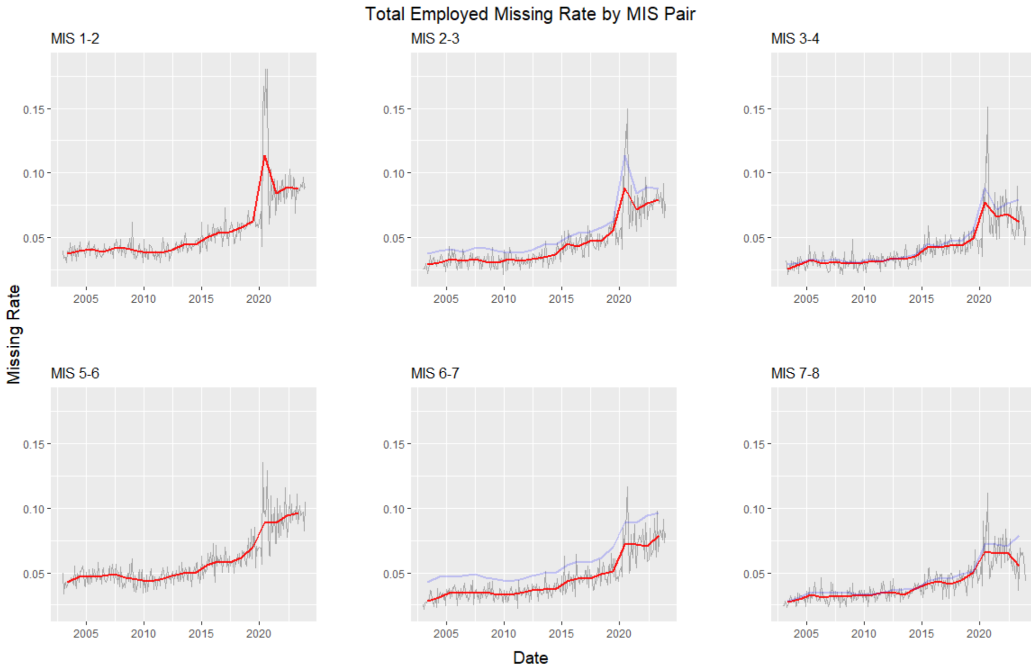

Figure 8: Missing Rates for Table B Employed.

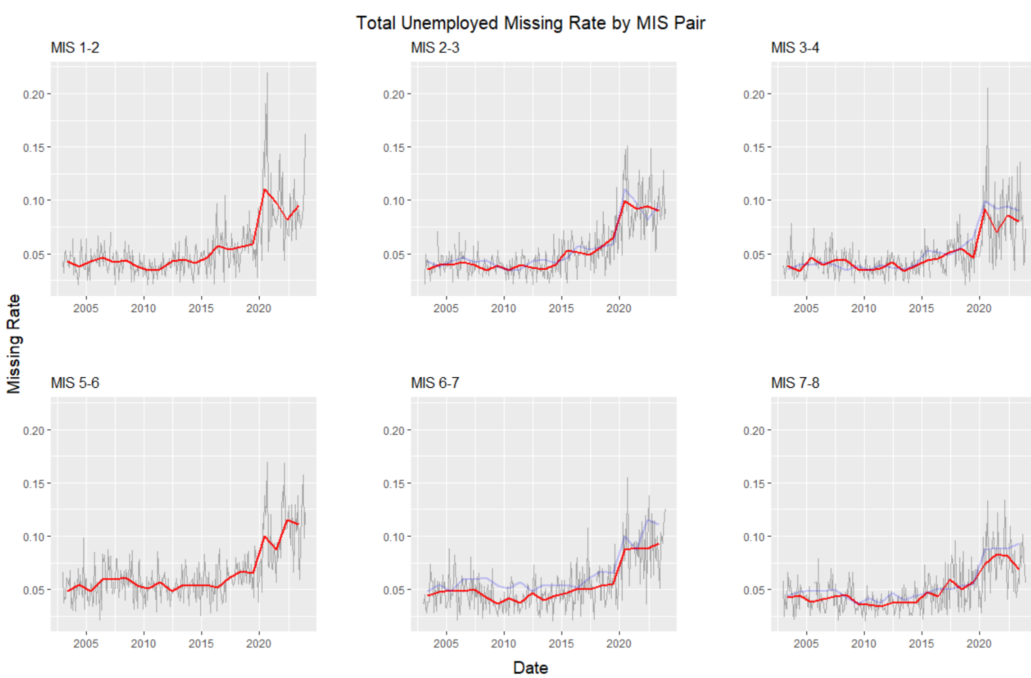

Figure 9: Missing Rates for Table B Unemployed.

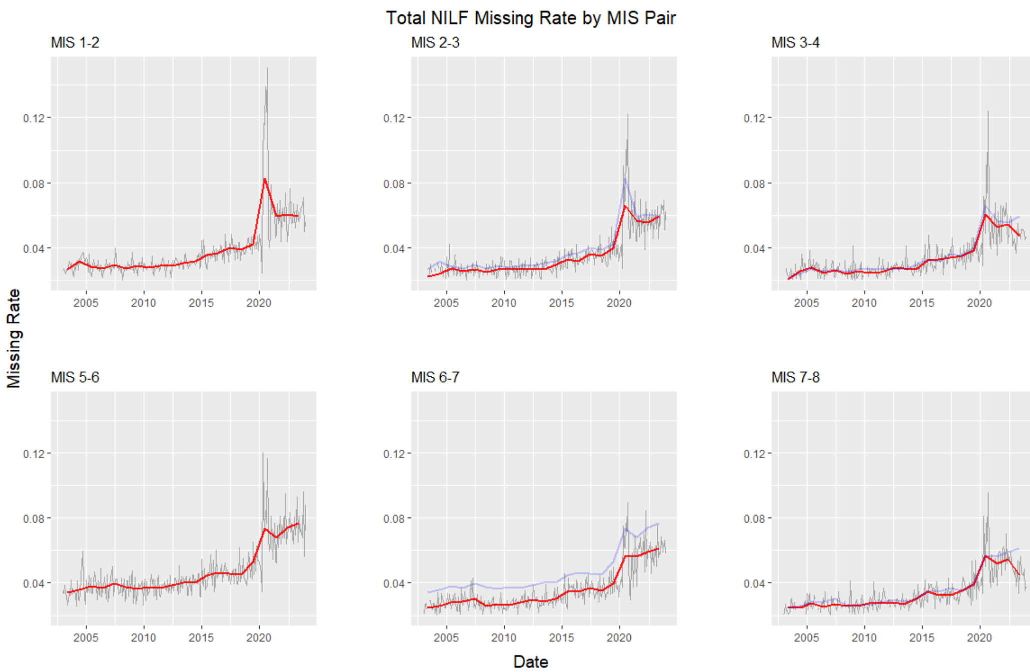

Figure 10: Missing Rates for Table *B* Not-n-Labor-Force.

4.4 Lagrange Multipliers

In this section we examine the Lagrange multipliers discussed in equation 3.28. In Figure 11 we present six panel graphs for the six lambda values which we label $L_1, L_2, L_3, L_4, L_5, L_6$. These panels are different than the style of the ones presented previously. Here we plot estimates for 2003 through 2023 but since the estimates appear to have seasonal behavior we plot them as schematic plots for each of the twelve months. In order to better understand the plots, it is important to recall that L_1 corresponds to the constraint for total employed for the previous month, L_2 corresponds to the constraint for total unemployed for the previous month and L_3 corresponds to the constraint for total not-in-labor force for the previous month. Similarly, L_4 corresponds to the constraint for total employed for the current month, L_5 corresponds to the constraint for total unemployed for the current month and L_6 corresponds to the constraint for total not-in-labor force for the current month. Another way to think about it is that L_1 and L_4 pair off as constraints for employed, L_2 and L_5 pair off as constraints for unemployed while L_3 and L_6 pair off as constraints for not-in-labor force. With that in mind we can make several observations. Note that January has the widest dispersion of any month, and that corresponds to the annual population control adjustment CPS undergoes each year. Also note that the labor force pairs (L_1, L_4), (L_2, L_5), (L_3, L_6) tend to balance each other out around one. For example L_1 tends to be below 1 while L_4 tends to be above 1 on average. We did not perform a formal test as was suggested in 3.28, but we can see that Lagrange multipliers do balance out around one between pairs. One interesting phenomena occurs in July in which the lambdas for all three labor force categories in the previous time period are significantly below 1 on average, while the lambdas are all significantly above 1 in the current time period. It is not readily apparent what is causing this but it could reflect an annual revision of population controls around that time of year.

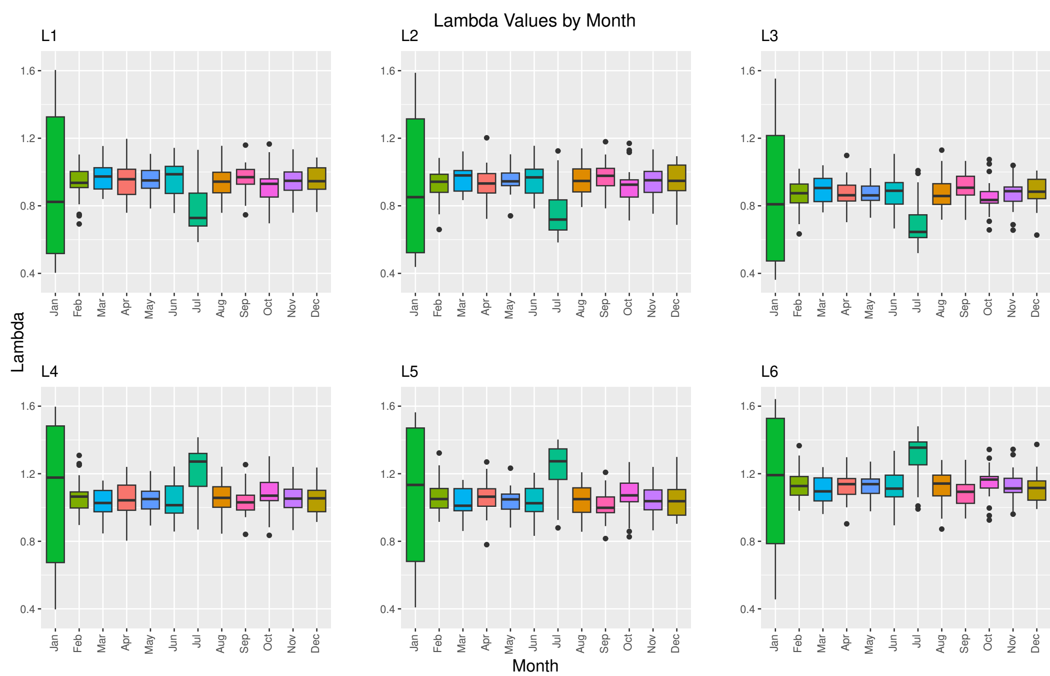

Figure 11: Lagrange Multipliers for Total.

5. Conclusion

We have examined an alternative gross flows estimation method and compared it to the official method used by the U.S. Bureau of Labor Statistics. We find that our method produces estimates which are very close to the main entries of the table of official estimates but differ for the outflow column and inflow row. The differences in the inflows and outflows reflect in part the fact that we do not use any external information to model them, while the official estimates do use external data. Future research could explore incorporating some external data either directly or in bounding the estimates.

An interesting result of our research is the development of missing rates by month-in-sample (MIS) groups. Our results indicate in general that the labor category with the largest missing rate is unemployed, followed by employed, while not-in-labor force tends to have the smallest missing rate. In addition, the MIS pair $(1, 2)$ has the largest missing rates for all labor force categories, and the missing rates diminish monotonically for the pairs $(2, 3)$ and $(3, 4)$. Upon reentry into the CPS, the pairs $(5, 6)$, $(6, 7)$, $(7, 8)$ also exhibit a general monotonic decreasing missing rate.

In future research we intend to use our method to expand the rows of the gross flows table to break out Employed into Full-time and Part-time and to break out Not-in-labor Force into marginally attached and the remainder. It is hoped that such an expansion of the table could yield further insights into the dynamics of the labor force.

REFERENCES

Design and Methodology, Current Population Survey, Technical Paper 77 (2019). U.S. Census Bureau.

Harley Frazis, Edwin Robison, Thomas Evans and Martha Duff, “Estimating gross flows consistent with stocks in the CPS”, Monthly Labor Review September 2005

Randy Ilg, “Analyzing CPS data using ”, Monthly Labor Review September 2005

Stasny, E.A., Fienberg, S.E. (1985), ”Some Stochastic Models for Estimating Gross Flows in the Presence of Non-Random Nonresponse, in em Proceedings on the Conference on Gross Flows in Labor Force Statistics, Washington DC: U.S. Department of Commerce and U.S. Department of Labor, pp. 25-39.

Stasny, Elizabeth A., (1988), ”Modeling Nonignorable Nonresponse in Categorical Panel Data with an Example in Estimating Gross Labor-Force Flows”, *Journal of Business & Economic Statistics* Vol. 6 No. 2, pp. 207-219.

McIllece, Justin (2020) ”Covid-19 and the Current Population Survey Response Rates and Estimation Bias” Proceedings the American Statistical Association Joint Statistical Meetings.