Q1.States	Duration of benefits before change in benefits (durat)	Duration of benefits after change in benefits (durat)
Kentucky	Average: 8.330071	Average: 9.566778
Findings#1	Duration of benefits after change in benefits are longer than before change in benefits in Kentucky.	
Michigan	Average: 12.06159	Average: 15.47055
Findings#2	Duration of benefits after change in benefits are longer than before change in benefits in Michigan.	

Q1Cont'	Hypothesis	t-statistic	p-value	Findings by t-test
Kentucky	H_1: There is the difference of duration of benefits between Before and After change in benefits.	-2.1251	0.03362 (sig. & supported)	They are significantly different in terms of duration of benefits. The duration of benefits after change in benefits are longer than before change.
Michigan	H_2: There is the difference of duration of benefits between Before and After change in benefits	-2.0099	0.04465 (sig. & supported)	They are significantly different in terms of duration of benefits. The duration of benefits after change in benefits are longer than before change.
Comments	In both states, compensation on the injury duration rises after change in benefits. However, we can't simply do regression of the compensation on the duration. Because there is comparable difference of upper and lower limits (E1 and E2) on previous earnings for the low-earners groups and lower earnings limit E3 for high-earner groups (shown in Q2 Ans). So, in order to measure the impact of a policy change on compensations, we should consider treatment groups (high-earners) and control groups (low-earners) in our model.			

Q2		Kentucky (n=5626)			Michigan (n= 1524)		
Variables		Before change [A]	After change [B]	% change [C]	Before change [D]	After change [E]	% change [F]
(IA)Duration of benefits (durat) with high-earners (highearn=1)	Mean	11.18	12.89	15.30	14.78	19.43	31.46
	SD	28.99	28.25	-2.55	34.74	39.51	13.73
	Min	0.25	0.25	0.00	0.25	0.25	0.00
	Max	182	182	0.00	182	182	0.00
	Count	1233	1161	-5.84	239	219	-8.37
(IB)Duration of benefits (durat) with low-earners (high-earners=0)	Mean	6.27	7.04	12.28	10.96	13.65	24.54
	SD	12.43	16.12	29.69	26.39	34.06	29.06
	Min	0.25	0.25	0.00	0.25	0.25	0.00
	Max	182	182	0.00	182	182	0.00
	Count	1705	1527	-10.44	589	477	-19.02
(IIA) Benefit with high-earners (highearn=1)	Mean	151.08	239.09	58.25	220.66	320.48	45.24
	SD	33.68	44.28	31.47	27.49	33.56	22.08
	Min	21.06	35.26	67.43	141.01	53.83	-61.83
	Max	742.22	294.61	-60.31	473.08	459.10	-2.96
	Count	1233	1161	-5.84	239	219	-8.37
(IIB) Benefit with low-earners (high-earners=0)	Mean	118.58	118.26	-0.27	250.18	182.77	-26.94
	SD	26.45	29.04	9.79	19.03	20.34	6.88
	Min	14.87	15.93	7.13	30.70	28.50	-7.17
	Max	236.67	255.94	8.14	250.18	324.01	29.51
	Count	1705	1527	-10.44	589	477	-19.02
(IIIA) Previous weekly wage(prewage) with high-earners(high-earners=1)	Mean	475.31	482.41	1.49	749.71	739.01	-1.43
	SD	86.03	93.12	8.24	112.14	110.85	-1.15
	Min	370.49	370.83	0.09	633.42	633.24	-0.03
	Max	1237.86	1049.91	-15.18	1136.06	1583.10	39.35
	Count	1233	1161	-5.84	239	219	-8.37
(IIIB) Previous weekly wage(prewage) with low-earners(high-earns=0)	Mean	179.09	177.54	-0.87	275.83	275.65	-0.07
	SD	36.86	37.73	2.36	18.10	18.10	0.00
	Min	81.78	81.90	0.15	245.64	245.91	0.11
	Max	242.86	243.43	0.23	308.18	308.18	0.00
	Count	1705	1527	-10.44	589	477	-19.02
Sample size: High-earners		1233	1161		239	219	
Sample size: Low-earners		1705	1527		589	477	
Comments: In both States, except "previous weekly wage" in Michigan, workers' compensations rise for the high-earners group which received the benefit increase "after change in benefit", but remains less rise for low-earners group. Especially for "Benefit" with low-earners in both states, they are decrease benefit "after change in benefit".							

Q3		High-earners in Kentucky (n=2394) ;Michigan (n=458)		Low-earners in Kentucky (n=3232);Michigan (n=1066)		Difference		Differences in differences
Variables		Before change [A]	After change [B]	Before change [C]	After change [D]	[B]-[A]=[E]	[D]-[C]=[F]	[E]-[F] =[G]
(Ia)Duration of benefits (durat) in Kentucky	Mean	11.18	12.89	6.27	7.04	1.71	0.77	0.94
	SD	28.99	28.25	12.43	16.12	-0.74	3.69	-4.43
	Min	0.25	0.25	0.25	0.25	0	0	0
	Max	182	182	182	182	0	0	0
	Count	1233	1161	1705	1527	-72	-178	106
(Ib) Duration of benefits (durat) in Michigan	Mean	14.43	19.43	10.96	13.65	5	2.69	2.31
	SD	34.74	39.51	26.39	34.06	4.77	7.67	-2.9
	Min	0.25	0.25	0.25	0.25	0	0	0
	Max	182	182	182	182	0	0	0
	Count	239	219	589	477	-20	-112	92
(IIa) Log Duration of benefits (ldurat) in Kentucky	Mean	1.38	1.58	1.13	1.13	0.2	0	0.2
	SD	1.30	1.30	1.22	1.27	0	0.05	-0.05
	Min	-1.39	-1.39	-1.39	-1.39	0	0	0
	Max	5.20	5.20	5.20	5.20	0	0	0
	Count	1166	1127	1545	1375	-39	-170	131
(IIb) Log Duration of benefits (ldurat) in Michigan	Mean	1.58	1.87	1.41	1.51	0.29	0.1	0.19
	SD	1.41	1.45	1.35	1.35	0.04	0	0.04
	Min	-1.39	-1.39	-1.39	-1.39	0	0	0
	Max	5.20	5.20	5.20	5.20	0	0	0
	Count	222	207	538	451	-15	-87	72
(IIIa) Total med. costs (totmed) in Kentucky	Mean	3298.97	1688.67	878.20	1155.98	-1610.3	277.78	-1888.08
	SD	66222.87	3972.68	3223.72	6145.38	-62250.19	2921.66	-65171.85
	Min	0	0	0	0	0	0	0
	Max	2323376.50	65920.28	116883.1	198726.8	-2257456	81843.7	-2339300
	Count	1233	1161	1705	1527	-72	-178	106
(IIIb) Total med. costs (totmed) in Michigan	Mean	2229.41	2585.23	1538.23	2017.65	355.82	479.42	-123.6
	SD	4544.08	5397.981	4574.42	6657.58	853.901	2083.16	-1229.259
	Min	0	0	0	0	0	0	0
	Max	33705.36	44500.00	72544.64	113030.00	10794.64	40485.36	-29690.72
	Count	239	219	589	477	-20	-112	92
(IVa) logTotal med. costs (ltotmed) in Kentucky	Mean	6.10	6.24	5.61	5.69	0.14	0.08	0.06
	SD	1.65	1.61	1.68	1.65	-0.04	-0.03	-0.01
	Min	0	0	0	0	0	0	0
	Max	14.66	11.10	11.69	12.20	-3.56	0.51	-4.07
	Count	1233	1161	1705	1527	-72	-178	106
(IVb) logTotal med. costs (ltotmed) in Michigan	Mean	6.56	6.59	5.85	6.10	0.03	0.25	-0.22
	SD	1.67	1.86	2.07	1.96	0.19	-0.11	0.3
	Min	0	0	0	0	0	0	0
	Max	10.43	10.70	11.19	11.64	0.27	0.45	-0.18
	Count	239	219	589	477	-20	-112	92
Sample size: Kentucky		1233	1161	1705	1527			
Sample size: Michigan		239	219	589	477			

Comments: (i) The purpose of taking the log of duration is likely to be more precisely measured and less susceptible to the influence of a few large observations.(ii) In Column [E], for two states, there are larger change in duration for high-earners individuals after the increase benefits than low earners. (iii) In Column [G], Difference in difference can be remove any bias due to changes over time in factors that are common to the high and low earners groups.

Cont'Q3	Log(durat)	F-statistics	p-values	Findings	Wald Test
Kentucky	Main Effect	55.37	< 2.2e-16	There is main effect model of ldurat on the relationship on highearn & afchnge.	The purpose of Wald test is to find out explanatory variables in a model are significant. We found that Chi-squared X2 is 7.7 & P(> X2) 0.0054. It shows the parameters are not zero, which means we should include variables in the Moderation effect model.
	Moderation Effect	39.54	< 2.2e-16	There is moderation effect model of ldurat on the relationship on highearn and highearn* afchnge.	
Michigan	Main Effect	0.01079	0.0002611	There is main effect model of ldurat on the relationship on highearn & afchnge.	We found that Chi-squared X2 is 1.6 & P(> X2) 0.21. It's indicating that the overall effect of rank is not statistically significant in Moderation effect model.
	Moderation Effect	6.049	0.0004303	There is NO moderation effect model of ldurat.	

To test the moderation effect, we include the interaction term between the predictor and the moderator, and test the coefficient of the interaction.
Main Effect for Kentucky: log(durat)= 1.08725 +0.34790 (highearn)+ 0.08886 (afchnge)
Moderation Effect for Kentucky: log(durat)= 1.125615+0.256479 (highearn)+ 0.007657(afchnge) +0.190601(highearn*afchnge)
Main Effect for Michigan: log(durat)= 1.38675+ 0.25916(highearn)+ 0.15545 (afchnge)
Moderation Effect for Michigan: log(durat)= 1.41274 + 0.16914 (highearn)+ 0.09738(afchnge) + 0.19199(highearn*afchnge)
where highearn=1 for high-earns; highearn= 0 for low-earners; afchange =1 for after change; afchange = 0 for before change.

Log(durat)	Main Effect Model		Moderation Effect Model		
	Kentucky	Michigan	Kentucky	Michigan	
High-earn & before change	1.43515	1.64591	1.382094	1.58188	
High-earn & after change	1.52401	1.80136	1.580352	1.87125	
low-earn & before change	1.08725	1.38675	1.125615	1.41274	
low-earn & after change	1.17611	1.5422	1.133272	1.51012	

Further study and Suggestions for (i) main effect and interaction effect on ldurat, highearn and afchnge by ANOVA: From results of ANOVA in Kentucky, we found that highearn, afchnge and the interaction of highearn and afchnge are significant predictor with p<2.2e-16, 0.001469 and 0.001077 respectively.
In contrast, in Michigan, we found that only highearn is significant predictor (p is 0.001612). While afchnge and the interaction of highearn & afchnge are NOT significant.

Q4 To account possible changes in the compositions of the samples after the benefit increase, we need to control for worker age, martial status, sex, industry and the severity of the injury like measured by medical costs, hospital days and type of injury.

Blue colour means p-value <0.05 and sig.	**High and low earners groups pooled**				**High earners groups**			
Log (durat) [dependent variables]	Kentucky		Michigan		Kentucky		Michigan	
Explanatory variables	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
After increase (afchnge)	-0.004193	0.037935	0.003153	0.073063	0.149449	0.043860	0.259751	0.113314
High-earners group (highearn)	-0.594548	0.930448	3.607160	4.162300				
After increase (afchnge) x (highearn)	0.162092	0.058680	0.202507	0.132379				
ln(previous earnings) prewage	0.207408	0.087896	0.138989	0.561996	0.228559	0.132879	-0.335325	0.414400
ln(previous earnings) (lprewage*highearn)	0.064615	0.158335	-0.588623	0.694855				
Male indicator	-0.070215	0.039099	-0.331703	0.085692	0.003635	0.108447	-0.489189	0.527452
Married indicator	0.054997	0.034629	0.065296	0.070235	0.111690	0.064922	-0.287258	0.154478
ln(age)	0.243992	0.044230	0.480636	0.098267	0.056278	0.075064	0.850060	0.195309
ln(total medical costs)	0.360620	0.011377	0.316001	0.018473	0.420759	0.018182	0.475270	0.037405
Hospital-stay indicator	0.251557	0.043813	0.243442	0.083854	0.190595	0.065175	-0.279256	0.148219
Industry indicators: manufacturing	-0.153458	0.035219	-0.125977	0.067752	-0.199989	0.060746	0.008071	0.141096
construction	0.043594	0.044348	0.347968	0.087642	0.034756	0.058244	0.466139	0.131284
Injury type indicators: Head	-0.431864	0.109058	-0.833484	0.234882	-0.335292	0.163143	-1.122075	0.410394
Neck	0.357815	0.136579	-0.282797	0.295556	0.479260	0.180361	-0.560021	0.477210
Upper extremities	0.131935	0.086065	-0.222958	0.163224	0.139025	0.128953	-0.160986	0.307717
Trunk	0.143441	0.091958	-0.135308	0.175506	0.154524	0.134563	-0.153161	0.316875
Low back	0.184857	0.085893	-0.340506	0.165465	0.259297	0.127208	-0.315403	0.309712
Lower extremities	0.184361	0.086767	-0.309005	0.166413	0.263831	0.128324	-0.392832	0.307266
Occupational diseases	0.525809	0.178116	0.352871	0.283137	0.759075	0.242183	-2.138409	0.867296
Sample size	5347		1475		2231		447	
R^2	0.3189		0.304		0.363		0.3995	

Observations: (i) In the absence of time varying explanatory variables, commonly use log-logistic hazard models with and without unobserved heterogeneity, which is special case of log-duration regression. It assumes some distributions for the error term which leave unspecified. Therefore, the sample sizes are slightly smaller than previous due to missing variables. (ii) "after-increase" variables in "High and low earners groups pooled" are common. The "highearn" and interaction "ln(previous earnings x highearn)" are net out any time invariant differences between high and low earners groups. (iii) The interaction "After increase (afchnge) x (highearn)" which experienced the increase in benefits in both states. (iv) For "High earners groups" found that large increase in durations after the benefit increases.

Q5. The potential problem of the above model is endogenous. The four pairs estimate the benefit increases are significant and similar in two States. In addition, we also found that medical costs, hospital stay and age indicators are significant variables in above regression model. These are the best factors to measure the injury severity but they are potentially endogenous. As higher medical costs, longer hospital stays and being older all lead to longer time to recover until they can return to work.

Further Study with Matching in Kentucky & Michigan, which are Estimated by the propensity model. Treatment group: high earners; Control group: low-earners.
In Kentucky, there is a significant impact with p is 0.011935. The log duration is increased by 0.12337 on average. For covariate balance, we found that male, married, log (age), hosp, manuf, construc are insignificant with p> 0.05.
In Michigan, there is not a significant impact with p is 0.54649. For covariate balance, we found that male, married, ltotmed, hosp are insignificant with p> 0.05.