1 Content Description

This zip file contains all data and programs necessary to reproduce the tables in "Estimating the Peace Dividend: The Impact of Violence on House Prices in Northern Ireland". Due to the fact that we generate part of the variables with a GAUSS program and use several different datasets there are four different "sets" of programs contained in the zip file:

- most main results in tables 1, 3 and 4 can be reproduced with the STATA data file "maindata.dta" and programme files, "table 1.do", "table 3.do" and "table 4.do" respectively. The do files also contain additional descriptions (section 2)
- the EM algorithm (used to estimate the Markov-Chain parameters (thetas) in the paper) can be run using the programme "pv5generation" in GAUSS. The programme does not only provide estimates of thetas but provides estimates of the conflict probability, the fitted value of killings, the present value of killings at 5 percent and the present value of conflict at 5 percent. **Note: the GAUSS programme provided in the zip file allows for an easy application of the EM-algorithm to different datasets (of violence/events). We have added detailed descriptions in the GAUSS programme itself and automated the process of creating estimates of the most interesting indicators. For details see section 3 below.**
- table 1, column 7 uses a three year rolling average of yearly tourism income to confirm that the effect of the conflict was not restricted to the housing market but was felt in the tourism industry as well. This data can be found in the STATA data file "tourismandkillings.dta" (section 4)
- table 6 uses data on the Isreali/Palestine conflict and data from the Tel Aviv stock market in an application of our methodology. This data is provided in the STATA data file "israelpalestinequart.dta" (section 5)

For additional materials (on the poisson and single index EM estimation, parametric bootstrapping etc.), comments and questions write Hannes Mueller, IAE (CSIC): hannes.mueller (at) iae.csic.es.

2 Main Results

Instructions: load data into STATA and run the respective do file - depending on which table you want to recreate. Note that the do files use the outreg2 program in STATA to create excel files containing the results. Search for outreg2 in STATA to install this routine.

Variable Description: the file "maindata" contains several variables not used in the tables in order to provide the "raw" data that goes into them. Variables are as follows (for a description of the sources see the web appendix of the paper):

- region: variable indicating the region (1=Belfast, 2=North Down, 3= Lisburn, 4=East Antrim, 5=Londonderry/Strabane, 6=Antrim/Ballymena, 7=Coleraine/Limavady, 8=N Coast Enniskillen/Fermanagh/ S Tyrone, 9=Mid Ulster, 10=Mid and South Down, 11=Craigavon/Armagh)
- time: combination of the variables quarter and year (included in the dataset)
- lnhouseprice: ln of the houseprice (in the dataset)
- wtotaldeaths: number of total deaths divided by the standard deviation of that variable
- l1unempl: lag (one quarter) of the ln of the unemployment count (included in the dataset)
- l1wprivatstarts: lag (one quarter) of the weighted number of private housing starts
- l1wtotalpv5: lagged and weighted present value of killings at 5 percent interest rate (Standard Deviation used for weighting is 23.97197)
- l1wtotalpv1: lagged and weighted present value of killings at the 1 percent interest rate
- wl1pv5si: lagged and weighted present value of killings at the 5 percent rate (alternative, single-index, calculation)
- l1belfastbordsiwpv5, l1londondbordsiwpv5, l1midulbordsiwpv5, l1midsdsiwpv5 and l1craigsiwpv5: interaction of dummies (that take the value of one if the region has a common boundary with the region in the name of the variable) and the lagged, weighted present value of killings at 5 percent in the region named in the variable name - the 5 regions are also the 5 most violent regions
- l1wtotalpv5pc: lagged, weighted present value of killings at 5 percent per capita (region population)
- l1wstatepv5: lagged, weighted present value of the "state" at 5 percent
- uklnhouseprice: data that includes the ln of houseprices in 4 UK regions
- ukl1wtotalpv5: data of killings that includes a value of 0 for 4 UK regions

3 EM Algorithm

3.1 To Recreate our Findings

Instructions: The present value of killings is based on EM estimates attained with Hamilton's GAUSS programs that have been modified somewhat to make

them a) compatible with the NI violence data; b) automate the calculation of present values and fitted values; c) include more explanations of how the filter works. In order to run these programs one needs the GAUSS software. The following procedure needs to be followed:

- create a folder in the same location as the GAUSS software (default C:\GAUSS\NI). Move the following files to the folder: pv5generation, SMOOTHSIMPLE and totaldeathsall.txt
- open the file pv5generation with GAUSS, modify all lines that contain the word "MODIFY" - change the path to the folder created in the previous step. If the files have been dropped into the folder C:\GAUSS\NI no modifications are necessary.
- run pv5generation

The programme should create three text files:

- the file "thetas.txt" contains the EM estimated for the eleven regions (ordering is as in table 2 in the article)
- the file "PV5s.txt" contains the weighted present value of killings (5 percent) - this number will deviate slightly from the data used in the STATA file due to rounding errors
- the file "conflpv5s.txt" contains the weighted present value of conflict probability (5 percent) - this number will deviate slightly from the data used in the STATA file due to rounding errors
- the file "otherresults.txt" contains other results: the first column contains the time (regions are ordered 1 to 11), the second column the conflict probability, the third column contains the fitted value of killings ($\hat{y}$), the third column the unweighted present value (5 percent) and the fourth column the unweighted present value of the conflict probability.

The produced text files can be imported easily into STATA using the "import" command.

3.2 To Use the Methodology with New Data

3.2.1 Generic Procedure

The use of the EM methodology should be fairly straightforward with new data. In addition to the orders above one needs to modify the main GAUSS programme, pv5generation. Lines in the programme that need to be modified are marked with "MOD".

1) the violence data needs to be saved in a text file that looks like the "totaldeathsall.txt" file, i.e. all incidents need to be listed in one column sorted first by region then time

2) the number n needs to be adapted to the new dataset, it captures the number of time periods available per region

3) the number of regions needs to be modified depending on how many geographic units the data has

4) the first column for the "otherresults.txt" file needs to be changed to account for the different time structure (this will not affect the running of the programme)

5) clearly the "PV5s.txt" file and the "conflpv5s.txt" file are relatively meaningless when new data is used. This is because these two matrixes use a standard deviation weight. Again, if this is ignored it does not affect the running of the programme.

6) in some cases it can be that the programme gives out an error message because it does not converge. If that happens go to line 57 and modify the starting values. Typically a good starting value for conflict mean will be something slightly below the max value and for peace mean something close to the minimum value. Modify the variances accordingly (the square of the means work well).

3.2.2 Example

The example of the Isreali/Palestinian conflict illustrates how easy it is to use the programme with different datasets. The steps to derive the conflict probability used in the article are:

- Move the file "isrealquart.txt" to the same folder as the main files.
- Modify the two references to "totaldeathsall.txt" to "isrealquart.txt".
- The dataset starts in 2000q4 and has 21 periods in one region. This means we need to set n = 21 and region = 1
- Change the statement in line 186 to seqa(2000.75, 0.25,n)
- Finally, the programme gives an error message if the starting values of the theta vector are not modified (try, for example, $\theta = (300, 10, 0.5, 0.5, 1000, 100)$)

Run the modified file. The second column of the file "otherresults" contains the conflict probabilities used in the paper.

4 Table 1, Column 7

Instructions: load data "tourismandkillings.dta" into STATA

run the following STATA command: "xtreg Pounds L4.yravrgdeath year12 year13 year14 year15 year16 year17 year18 year19, fe"

We run the 4-year lag of the yearly killings because the "Pounds" variable is a three year rolling average of the yearly income in the tourism industry.

5 Table 6

Instructions: load the dataset "israelpalestinequart.dta" into STATA

run the following STATA commands: "regress lnstockmarket confprobtot" and "regress lnstockmarket wtotaldeaths" to recreate columns 1 and 2 of table 6.

Variable Description: the STATA data includes the "raw" data as well as the data used to create table 6. The variables are

- lnstockmarket: the ln of the stockmarket value (see web appendix for details)
- confprobtot: the EM estimate of the probability of conflict estimated with the GAUSS programme discussed under section 3
- wtotaldeaths: weighted number of deaths (weight is standard deviation)