Some General Comments on the Group Projects

Spring 2012

Most Important

- I know the projects were a typical MIT fire hose experience
- The grading reflects that

Comment topics

- Writing
- Interpreting Coefficients
- Formatting the paper
- Organization of data analysis paper
- Presenting regression and other statistical results in tables
- Do-files

Writing

- Writing must be logical and free of *all* grammatical and typographical errors
- Write economically.

Interpreting Coefficients

- The most basic interpretation point: $\beta = \frac{\partial Y}{\partial X}$
- But, that's only the beginning
  - Always keep track of units and scale
  - Aids to dealing with different units
    - Rescale everything to the 0-1 interval
    - Run a standardize regression
    - Multiply the coefficient by the standard deviation of X
    - If appropriate, convert everything to logs
- Remember, it is rare for *any* regression coefficient to predict a knock-your-socks-off effect

Formatting the paper

- I am serious about you following a style book/sheet
  - Turabian's style manual
  - American Political Science Association style manual: http://www.ipsonet.org/data/files/APSAStyleManual2006.pdf
  - Always use author-date citation system
  - Put figures and tables at back of paper.
  - Number figures and tables consecutively

Organization of a Data Analysis Paper

- Think inverted pyramid
  - Describe the data
    - Report means, standard deviations, min, and max of all variables
    - Show bivariate scatterplots of important relationships between the dependent variable and independent variables
    - Run regression and report regression coefficients
      - Interpret regression coefficients
      - Discuss goodness-of-fit

Presenting Regression and Other Statistical Results in Tables

- Leaf through the American Political Science Review, American Journal of Political Science, and Journal of Politics to see what tables look like.

Descriptive statistics

- Usually not in journal articles, but in conference papers and in online supplemental materials to journal articles

Variable name*	Mean	s.d.	Min	Max

*Not the name of the variable in the codebook (e.g., V4067a), but a short descriptive name (e.g., number of parties in the legislature).

From Efraim Benmelech, Claude Berrebi, and Esteban F. Klor, "Economic Conditions and the Quality of Suicide Terrorism," *Journal of Politics,* vol 74(1), 2012: 113-128

TABLE 1 **Summary Statistics on the Characteristics of Suicide Terrorists and their Attacks**

	Number of Observations	Mean	Standard Deviation	Median	Minimum	Maximum
A. Characteristics of Suicide Terrorists						
Suicide Terrorist has Academic Education	157	0.197	0.399	0	0	1
Age of Suicide Terrorist	156	21.397	5.727	21	12	64
Suicide Terrorist Previously Involved in Terror	157	0.121	0.327	0	0	1
B. Characteristics of Targeted Locality						
Targeted Locality's Population above 50,000	157	0.745	0.437	1	0	1
Regional Capital	157	0.535	0.500	1	0	1
C. Outcome of Suicide Attack						
Thwarted Suicide Terrorist	157	0.248	0.433	0	0	1
Casualties from Suicide Attack	157	29.535	39.397	10	0	181

Notes: The summary statistics reflect authors' calculations based on Israeli Security Agency reports of suicide terrorists. The targeted cities population refers to their population within the metro area of the city according to the population figures for the year 2003 of the Israeli Central Bureau of Statistics.

Regression results

Coefficient	(1)	(2)
Variable 1	Coeff. (s.e.)	Coeff. (s.e.)
Variable 2	Coeff (s.e.)	
Etc.		
N		
R^2		
RMSE		

Regression results

TABLE 4 The Effect of Economic Conditions on the Outcomes of Suicide Attacks

	Thwarted Suicide Terrorist			Casualties from Suicide Attack		
	(1)	(2)	(3)	(4)	(5)	(6)
Unemployment Rate	-0.710***	0.606**	1.516***	2.551***	0.259	-8.716***
	[0.279]	[0.296]	[0.183]	[0.349]	[0.316]	[0.233]
Palestinian Fatalities	0.006***	0.005***	0.010***	-0.007***	-0.003***	-0.006**
	[0.001]	[0.000]	[0.002]	[0.002]	[0.001]	[0.003]
Days with Curfews			-0.002*			0.004
			[0.001]			[0.004]
Group Specific	-0.385	-0.156*	0.421	1.766***	0.747*	-2.211**
Unemployment Rate	[0.292]	[0.092]	[0.629]	[0.198]	[0.448]	[1.042]
Palestinian Fatalities	0.006***	0.005***	0.010***	-0.007***	-0.003***	-0.003**
	[0.001]	[0.000]	[0.001]	[0.002]	[0.001]	[0.002]
Days with Curfews			-0.002***			0.002
			[0.001]			[0.006]
Income Inequality	-0.516***	-0.620***	-1.206***	0.638***	0.579**	0.942***
	[0.059]	[0.195]	[0.121]	[0.154]	[0.296]	[0.032]
Palestinian Fatalities	0.005***	0.004***	0.007***	-0.005*	-0.001	0.004
	[0.001]	[0.000]	[0.000]	[0.003]	[0.002]	[0.003]
Days with Curfews			-0.001*			-0.002
			[0.001]			[0.005]
Region Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	No	Yes	Yes	No	Yes	Yes
Observations	156	156	79	155	155	79

Notes: Each column in each panel reports the estimated effects of three separate regressions in which the dependent variable is, respectively, whether the suicide terrorist was caught (estimated using a Probit model), or the number of casualties from suicide attacks (estimated using a Poisson model). In addition to the explanatory variables listed on the left, regressions in Columns 1 and 4 add each district average years of education, population size over the age of 15, proportion of males, married, proportion living in a refugee camp and whether the district is in the West Bank as explanatory variables. We subsequently add to these regressions years fixed effects (regressions in Columns 2 and 5) and days with curfews (regressions in Columns 3 and 6). Robust standard errors (adjusted for clustering at the regional level) are in parentheses. *, **, *** denote statistically significance at the 10, 5, and 1% level, respectively.

Do-files

- Carefully document your do-files so that your TA and professor can figure out what you're doing.
- Label each section of the do-file according to what it produces (e.g., figure1.do, table1.do)