A Concordance of Conway with Depken and Gaggl (DG).

(1) Preface. Both papers use the monthly samples of Current Population Survey (CPS) to determine the effect of UI policy reform on employment (E), unemployment (U) and those not in the labor force (N). We seem to get quite different results.

In this note I demonstrate that the differences come from (a) a different estimation methodology and (b) a different age profile of those included.

(2) **Methodology.**

In DG, the authors use a regression to estimate the expectation of being unemployed contingent on the policy in place (and other factors). If we use expectations notation, DG constructs $E(U_t \mid P_t)$, $E(E_t \mid P_t)$ and $E(N_t \mid P_t)$ as the expectations at time t for given UI payment policy P_t.

As DG points out, $E(U_t \mid P_t) + E(E_t \mid P_t) + E(N_t \mid P_t) = 1$ by construction.

In Conway, the regressions are contingent as well on the employment status in the previous period. For example, $E(U_t \mid X_{t-1}, P_t)$ where $X_{t-1} \in (U_{t-1}, E_{t-1}, N_{t-1})$. These contingent expectations are related to the DG expectations: for example, in the case of unemployment in period t.

$$E(U_t \mid P_t) = \xi_{Ut-1}\, E(U_t \mid U_{t-1}, P_t) + \xi_{Et-1}\, E(U_t \mid E_{t-1}, P_t) + \xi_{Nt-1}\, E(U_t \mid N_{t-1}, P_t) \qquad (1)$$

Where ξ_{Xt-1} is the share of the survey sample with labor status X_{t-1} in period t-1.

The message of Conway is that the general-equilibrium outcome for an individual choosing to go from U_{t-1} to N_t is different from that of an individual choosing to go from from E_{t-1} to N_t. For that reason, Conway estimates each contingent expectation in (1) separately rather than taking the DG approach of estimating just the left-hand side expectation over all individuals.

Both papers use a difference-in-difference framework to undertake this estimation. DG includes potential determinants of the aggregate expectations: age, race, education, industry in which the individual works. Conway focuses on the contingent expectations (like the ones on the right-hand side of (1)).

(3) **Replicating the DG results using Conway's data.**

In addition to the difference in methodology, DG make two different data choices from Conway. First, the DG sample was of individuals 16 and above in the CPS, while Conway used the prime working-age group of ages 25-54. Second, DG imposed a constant policy effect over six quarters while Conway tested policy effects quarter by quarter.

To begin, I use the Conway data to replicate the result in DG. Table A1 reports the results for the "difference-in-difference" coefficients in $E(U_t \mid P_t)$, $E(E_t \mid P_t)$ and $E(N_t \mid P_t)$ using the CPS base

data for individuals of age 16 and up for the period 2011q3 – 2015q2.[1] The larger size in the replication sample is perhaps due to the careful work of IPUMS at the University of Minnesota in cleaning the data; I used their cleaned data, while DG created theirs from the "raw" data on the NBER website (as I did in my first draft). I cannot explain the relatively large difference in NILF coefficient. I did use the CPS weights in my weighted regression, but I believe DG did as well.

Table A1: Replicating the results in DG

	Difference-in-differences coefficient			
	Unemployed	Employed	NILF	Observations
DG (Table 4)	-0.0097	0.012	-0.0026	425978
	(0.0019)	(0.0038)	(0.0036)	
R^2	0.002	0.002	0.002	
Replication	-0.0083	0.01321	-0.00488	457094
	(0.0024)	(0.0068)	(0.0066)	
R^2	0.002	0.002	0.002	

The impact of policy shift in North Carolina is to reduce unemployment, increase employment, and reduce (insignificantly) those out of the labor force. Is this consistent with Conway?

(4) **Is the sign pattern robust?**

To begin, I look at the two important differences in data choice. Do they matter to the sign pattern of these coefficients?

(a) Does the age of respondents matter? Surprisingly (at least to me), the prime-age group have the same pattern. This is reported in the second panel of Table A2. (The first panel reports once again the DG results.) While the sample size is smaller, the pattern of signs is the same.

(b) Does the length of the "treatment period" matter? The DG assumption that the differential impact of the UI policy shift in North Carolina will be constant from 2013q3 through 2015q2 is extreme. The qualitative difference in policy is only observed in 2013q3 and 2013q4. After that, the other states also see an end to extended UI benefits. If we consider only 2013q3 and 2013q4 as "policy" quarters, we get the results in the third panel of Table A2. The effect on unemployment remains negative and significant. The effect on the employed is positive but insignificant. The effect on NILF has reversed sign (to positive) but remains insignificant.

[1] These are the coefficients in column 1, Table 4 in DG. The columns after that were estimated with unexplained polynomial expressions in time and unstated corrections for industry, race and age and so were impossible to replicate.

Table A2: Robustness checks for DG.

	Difference-in-differences coefficient			
	Unemployed	Employed	NILF	Observations
DG (Table 4)	-0.0097	0.012	-0.0026	425978
	(0.0019)	(0.0038)	(0.0036)	
R^2	0.002	0.002	0.002	
Replication	-0.0116	0.0140	-0.0036	232716
(25-54)	(0.0035)	(0.0082)	(0.0073)	
R^2	0.003	0.002	0.002	
Replication	-0.0198	0.0113	0.0088	457094
(short horizon)	(0.0022)	(0.0061)	(0.006)	
R^2	0.001	0.002	0.002	

(5) **Are the results for Conway consistent with this?**

Table 2 below are the results reported in Conway for quarter-by-quarter difference-in-difference coefficients for the contingent expectations of equation (1) above. If we look at the table as a series of three horizontal panels (the top for observed N_t, the middle for observed U_t and the bottom for observed E_t) we can see the concordance. For calculation, in 2013q2 the labor market shares were $\xi_{Ut-1} = 0.05$, $\xi_{Et-1} = 0.57$, and $\xi_{Nt-1} = 0.38$

- In the middle panel (unemployment outcomes) and read across for 2013q3 and 2013q4 we see that the estimated coefficients are almost all negative. Thus, the weighted average should be negative as well – just as in DG or the short-horizon replication above. The aggregate expectation is -0.022.
- Across the bottom panel for 2013q3 and 2013q4 (employment outcomes) we see mixed signs, but one significant positive effect from NILF to employment in 2013q3. This could have generated the positive effect for DG above even though the transition from unemployment to employment is negative in both quarters. The aggregate expectation is .01232.
- In the top panel for 2013q3 and 2013q4 (NILF outcomes) we observe a predominance of positive signs, consistent with the insignificant positive outcome from a short-horizon DG (above). The headline result in Conway is the significant effect in the right-most column (the transition from U to N), but that is only one of the three effects in the weighted average that makes up the DG result. The aggregate expectation is 0.00893.

Table 2: Hypothesis Test of whether NC differs from ROUS in the period of the UI Reform										
	γ_{ikt}	s_{ikt}	z		γ_{ikt}	s_{ikt}	z		γ_{ikt}	s_{ikt}	z
	Δc_{EN}				Δc_{NN}				Δc_{UN}		
2013q1	0.001	0.004	0.16		-0.016	0.018	0.90		0.065	0.039	1.67
2013q2	-0.006	0.004	1.66		0.008	0.018	0.41		-0.024	0.041	0.59
2013q3	0.004	0.004	0.90		-0.030	0.020	1.51		**0.150**	0.044	3.43
2013q4	0.004	0.004	1.00		0.009	0.019	0.48		**0.097**	0.046	2.13
	Δc_{EU}				Δc_{NU}				Δc_{UU}		
2013q1	-0.003	0.003	0.90		0.008	0.011	0.70		**-0.106**	0.049	2.14
2013q2	-0.005	0.003	1.64		-0.003	0.011	0.31		0.079	0.052	1.52
2013q3	-0.002	0.003	0.83		-0.021	0.012	1.84		-0.088	0.055	1.58
2013q4	-0.004	0.003	1.28		0.004	0.011	0.32		**-0.128**	0.058	2.23
	Δc_{EE}				Δc_{NE}				Δc_{UE}		
2013q1	0.002	0.005	0.45		0.011	0.015	0.74		0.041	0.043	0.96
2013q2	**0.011**	0.005	2.36		-0.005	0.015	0.32		-0.054	0.045	1.21
2013q3	-0.001	0.005	0.18		**0.051**	0.016	3.15		-0.062	0.048	1.30
2013q4	0.000	0.005	0.04		-0.013	0.015	0.84		0.031	0.049	0.63
N Obs			4,956,457				1,754,902				239,590
Avg Obs per quarter			55,136				11,680				2,836
Wald (159)											
	814.1	E to N			**716.1**	N to N			**946.8**	U to N	
	2062.8	E to U			**2042.2**	N to U			**4104.4**	U to U	
	748.5	E to E			**1203.0**	N to E			**3394.2**	U to E	

GLS estimation, with errors clustered by household.

In summary, the two papers begin with the same database and work with a very similar dataset (with only difference being the added observations "saved" by IPUMS). DG use the entire working-age population while Conway uses only those aged 25-54 – but that leads to less of a difference than I expected. Conway uses a methodology for linking data across months that effectively eliminates half of the observations – but that does not generate a large difference.

There are two modeling choices that explain the difference in results in the two papers. First, DG assumed that the impact of the UI reform would be constant throughout the period 2013q3-2015q2, while Conway allowed the impact to differ by quarter and focused his testing on the quarters 2013q3 and 2013q4. When this short testing horizon is introduced, the results converge. Second, Conway modeled the response to the UI reform as differing by the previous-period labor force status of the individual while DG modeled the labor-force choice as being the same for all working-age individuals. When Conway's results are aggregated up across the unemployed, employed and NILF groups, the aggregate effects are similar to those of DG.