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

Connor Doherty [1] Stephen M. Miller [2]

[1] U.S. Bureau of Labor Statistics (OEUS Statistical Methods Staff)

[2]U.S. Bureau of Labor Statistics (Office of Survey Methods Research)

Joint Statistical Meetings
August, 2024

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

Outline

Monthly Labor Force Estimates from BLS

- The Bureau of Labor Statistics (BLS) publishes monthly estimates for individuals aged 16+ in the Civilian Non institutional Population (CNP)
  - Total for those Employed
  - Total for those Unemployed
  - Total for those Not in the Labor Force
- Month-to-month transitions between those labor force states which are called ”Gross Flows”
- There are some individuals who transition in or out of the CNP each month which we call Flows (either inflows or outflows)

Monthly Gross Flows Labor Force Estimation

Table: Typical Monthly Table (in thousands).

		April 2024				
		E	U	N	F	T
March 2024	E	155,916	1,361	4,053	26	161,356
	U	1,793	3,176	1,633	2	6,604
	N	3,760	1,347	94,609	209	99,924
	F	121	10	288	–	418
	T	161,590	5,894	100,582	237	–

E = Employed
U = Unemployed
N = Not-in-Labor-Force
F = Flow (In-flow or Out-flow)
T = Total

Theoretical Monthly Gross Flows Table

Table: Population Gross Flows Table.

		Current Month				
		E	U	N	F	T
Previous Month	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}	$-$	T_{FT}
	T	T_{TE}	T_{TU}	T_{TN}	T_{TF}	$-$

E = Employed
U = Unemployed
N = Not-in-Labor-Force
F = Flow (In-flow or Out-flow)
T = Total

Motivation for Our Research

- The current raking method of Gross Flows Estimation used by BLS was described in Frazis, Robison, Evans and Duff (2005)
- The purpose of our research is to evaluate an alternative estimation method which builds off a model proposed by Stasny and Fienberg (1985)
  - Does not use external death data from the NCHS
  - Explicitly models differences in Month-in-sample pairs from the Current Population Survey (CPS)

Outline

Current Population Survey

- Household survey conducted each month, with particular emphasis on the month unemployment rate
- Survey of approximately 42,000 households each month
- Basic weighting reflects: Probability of selection, various nonresponse adjustments, and Post-stratification to known population control totals each month
- 4-8-4 Rotation Design.

The CPS 4-8-4 Rotation Design

- In the CPS a household is in sample for 4 consecutive months, then exits the sample for 8 months, then reenters the sample for a final 4 consecutive months.
- MIS1, MIS2, . . ., MIS8 refers the number of months a sample household has been in Sample.
- Each month there are 8 month-in-sample groups.

The CPS 4-8-4 Rotation Design

Table: Month in Sample Pairs

	Month		
MIS	Previous	Current	Pair
1	$\searrow$	R	
2	$\searrow$	$\longrightarrow$	(1,2)
3	$\searrow$	$\longrightarrow$	(2,3)
4	R	$\longrightarrow$	(3,4)
5	$\searrow$	R	
6	$\searrow$	$\longrightarrow$	(5,6)
7	$\searrow$	$\longrightarrow$	(6,7)
8	R	$\longrightarrow$	(7,8)

- R = unpaired due to sample rotation

Outline

Definition of Rotation and Missing

- Each month some units were eligible to be sampled but were not collected. We refer to these as Missing.
- Each month some sample units are not available because they have been rotated out of the sample. We refer to this unavailability as Rotation.
- Note: both Rotation and Missing are distinct from Flows

Partial Flow Table A Using Previous Month's Weights

Table: Sample A Table

		Current Month						
		E	U	N	F	M	R	T
Previous Month	E	A_{EE}	A_{EU}	A_{EN}	A_{EF}	A_{EM}	A_{ER}	A_{ET}
	U	A_{UE}	A_{UU}	A_{UN}	A_{UF}	A_{UM}	A_{UR}	A_{UT}
	N	A_{NE}	A_{NU}	A_{NN}	A_{NF}	A_{NM}	A_{NR}	A_{NT}

E = Employed, U = Unemployed, N = Not-in-Labor Force
F = Flow (In-flow or Out-flow)
M = Missing, R = Rotation, T = Total

Partial Flow Table B Using Current Month's Weights

Table: Sample B Table

		Current Month		
		E	U	N
Previous Month	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}
	T	B_{TE}	B_{TU}	B_{TN}

E = Employed, U = Unemployed, N = Not-in-Labor Force
F = Flow (In-flow or Out-flow)
M = Missing, R = Rotation, T = Total

Schematic Representation of Our Estimation Method

Table: Reallocation of Those due to Missing and Rotation

		Current Month						
		E	U	N	F	M	R	T
Previous Month	E	$\hat{T}_{EE}$	$\hat{T}_{EU}$	$\hat{T}_{EN}$	$\hat{T}_{EF}$	$\Longleftarrow$	$\Longleftarrow$	$\hat{T}_{ET}$
	U	$\hat{T}_{UE}$	$\hat{T}_{UU}$	$\hat{T}_{UN}$	$\hat{T}_{UF}$	$\Longleftarrow$	$\Longleftarrow$	$\hat{T}_{UT}$
	N	$\hat{T}_{NE}$	$\hat{T}_{NU}$	$\hat{T}_{NN}$	$\hat{T}_{NF}$	$\Longleftarrow$	$\Longleftarrow$	$\hat{T}_{NT}$
	F	$\hat{T}_{FE}$	$\hat{T}_{FU}$	$\hat{T}_{FN}$				
	M	$\Uparrow$	$\Uparrow$	$\Uparrow$				
	R	$\Uparrow$	$\Uparrow$	$\Uparrow$				
	T	$\hat{T}_{TE}$	$\hat{T}_{TU}$	$\hat{T}_{TN}$				

E = Employed, U = Unemployed, N = Not-in-Labor Force

F = Flow (In-flow or Out-flow)

M = Missing, R = Rotation, T = Total

Example of Modeling for the A Table

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

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

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

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

Example of Modeling for the B Table

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

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

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

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

Interpretation of the Parameters

- The P parameters indicate the $\%$ missing rate
  - These are computed for E, U, N
  - Also computed for each A and B table
  - And computed for each MIS pair (1,2), (2,3), (3,4), (5,6), (6,7), (7,8)
- The θ indicate the month-in-sample effects and are computed monthly
- These parameters model the differences between MIS pairs and are a unique contribution of our research

Estimation

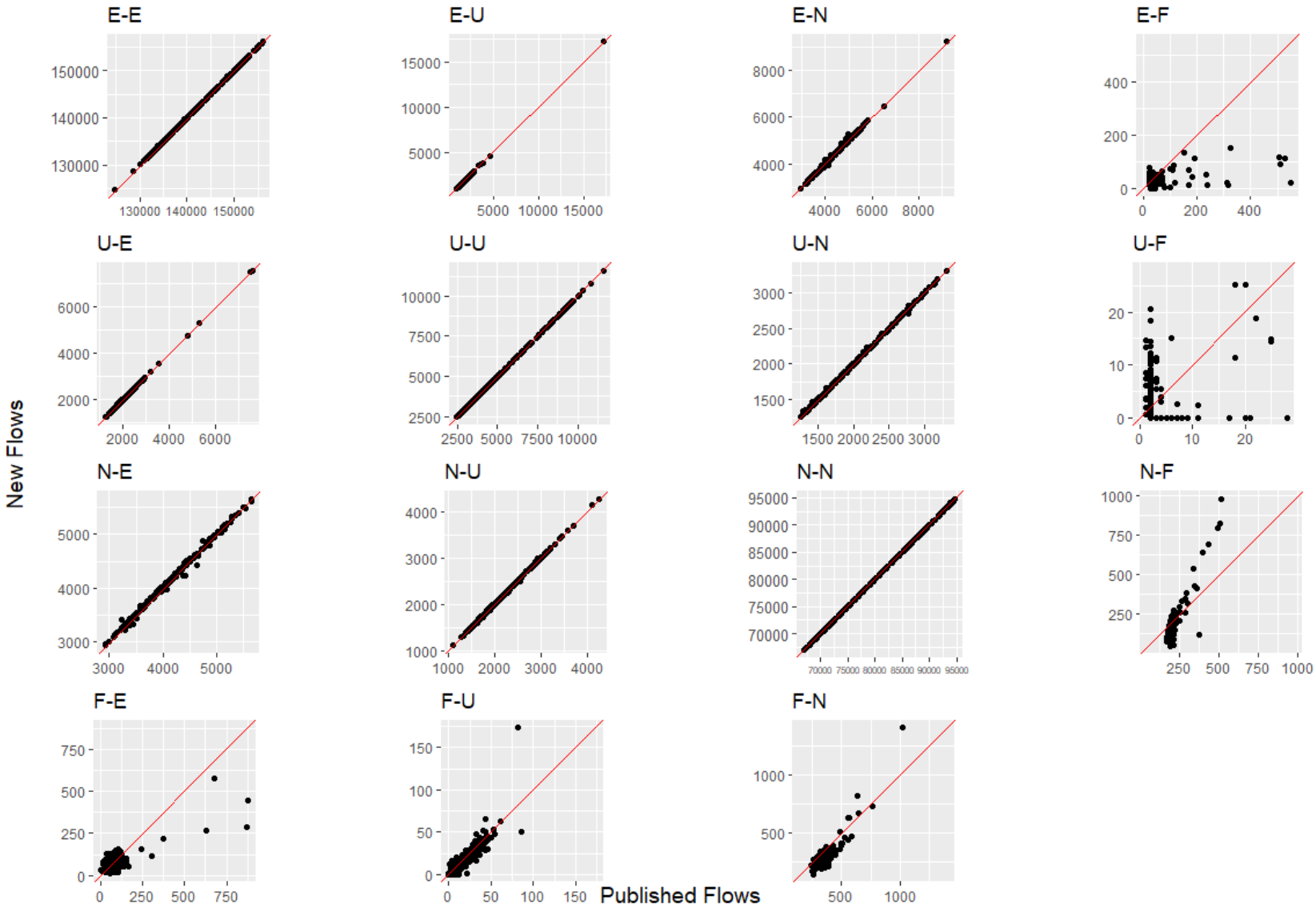
- We use a pseudo likelihood method based on a multinomial likelihood.
- We estimate the parameters by nonlinear optimization while constraining the row and column totals of the the Gross Flows Table for E, U and N to match the official monthly estimates (6 constraints). The 6 constraints correspond to 6 Lagrange multipliers which we compute each month.

Outline

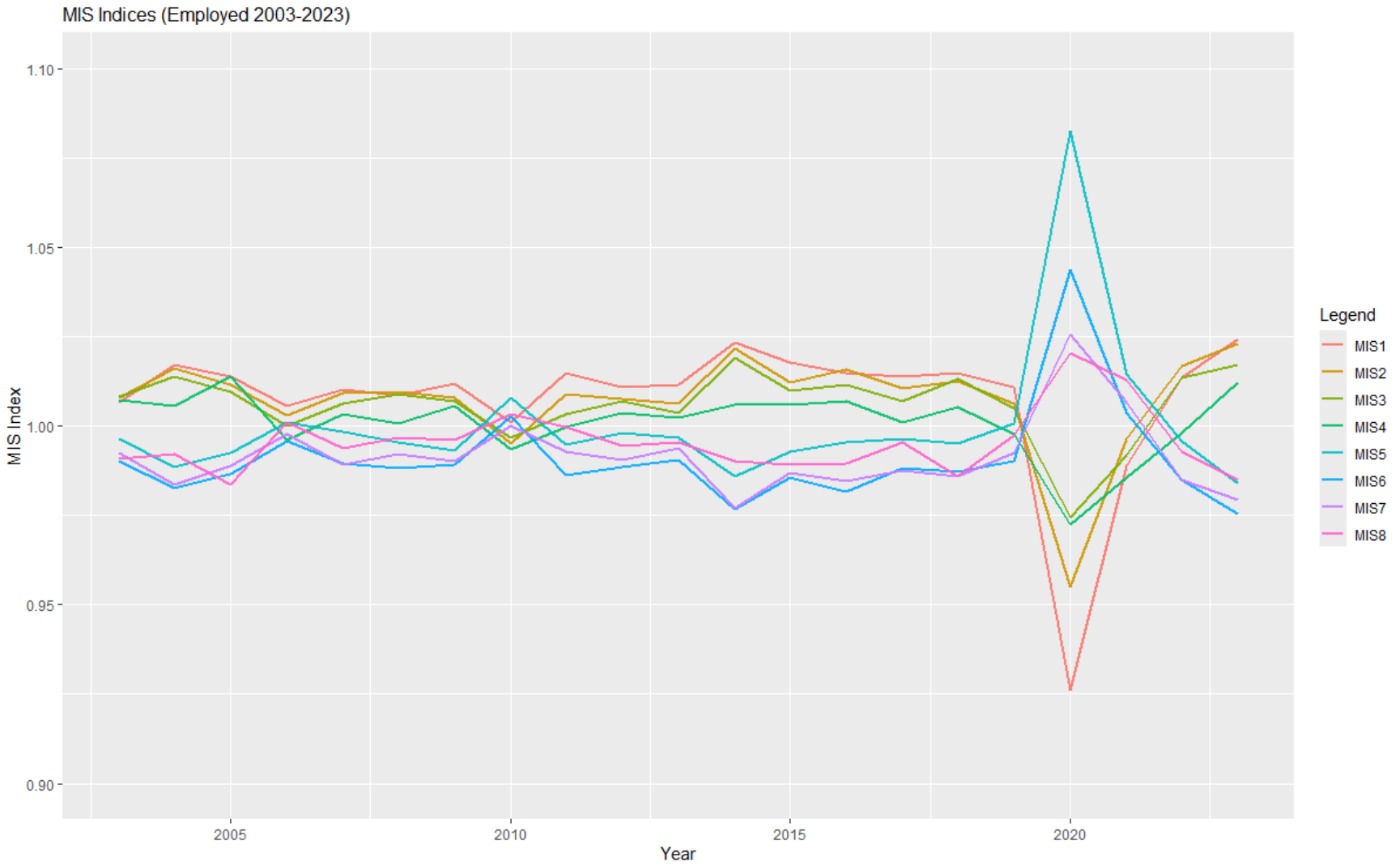
Data Analyzed

- Monthly estimates from December 2002 to December 2023 were used.
- Gross Flows Tables were calculated by Total and Total split by Sex, but only Total are presented here.

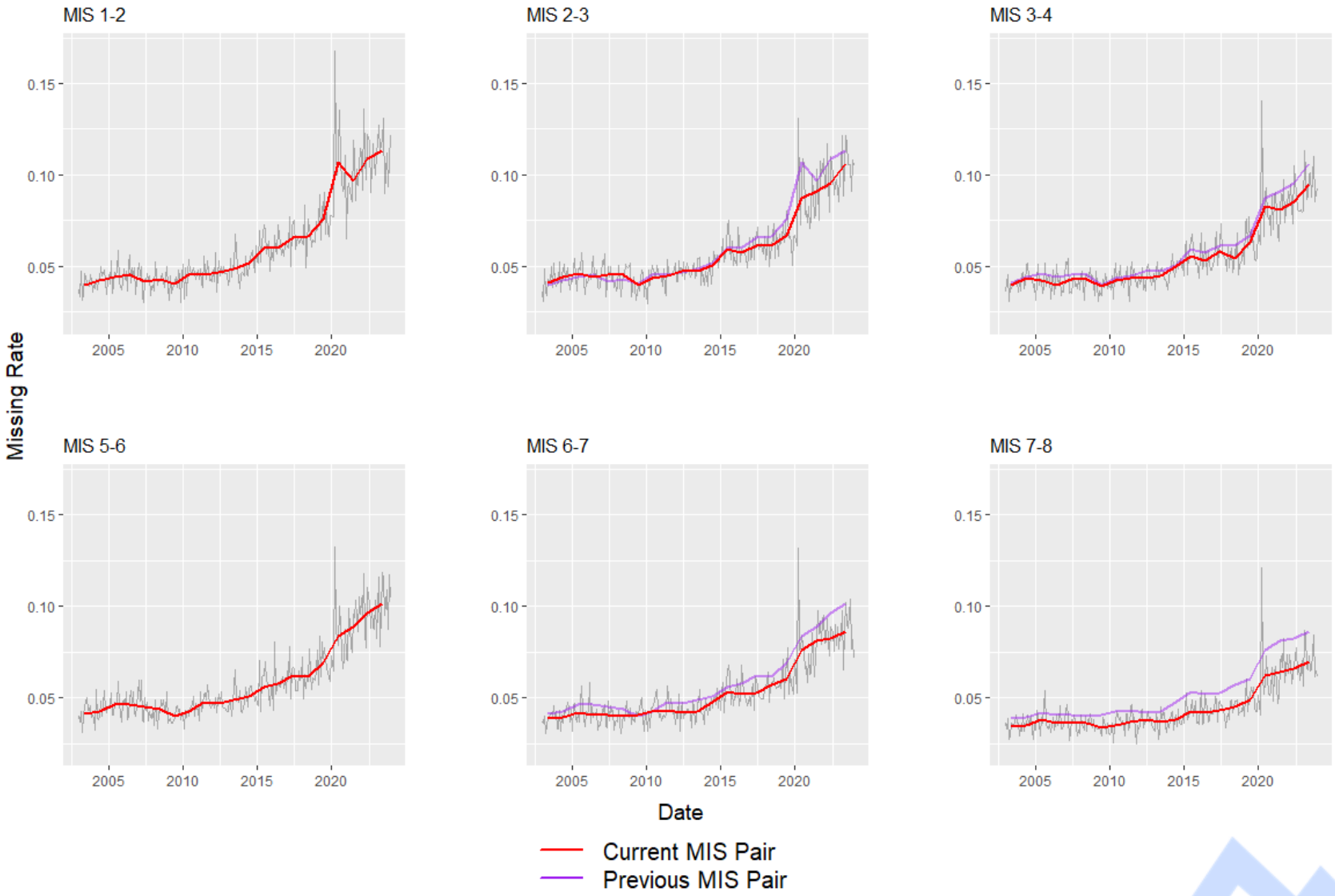
Gross Flow Estimate Comparison



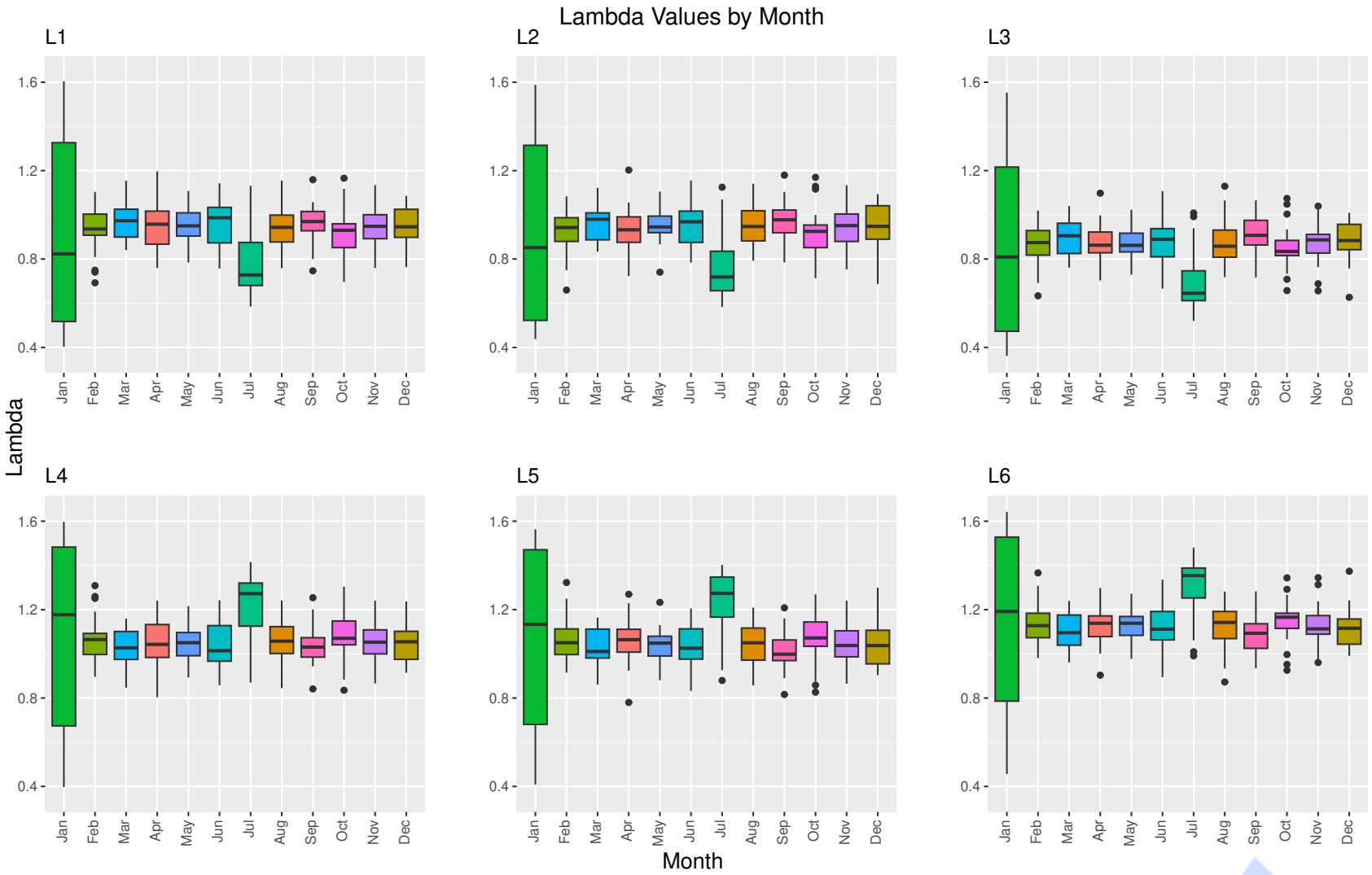
Month-in-Sample Effects (2003-2023)



Employed Missing Rates by MIS Pair (A Table)



Lagrange Constraints



Outline

- Our new estimates are very close to the reported estimates except for the in-flows and out-flows.
- Month-in-sample indexes reflect an unusual change around 2020.
- The Percent missing parameters are revealing about the dynamics of survey response.
- Lagrange multipliers reflect the impacts of annual population control changes in January,
- Future work will examine expanding the Gross Flow Table to expand Employed to include Part-time and Full-time, and expand Not-in-Labor-Force to include the Marginally Attached and the Non Marginally Attached.

CONTACT INFORMATION

Connor Doherty
doherty.connor@bls.gov
OEUS Statistical Methods Staff
Bureau of Labor Statistics

Stephen M. Miller
miller.steve@bls.gov
Office of Survey Methods Research
Bureau of Labor Statistics