FINAL REPORT

IMPACT OF MULTIPLE FACTORS ON TOTAL TRAFFIC FATALITIES: PANEL DATA EVIDENCE

BRIOUGHT TO YOU BY
TASHVEEN KAUR, CHESTA VIRANI, DHRUVI SONI

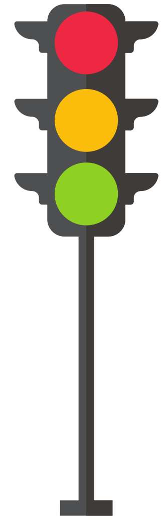
ABSTRACT

This project examines the impact of multiple factors like seatbelts, alcohol level, speed limit, etc on total traffic fatalities. In terms of estimation, the research employs a data set that is a panel data collection containing information on the effects of various traffic legislation on mortality in 48 states between 1980 and 2004. We are going to test our model in advanced methods which are Fe (Fixed Effects) and Re (Random Effects) and Hausman test.
The results derived from this study suggest that:
(1) Vehicle miles traveled in billions and Blood alcohol limit more than 10% and 8% is positively related to total fatalities
(2)R square is 99.36% for our model
(3)After Hausman Test we have chosen Random Effects Model

BACKGROUND RESEARCH

The rising number of road fatalities is a national crisis; we cannot and must not accept these deaths as unavoidable. According to the data, drivers of passenger vehicles engaged in at least one of the following risky behaviors in 45% of fatal crashes: speeding, alcohol impairment, or not wearing a seat belt and many more. Many states, for example, have enacted legislation to help prevent this. Graduated driver licensing (GDL) is a system that allows new drivers to gain experience in low-risk situations by gradually granting driving privileges..

INTRODUCTION



This model depicts how multiple factors influenced traffic fatalities in the United States from 1980 to 2004 across 48 states. We have 1200 observations and "State" is our Entity and our time variable for panel data are the number of years.

ADVANCED METHOD

In an essay, article, or book, an introduction is a beginning section which states the purpose and goals of the following writing. This is generally followed by the body and conclusion.

The introduction typically describes the scope of the document and gives the brief explanation or summary of the document. It may also explain certain elements that are important to the essay.

TYPES OF ADVANCED METHODS

FE: In FE Unobserved, time-invariant effects that are correlated with error are controlled for. When omitted variable bias is an issue, this is a huge benefit.

RE: The weighted average of the within/fixed effect estimator and the between estimator compensates the random effects estimator. If fixed effects are not associated with x's, random effects estimators will be reliable and fair. Fixed effects estimators are always reliable and impartial. Because they use more of the variation in X, random effects estimators will be more effective than fixed effects estimators.

IV(instrumental variable): When controlled experiments are not feasible or a treatment is not successfully delivered to every unit in a randomized experiment, the method of instrumental variables (IV) is used to estimate causal relationships. Changes in the dependent variable affect the value of at least one of the covariates in this method. There are missing variables that influence both the dependent and independent variables, or the covariates have non-random measurement error.

METHODOLOGY

Our primary goal is to analyze how advanced techniques affect our model. The panel estimations we use consider a variety of techniques. We have employed Fe and Re as our techniques.
We will first make a list of the variables that will be included in our model and utilized to represent our panel data.
To obtain this, we utilized Stata's xtset year state command. We have used multiple regression and robust using pooled OLS and robust standard errors that have been corrected for heteroskedasticity and serial correlation.
The random effects estimator is the next technique we performed which removed partial serial correlation.
To run the fixed and random impact models, we established dummy variables for each year (y80, y81, y82, ..., y24). The R2 rose as a result of the dummy variables.
Each of these dummy variables absorbs the effects that are unique to each year.
The dummy variable was employed in FE and RE approaches because it allowed each pair of observations to get a tailored coefficient, allowing the R2 to be calculated.

DATA SOURCE

The driving data set was gathered from the datasets supplied by the professor. The data set is a panel data collection that contains information on the impact of various traffic legislation on deaths in 48 states between 1980 and 2004.

It in the form of Panel data or longitudinal data. Panel data is a collection of observations gathered chronologically at regular intervals regarding various cross sections of time. It gives data on individual behavior across time and across persons. The data and models are cross-sectional and time-series in nature. When all people are observed in all time periods, panel data is balanced. When individuals are not observed in all time periods, panel data is unbalanced. A pool of data is utilized to infer outcomes for an individual in this case. This aids in precise prediction.

DESCRIPTION OF EACH VARIABLE

Number	Variable	Description	Values
1	Year	This variable defines the year	The value varies from 1980-2004
2	State	48 continental states, alphabetically	The values range from 1 to 51
3	Speed limit 55	This variable defines the speed limit as 55	1= if the speed limit is 55
			0 = if the speed limit is not 55
4	Speed limit 65	This variable defines the speed limit as 65	1= if the speed limit is 65
			0 = if the speed limit is not 65
5	Speed limit 70	This variable defines the speed limit as 70	1= if the speed limit is 70
			0 = if the speed limit is not 70
6	Speed limit 75	This variable defines the speed limit as 75	1= if the speed limit is 75
			0 = if the speed limit is not 75
7	No speed Limit	This variable defines the places where there is no speed limit	0=defines the places with no speed limit zones
			1=defines speed limit zones
8	Seat belt	This variable defines people wearing seat belts while driving	0= if no one wearing seatbelt
			1= primary person wearing seatbelt
			2= secondary also wearing seat belt
9	Minimum drinking age	This variable defines the minimum age of drinking in different years in different states.	The value ranges from 18-21 according to different years and states
10	Zero tolerance Law	This variable defines the zero-tolerance law among the drivers	0= if the law was not there
			1= if the law was there
11	Graduated driving license	This variable defines people driving with a driving license	0 = no people driving without a license
			1= people driving without a license
12	Blood alcohol limit .10	This variable the alcohol level found in person while driving is whether 10% or not	1 = people having alcohol level is 10%
			0 = having alcohol level less than 10%
13	Blood alcohol limit 0.08	This variable the alcohol level found in person while driving is whether 8% or not	1 = people having alcohol level is 8%
			0 = having alcohol level less than 8%
14	Administrative revocation of license	This variable defines revocation of license as per law.	0 = revocation of license
			1= not revocation of license
15	Total traffic fatalities	This variable defines total traffic fatalities in a particular year	The value varies every year, this value is the number of deaths that result from the crash within 30 days

16	Total nighttime fatalities	This variable defines total nighttime fatalities in a particular year	The value varies every year, this value is the number of deaths that result from the crash in the nighttime.
17	Total weekend fatalities	This variable defines total weekend fatalities in a particular year	The value varies every year, this value is the number of deaths that result from the crash on the weekend.
18	Total fatalities per 100 million miles	This variable defines total fatalities per 100 million miles in a particular year	The value differs from 0%-6% in total fatalities
19	Nighttime fatalities per 100 million miles	This variable defines nighttime fatalities per 100 million miles in a particular year	The value differs from 0%-3% in nighttime fatalities
20	Weekend fatalities per 100 million miles	This variable defines total weekend fatalities per 100 million miles in a particular year	The value differs from 0%-2% in weekend fatalities
21	State population	The variable defines the population.	The value differs in different years and different states. n
22	Total fatalities per 100,000 population	This variable defines total fatalities in a particular year per 100,000 population	The value differs from 0%-40% in total fatalities
23	Nighttime fatalities per 100,000 population	This variable defines nighttime fatalities in a particular year per 100,000 population	The value differs from 0%-20% in nighttime fatalities
24	Weekend fatalities per 100,000 population	This variable defines weekend fatalities in a particular year per 100,000 population	The value differs from 0%-10% in total fatalities
25	Vehicle miles travelled in billions	This variable defines the vehicle's miles that have been in billions	The value differs every year.
26	Unemployment percentage	This variable defines unemployment	The value differs from 0%-10%
27	Percentage population aged through 14 through 24	This variable defines the percentage of population driving through the age of 14-24	The value differs from 0%-20%.
28	Speed limit 70/75/no speed limit	The variable defines no speed limit and speed limit more than 70	1 = if there is no speed limit and speed limit more than 70
			0= if there is speed limit and speed limit more than 70
29	Primary seat belt law	This variable defines that people wear primary seat belt.	0 = not wearing primary seat belt
			1 = wearing primary seat belt
30	Secondary seat belt law	This variable defines that people wear secondary seat belt.	0 = not wearing secondary seat belt
			1 = wearing secondary seat belt
31	D80-D04	These variables define the year with driving	1= if the D is equal to respective years.

	year	totfat	vehicm~s	seatbelt	gdl	bac10	bac08	nghtfat	wkndfat	statepop
1.	1980	940	29.375	0	0	1	0	422	236	3.9e+06
2.	1981	933	27.852	0	0	1	0	434	248	3.9e+06
3.	1982	839	29.8577	0	0	1	0	376	224	3.9e+06
4.	1983	930	31	0	0	1	0	397	223	3.9e+06
5.	1984	932	32.9329	0	0	1	0	421	237	4.0e+06
6.	1985	882	35.1394	0	0	1	0	358	224	4.0e+06
7.	1986	1080	33.9937	0	0	1	0	500	279	4.0e+06
8.	1987	1111	37.4074	0	0	1	0	499	300	4.0e+06
9.	1988	1024	39.6899	0	0	1	0	423	226	4.0e+06
10.	1989	1029	40.8333	0	0	1	0	418	247	4.0e+06
11.	1990	1121	42.3019	0	0	1	0	466	271	4.0e+06
12.	1991	1116	42.9231	0	0	1	0	474	276	4.1e+06
13.	1992	1031	45.8222	2	0	1	0	408	218	4.2e+06
14.	1993	1044	47.2398	2	0	1	0	441	235	4.2e+06
15.	1994	1083	49.0045	2	0	1	0	451	267	4.3e+06
16.	1995	1114	50.6364	2	0	.583	.417	463	260	4.3e+06
17.	1996	1146	51.3901	2	0	0	1	512	268	4.3e+06
18.	1997	1192	53.4529	2	0	0	1	518	255	4.4e+06
19.	1998	1071	55.2062	2	0	0	1	433	233	4.4e+06
20.	1999	1138	56.0591	1	0	0	1	486	257	4.4e+06
21.	2000	996	56.5909	1	0	0	1	418	220	4.4e+06
22.	2001	991	56.6286	1	0	0	1	398	198	4.5e+06
23.	2002	1038	57.3889	1	.75	0	1	468	230	4.5e+06
24.	2003	1004	58.538	1	1	0	1	440	242	4.5e+06
25.	2004	1154	59.1795	1	1	0	1	486	269	4.5e+06
26.	1980	947	17.8679	0	0	0	0	503	259	2.7e+06
27.	1981	917	18.57	0	0	0	0	434	222	2.8e+06
28.	1982	724	20.3944	0	0	.417	0	337	177	2.9e+06
29.	1983	675	20.5793	0	0	1	0	319	177	3.0e+06
30.	1984	869	20.9398	0	0	1	0	386	171	3.1e+06
31.	1985	893	21.5701	0	0	1	0	444	215	3.2e+06
32.	1986	1007	22.6802	0	0	1	0	492	252	3.3e+06
33.	1987	939	31.723	0	0	1	0	428	196	3.4e+06
34.	1988	944	34.2029	0	0	1	0	386	199	3.5e+06
35.	1989	879	34.881	0	0	1	0	366	162	3.6e+06
36.	1990	869	35.4694	0	0	1	0	389	195	3.7e+06
37.	1991	816	34.8718	2	0	1	0	353	177	3.8e+06
38.	1992	809	35.0216	2	0	1	0	353	167	3.9e+06
39.	1993	801	37.6056	2	0	1	0	368	188	4.1e+06
40.	1994	904	38.7983	2	0	1	0	395	197	4.2e+06
41.	1995	1035	39.6552	2	0	1	0	475	245	4.4e+06
42.	1996	994	42.1186	2	0	1	0	469	249	4.6e+06
43.	1997	951	43.4247	2	0	1	0	447	230	4.7e+06
44.	1998	980	45.1613	2	0	1	0	419	220	4.9e+06
45.	1999	1024	46.758	2	0	1	0	435	207	5.0e+06
46.	2000	1036	49.0995	2	0	1	0	454	209	5.1e+06
47.	2001	1051	49.5755	2	0	.667	.333	502	260	5.3e+06
48.	2002	1132	51.2385	2	0	0	1	507	284	5.4e+06
49.	2003	1118	53.8461	2	0	0	1	544	296	5.6e+06
50.	2004	1150	57.2139	2	0	0	1	527	250	5.7e+06

REGRESSION

CODE: REG TOTFAT VEHICMILES YEAR SEATBELT GDL BAC10 BAC08 NGHTFAT WKNDFAT STATEPOP

Source	SS	df	MS
Model	920585348	9	102287261
Residual	5974286	1,190	5020.4084
Total	926559634	1,199	772777.009

Number of obs	=	1,200
F(9, 1190)	=	20374.29
Prob > F	=	0.0000
R-squared	=	0.9936
Adj R-squared	=	0.9935
Root MSE	=	70.855

totfat	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
vehicmiles	3.527085	.2186774	16.13	0.000	3.098049	3.956122
year	-.2162204	.6064351	-0.36	0.721	-1.406021	.9735806
seatbelt	8.492169	3.207807	2.65	0.008	2.198581	14.78576
gdl	-27.21807	7.468595	-3.64	0.000	-41.87115	-12.56499
bac10	27.92899	6.433433	4.34	0.000	15.30686	40.55113
bac08	22.38262	8.833922	2.53	0.011	5.050825	39.71442
nghtfat	2.309643	.0646575	35.72	0.000	2.182788	2.436499
wkndfat	-1.119373	.119695	-9.35	0.000	-1.35421	-.8845364
statepop	-8.28e-06	1.73e-06	-4.80	0.000	-.0000117	-4.90e-06
_cons	447.2231	1203.629	0.37	0.710	-1914.249	2808.695

CODE: REG TOTFAT VEHICMILES YEAR SEATBELT GDL BAC10 BAC08 NGHTFAT WKNDFAT STATEPOP, ROBUST

Linear regression

Number of obs	=	1,200
F(9, 1190)	=	7960.45
Prob > F	=	0.0000
R-squared	=	0.9936
Root MSE	=	70.855

totfat	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
vehicmiles	3.527085	.3284516	10.74	0.000	2.882677	4.171494
year	-.2162204	.5916012	-0.37	0.715	-1.376918	.9444772
seatbelt	8.492169	2.840636	2.99	0.003	2.918957	14.06538
gdl	-27.21807	6.811859	-4.00	0.000	-40.58266	-13.85348
bac10	27.92899	6.864964	4.07	0.000	14.46021	41.39777
bac08	22.38262	8.301569	2.70	0.007	6.095281	38.66996
nghtfat	2.309643	.0998296	23.14	0.000	2.113782	2.505505
wkndfat	-1.119373	.1870239	-5.99	0.000	-1.486306	-.7524399
statepop	-8.28e-06	2.72e-06	-3.04	0.002	-.0000136	-2.94e-06
_cons	447.2231	1174.365	0.38	0.703	-1856.834	2751.28

FIXED EFFECTS MODEL

A statistical model called a "fixed effects model" uses fixed or non-random model parameters. In panel data, if the same individual has longitudinal observations, fixed effects indicate the subject-specific means. An estimator for the coefficients in the regression model considering those fixed effects is referred to as a fixed effects estimator (also known as the inside estimator) in panel data analysis. The fixed effects model's main benefit is that it enables us to account for all time invariant missing variables.
This is crucial when dealing with variables that are challenging or impossible to observe.
The main drawback is that we need to estimate a lot of extra parameters.

INTERPRETATION

In order to remove bias, the fixed effects model removes the effect of time invariant characters. Because of collinearity, the variable y24 was omitted. According to our model, the most significant advantage of FE is that it controls all time invariants and time differences. We concluded from the model after running the FE command that three variables are significant: vehicle miles, seat belt, and nightfat. The standard error for vehicle miles is 0.32. Our RHO value is 0.81, which equals 81%, and our R square is 91%.

CODE: XTREG TOTFAT VEHICMILES YEAR SEATBELT GDL BAC10 BAC08 NGHTFAT WKNDFAT STATEPOP Y80 Y81 Y82 Y83 Y84 Y85 Y86 Y87 Y88 Y89 Y90 Y91 Y92 Y93 Y94 Y95 Y96 Y97 Y98 Y99 Y20 Y21 Y22 Y23 Y24, FE

```
Fixed-effects (within) regression               Number of obs     =      1,200
Group variable: state                           Number of groups  =         48

R-squared:                                      Obs per group:
     Within  = 0.9183                                         min =         25
     Between = 0.9949                                         avg =       25.0
     Overall = 0.9925                                         max =         25

                                                F(33,1119)        =     381.09
corr(u_i, Xb) = 0.7221                          Prob > F          =     0.0000
```

totfat	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
vehicmiles	4.099359	.3702609	11.07	0.000	3.372875	4.825842
slcom	-9.28e+07	7.55e+08	-0.12	0.902	-1.57e+09	1.39e+09
seatbelt	8.088135	2.518826	3.21	0.001	3.145982	13.03029
gdl	-8.080032	6.305495	-1.28	0.200	-20.45196	4.291893
bac10	-4.806163	5.786394	-0.83	0.406	-16.15957	6.547242
bac08	-11.22682	8.311942	-1.35	0.177	-27.53557	5.081926
nghtfat	2.083953	.0552958	37.69	0.000	1.975458	2.192448
wkndfat	-.7144458	.1016539	-7.03	0.000	-.9138996	-.514992
statepop	-.0000242	5.30e-06	-4.57	0.000	-.0000346	-.0000138
y80	-27.47787	14.33129	-1.92	0.055	-55.59711	.6413623
y81	-29.13733	14.29762	-2.04	0.042	-57.19049	-1.084171
y82	-25.94719	14.0997	-1.84	0.066	-53.61202	1.717646
y83	1.298757	13.63975	0.10	0.924	-25.46361	28.06112
y84	15.64253	13.22542	1.18	0.237	-10.30688	41.59194
y85	28.14817	13.01434	2.16	0.031	2.612917	53.68342
y86	21.63482	12.4453	1.74	0.082	-2.783929	46.05356
y87	32.8252	12.21905	2.69	0.007	8.85036	56.80004
y88	35.42164	11.99138	2.95	0.003	11.89351	58.94978
y89	33.46947	11.83949	2.83	0.005	10.23938	56.69956
y90	22.61082	11.78857	1.92	0.055	-.5193777	45.74101
y91	18.3286	11.65817	1.57	0.116	-4.545745	41.20293
y92	12.52502	11.72182	1.07	0.286	-10.47419	35.52423
y93	30.5125	11.59849	2.63	0.009	7.755268	53.26973
y94	15.35628	11.32179	1.36	0.175	-6.858045	37.5706
y95	9.813398	11.22667	0.87	0.382	-12.2143	31.8411
y96	6.491759	11.18828	0.58	0.562	-15.46062	28.44413
y97	23.22156	10.91427	2.13	0.034	1.806825	44.63629
y98	18.64482	10.55796	1.77	0.078	-2.070815	39.36046
y99	17.02593	10.02386	1.70	0.090	-2.641748	36.69361
y20	7.051568	9.755834	0.72	0.470	-12.09022	26.19336
y21	7.717029	9.371088	0.82	0.410	-10.66985	26.10391
y22	-6.74436	9.143159	-0.74	0.461	-24.68403	11.19531
y23	4.535091	8.992334	0.50	0.614	-13.10864	22.17883
y24	0	(omitted)				
_cons	9.28e+07	7.55e+08	0.12	0.902	-1.39e+09	1.57e+09
sigma_u	92.394173					
sigma_e	43.63129					
rho	.81766077	(fraction of variance due to u_i)				

```
F test that all u_i=0: F(47, 1119) = 40.22                  Prob > F = 0.0000
```

FIXED EFFECTS CLUSTER

With fixed effects, a main reason to cluster is you have heterogeneity in treatment effects across the clusters. There are other reasons, for example if the clusters (State) are a subset of the clusters in the population (about which we are inferring).

CODE: XTREG TOTFAT VEHICMILES YEAR SEATBELT GDL BAC10 BAC08 NGHTFAT WKNDFAT STATEPOP Y80 Y81 Y82 Y83 Y84 Y85 Y86 Y87 Y88 Y89 Y90 Y91 Y92 Y93 Y94 Y95 Y96 Y97 Y98 Y99 Y20 Y21 Y22 Y23 Y24, FE CLUSTER(STATE)

NOTE: Y23 OMITTED BECAUSE OF COLLINEARITY
Y24 OMITTED BECAUSE OF COLLINEARITY

```
Fixed-effects (within) regression               Number of obs     =      1,200
Group variable: state                           Number of groups  =         48

R-squared:                                      Obs per group:
     Within  = 0.9183                                         min =         25
     Between = 0.9949                                         avg =       25.0
     Overall = 0.9925                                         max =         25

                                                F(33,47)          =     360.94
corr(u_i, Xb) = 0.7221                          Prob > F          =     0.0000

                                  (Std. err. adjusted for 48 clusters in state)
```

totfat	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
vehicmiles	4.099359	1.400386	2.93	0.005	1.282146	6.916572
slcom	-9.28e+07	1.08e+11	-0.00	0.999	-2.17e+11	2.17e+11
seatbelt	8.088135	8.576363	0.94	0.350	-9.165281	25.34155
gdl	-8.080032	14.68361	-0.55	0.585	-37.61965	21.45959
bac10	-4.806163	8.73356	-0.55	0.585	-22.37582	12.76349
bac08	-11.22682	11.21697	-1.00	0.322	-33.79245	11.33881
nghtfat	2.083953	.1885354	11.05	0.000	1.704669	2.463238
wkndfat	-.7144458	.2663821	-2.68	0.010	-1.250338	-.1785541
statepop	-.0000242	.0000244	-0.99	0.326	-.0000733	.0000248
y80	-34.52944	13.29699	-2.60	0.013	-61.27954	-7.779338
y81	-36.1889	13.0568	-2.77	0.008	-62.4558	-9.922002
y82	-32.99876	13.03017	-2.53	0.015	-59.21208	-6.785433
y83	-5.752811	13.13126	-0.44	0.663	-32.1695	20.66388
y84	8.59096	11.99068	0.72	0.477	-15.53119	32.71311
y85	21.0966	11.58327	1.82	0.075	-2.205934	44.39913
y86	14.58325	12.04753	1.21	0.232	-9.653261	38.81976
y87	25.77363	11.92665	2.16	0.036	1.780313	49.76694
y88	28.37008	12.79403	2.22	0.031	2.631804	54.10835
y89	26.4179	13.26537	1.99	0.052	-.2685823	53.10439
y90	15.55925	10.96416	1.42	0.162	-6.497804	37.6163
y91	11.27703	12.52246	0.90	0.372	-13.91491	36.46896
y92	5.473449	10.56939	0.52	0.607	-15.78943	26.73633
y93	23.46093	14.12734	1.66	0.103	-4.95962	51.88148
y94	8.304711	13.2427	0.63	0.534	-18.33616	34.94558
y95	2.761829	11.80147	0.23	0.816	-20.97967	26.50333
y96	-.5598095	129.5963	-0.00	0.997	-261.274	260.1544
y97	16.16999	10.73366	1.51	0.139	-5.42334	37.76332
y98	11.59325	8.405719	1.38	0.174	-5.316874	28.50338
y99	9.974361	7.942967	1.26	0.215	-6.004828	25.95355
y21	.6654604	8.820905	0.08	0.940	-17.07991	18.41083
y22	-13.79593	10.19034	-1.35	0.182	-34.29626	6.704399
y23	-2.516478	10.95524	-0.23	0.819	-24.55558	19.52262
y23	0	(omitted)				
y24	-7.051568	11.47435	-0.61	0.542	-30.13499	16.03185
_cons	9.28e+07	1.08e+11	0.00	0.999	-2.17e+11	2.17e+11
sigma_u	92.394173					
sigma_e	43.63129					
rho	.81766077	(fraction of variance due to u_i)				

RANDOM EFFECTS MODEL

The Random Effects regression model is used to quantify the effect of intrinsically unmeasurable individual-specific attributes such as grit or acumen. Individual-specific effects like these are common in panel data research. The estimation of the RE model entails the following two steps:
Estimation of the variance components and their relationship to the composite residual error.
Estimating the regression coefficients and determining the common bias (which we forms the intercept term of the regression model).Over fixed effect models, random effects models have at least two significant advantages: 1) the ability to estimate shrunken residuals; 2) the ability to account for unequal school effectiveness using random coefficients models

INTERPRETATION

To estimate the effect of individual-specific characteristics, the Random Effects regression model is used. Random effects have the advantage of including time invariant variables such as gender, whereas fixed effects absorb all time invariant variables. Individual error terms are not correlated with predictors in this case, allowing time invariant variables to serve as explanatory variables. Two common assumptions can be made the random effects assumption and the fixed effects assumption. The assumption of random effects is that individual unobserved heterogeneity is uncorrelated with the independent variables. According to the fixed effect assumption, the individual specific effect is related to the independent variables.

CODE: XTREG TOTFAT VEHICMILES YEAR SEATBELT GDL BAC10 BAC08 NGHTFAT WKNDFAT STATEPOP Y80 Y81 Y82 Y83 Y84 Y85 Y86 Y87 Y88 Y89 Y90 Y91 Y92 Y93 Y94 Y95 Y96 Y97 Y98 Y99 Y20 Y21 Y22 Y23 Y24, RE

NOTE: SLCOM OMITTED BECAUSE OF COLLINEARITY
Y24 OMITTED BECAUSE OF COLLINEARITY

```
Random-effects GLS regression                   Number of obs     =      1,200
Group variable: state                           Number of groups  =         48

R-squared:                                      Obs per group:
     Within  = 0.9176                                         min =         25
     Between = 0.9959                                         avg =       25.0
     Overall = 0.9937                                         max =         25

                                                Wald chi2(32)     =   22506.72
corr(u_i, X) = 0 (assumed)                      Prob > chi2       =     0.0000
```

totfat	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
vehicmiles	3.476031	.3037003	11.45	0.000	2.880789	4.071272
slcom	0	(omitted)				
seatbelt	7.955665	2.512351	3.17	0.002	3.031547	12.87978
gdl	-8.779082	6.299889	-1.39	0.163	-21.12664	3.568474
bac10	-2.988909	5.735316	-0.52	0.602	-14.22992	8.252103
bac08	-12.09281	8.197031	-1.48	0.140	-28.15869	3.973078
nghtfat	2.164225	.0507792	42.62	0.000	2.0647	2.263751
wkndfat	-.8116944	.0975048	-8.32	0.000	-1.0028	-.6205886
statepop	-.0000107	3.44e-06	-3.10	0.002	-.0000174	-3.93e-06
y80	-38.98806	12.27586	-3.18	0.001	-63.0483	-14.92782
y81	-40.87632	12.21731	-3.35	0.001	-64.8218	-16.93084
y82	-35.89933	12.01965	-2.99	0.003	-59.45742	-12.34125
y83	-8.130549	11.59117	-0.70	0.483	-30.84882	14.58772
y84	5.58027	11.26233	0.50	0.620	-16.4935	27.65403
y85	18.1143	11.05071	1.64	0.101	-3.544691	39.77329
y86	11.35245	10.45306	1.09	0.277	-9.135177	31.84008
y87	23.32101	10.24254	2.28	0.023	3.245989	43.39602
y88	27.01246	10.04431	2.69	0.007	7.325963	46.69895
y89	25.58688	9.922066	2.58	0.010	6.13999	45.03377
y90	15.34755	9.890576	1.55	0.121	-4.037624	34.73272
y91	11.94819	9.768673	1.22	0.221	-7.198055	31.09444
y92	5.504384	9.81564	0.56	0.575	-13.73392	24.74269
y93	23.50225	9.742297	2.41	0.016	4.4077	42.5968
y94	8.171733	9.632375	0.85	0.396	-10.70738	27.05084
y95	2.306455	9.586737	0.24	0.810	-16.48321	21.09611
y96	-1.818829	9.580933	-0.19	0.849	-20.59711	16.95946
y97	15.61432	9.423253	1.66	0.098	-2.854919	34.08355
y98	11.1759	9.225541	1.21	0.226	-6.905825	29.25763
y99	9.513609	8.98909	1.06	0.290	-8.104683	27.1319
y21	.0620121	9.044464	0.01	0.995	-17.66481	17.78884
y22	-14.65521	9.22007	-1.59	0.112	-32.72621	3.415794
y23	-2.916737	9.391755	-0.31	0.756	-21.32424	15.49077
y23	0	(omitted)				
y24	-6.620246	9.755732	-0.68	0.497	-25.74113	12.50064
_cons	46.8579	15.4853	3.03	0.002	16.50726	77.20853
sigma_u	59.099937					
sigma_e	43.612102					
rho	.64743658	(fraction of variance due to u_i)				

RANDOM EFFECTS CLUSTER

CODE: XTREG TOTFAT VEHICMILES YEAR SEATBELT GDL BAC10 BAC08 NGHTFAT WKNDFAT STATEPOP Y80 Y81 Y82 Y83 Y84 Y85 Y86 Y87 Y88 Y89 Y90 Y91 Y92 Y93 Y94 Y95 Y96 Y97 Y98 Y99 Y20 Y21 Y22 Y23 Y24, RE CLUSTER(STATE)

NOTE: Y23 OMITTED BECAUSE OF COLLINEARITY
Y24 OMITTED BECAUSE OF COLLINEARITY

```
Random-effects GLS regression                   Number of obs     =      1,200
Group variable: state                           Number of groups  =         48

R-squared:                                      Obs per group:
     Within  = 0.9176                                         min =         25
     Between = 0.9959                                         avg =       25.0
     Overall = 0.9937                                         max =         25

                                                Wald chi2(32)     =   96964.71
corr(u_i, X) = 0 (assumed)                      Prob > chi2       =     0.0000

                                 (Std. err. adjusted for 48 clusters in state)
```

totfat	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
vehicmiles	3.476031	.7564982	4.59	0.000	1.993321	4.95874
slcom	0	(omitted)				
seatbelt	7.955665	2.897157	2.75	0.006	2.277342	13.63399
gdl	-8.779082	12.87003	-0.68	0.495	-34.00387	16.44571
bac10	-2.988909	8.575011	-0.35	0.727	-19.79562	13.8178
bac08	-12.09281	9.420067	-1.28	0.199	-30.5558	6.370185
nghtfat	2.164225	.2157789	10.03	0.000	1.741307	2.587144
wkndfat	-.8116944	.2956721	-2.75	0.006	-1.391201	-.2321878
statepop	-.0000107	.0000106	-1.01	0.313	-.0000314	.0000101
y80	-38.98806	15.00862	-2.60	0.009	-68.40441	-9.571703
y81	-40.87632	14.58971	-2.80	0.005	-69.47162	-12.28102
y82	-35.89933	12.53567	-2.86	0.004	-60.46879	-11.32987
y83	-8.130549	11.31903	-0.72	0.473	-30.31544	14.05434
y84	5.58027	10.77443	0.52	0.605	-15.53722	26.69776
y85	18.1143	9.808379	1.85	0.065	-1.109771	37.33837
y86	11.35245	10.76103	1.05	0.291	-9.738775	32.44368
y87	23.32101	11.21121	2.08	0.038	1.347443	45.29457
y88	27.01246	11.85558	2.28	0.023	3.775951	50.24896
y89	25.58688	12.36246	2.07	0.038	1.35691	49.81685
y90	15.34755	9.724451	1.58	0.115	-3.712026	34.40712
y91	11.94819	10.60708	1.13	0.260	-8.841301	32.73769
y92	5.504384	8.164451	0.67	0.500	-10.49765	21.50641
y93	23.50225	12.39708	1.90	0.058	-.7955805	47.80008
y94	8.171733	11.53542	0.71	0.479	-14.43728	30.78075
y95	2.306455	9.877293	0.23	0.815	-17.05268	21.66559
y96	-1.818829	10.57245	-0.17	0.863	-22.54044	18.90278
y97	15.61432	9.588274	1.63	0.103	-3.178354	34.40699
y98	11.1759	7.862624	1.42	0.155	-4.234556	26.58636
y99	9.513609	8.163794	1.17	0.244	-6.487133	25.51435
y21	.0620121	8.784067	0.01	0.994	-17.15444	17.27847
y22	-14.65521	9.470997	-1.55	0.122	-33.21802	3.907602
y23	-2.916737	9.701597	-0.30	0.764	-21.93152	16.09805
y23	0	(omitted)				
y24	-6.620246	9.585884	-0.69	0.490	-25.40823	12.16774
_cons	46.8579	15.64297	3.00	0.003	16.19824	77.51756
sigma_u	59.099937					
sigma_e	43.612102					
rho	.64743658	(fraction of variance due to u_i)				

RESULTS

Hausman test

In a regression model, the Hausman Test finds endogenous regressors (predictor variables). The values of endogenous variables are influenced by other variables in the system. Ordinary least squares estimators will not work in the presence of endogenous regressors since one of the underlying assumptions is that there is no correlation between the predictor variable and the error term.The Hausman test can assist you in deciding between a fixed effects model and a random effects model when conducting panel data analysis. The alternative hypothesis is that the preferred model is one with fixed effects; the null hypothesis is that the preferred model is one with random effects. The test basically checks to determine if there is a relationship between the unique errors and the model's regressors. There is no association between the two, according to the null hypothesis.It's quite simple to interpret the results of a Hausman test: if the p-value is low (less than 0.05), reject the null hypothesis, and vice versa. The p-value in our model is greater than 0.05 which is 0.9944, indicating that the random effect model is appropriate for the driving data set.

	Coefficients (b) fe	(B) re	(b-B) Difference	sqrt(diag(V_b-V_B)) Std. err.
vehicmiles	4.099359	3.476031	.6233279	.2118001
seatbelt	8.088135	7.955665	.1324702	.1804826
gdl	-8.080032	-8.779082	.6990499	.2658238
bac10	-4.806163	-2.988909	-1.817254	.7671485
bac08	-11.22682	-12.09281	.8659871	1.37734
nghtfat	2.083953	2.164225	-.0802722	.0218882
wkndfat	-.7144458	-.8116944	.0972486	.0287463
statepop	-.0000242	-.0000107	-.0000135	4.03e-06
y80	-34.52944	-38.98806	4.458618	1.645571
y81	-36.1889	-40.87632	4.68742	1.732379
y82	-32.99876	-35.89933	2.900575	1.452603
y83	-5.752811	-8.130549	2.377738	1.362727
y84	8.59096	5.58027	3.01069	1.282179
y85	21.0966	18.1143	2.982301	1.309475
y86	14.58325	11.35245	3.230796	1.130782
y87	25.77363	23.32101	2.452623	.8440197
y88	28.37008	27.01246	1.35762	.321314
y89	26.4179	25.58688	.8310214	.
y90	15.55925	15.34755	.2116987	.
y91	11.27703	11.94819	-.6711655	.495532
y92	5.473449	5.504384	-.030935	.1934991
y93	23.46093	23.50225	-.0413226	.4626876
y94	8.304711	8.171733	.132978	.
y95	2.761829	2.306455	.4553747	.
y96	-.5598095	-1.818829	1.259019	.8847866
y97	16.16999	15.61432	.5556727	.
y98	11.59325	11.1759	.4173484	.
y99	9.974361	9.513609	.4607515	.
y21	.6654604	.0620121	.6034482	.
y22	-13.79593	-14.65521	.8592823	.
y23	-2.516478	-2.916737	.4002589	.
y24	-7.051568	-6.620246	-.431323	.0445324

b = Consistent under H0 and Ha; obtained from **xtreg**.
B = Inconsistent under Ha, efficient under H0; obtained from **xtreg**.

```
Test of H0: Difference in coefficients not systematic

   chi2(31) = (b-B)'[(V_b-V_B)^(-1)](b-B)
            =  14.64
Prob > chi2 = 0.9944
(V_b-V_B is not positive definite)
```

CONCLUSION

Finally, we concluded that a variety of factors influence total traffic fatalities. We concluded that the speed limit, seat belt, and nightfat have a positive relationship after using various econometric analysis techniques. Furthermore, GDl, statepop has negative relationship. To obtain a more accurate model, we use advanced methods such as fe and re to process panel data. The Hausman test was then performed, and it was determined that the RE treatment would be more appropriate for this model.