Online Appendices for Brooks et al (2018) "Transition to a Demographic Peace? Why Young Societies are Conflict Prone and Old Societies are the Most Peaceful"

International Security, Vol. 43, No. 3 (Winter 2018/19)

Appendix A

1. Modeling Framework

Our unit of analysis is the (undirected) dyad-year. We restrict our data to "politically relevant" dyads. Following the convention in the literature, this involves considering only the pairs of states that meet at least one of the following criteria: (1) they are connected either by a land or river border or by less than 400 miles of water; or (2) at least one member of the dyad is a major power. Data on contiguity and on major power status were obtained from the Correlates of War project.[1]

We account for temporal dependence among the conflicts themselves—i.e., the fact that the probability of a conflict occurring within a given dyad is likely to be somewhat dependent on the time that has elapsed since the last conflict—using the cubic polynomial method proposed by Carter and Signorino (2010). This involves including a measure of the number of years that have elapsed since the end of the last conflict within each dyad ("peace years") along with its squared and cubed forms. As Carter and Signorino point out, including this cubic polynomial provides a highly flexible yet simple way of accounting for non-linearities in the relationship between time and the probability of conflict.[2]

2. Dependent Variable

Our dependent variable is a dichotomous measure of the onset of (one or more) Militarized Interstate Disputes (MID). This is a broad measure of conflict that includes full-scale wars as well

[1] Data on direct contiguity were obtained from the Correlates of War Project Direct Contiguity Data, Version 3.1. The list of major powers was obtained from the Correlates of War Project State System Membership List, v2011.

[2] Carter and Signorino (2010) suggest this method as a simpler alternative to the use of splines. They demonstrate empirically that in most cases it is as effective as splines in terms of its ability to account for temporal dependence. Moreover, it avoids the problem of separation that commonly arises when estimating logistic regression models in the presence of year dummy variables. See Carter and Signorino (2010).

as less hostile interactions among states that nonetheless involve "an explicit threat, display, or use of force" (Jones, Bremer and Singer, 1996). Data on the presence of MIDs at the dyadic level were obtained from the Dyadic Militarized Interstate Disputes Dataset (Maoz, 2005). As is commonplace in the study of interstate conflicts, we restrict our analysis to only those MIDs that have a score of either 4 ("use of force") or 5 ("war") on the hostility variable used in the Maoz data.

Given that we are interested in testing hypothesis about the factors influencing the onset of conflict—as opposed to the continuation or termination of existing conflicts—our dependent variable is a dichotomous indicator of whether a state initiated a conflict in a given year. We therefore coded the dependent variable as 1 only for the year in which the MID was initiated; subsequent years of the MID—as well as, of course, peace years—are coded as 0.

3. Demographic Variables

Time-series data on total fertility rates were obtained from the World Bank's World Development Indicators. According to the World Bank data, during the 1960-2001 period, total fertility rates for the countries included in this study range from a minimum of 1.09 births per woman (Bulgaria in 1997) to a maximum of 8.66 births per woman (Yemen in 1990).

Life expectancy data were also obtained from the World Bank's *World Development Indicators*. Over the 1960-2001 period, this variable ranges from a minimum of 27 years (Rwanda in 1993) to a maximum of 81 years (Japan in 2001).

The Youth Bulge ratio is defined as the ratio of the population (both male and female) aged between 15 and 24 to the total adult population, i.e.,

$$\frac{Ages\ 15-24}{Ages > 15}$$

Data on the number of people in each age group were obtained from the United Nations Population Division.

Median Age is simply the median age of the population in each year. Over the 1960-2001 period, the UN data on median ages range from a low of 13.9 (Yemen in 1990) to a high of 44.7 (Japan in 2001). For comparison, the median age of the United States in 2010 was 36.9 years.

As is standard practice in the conflict literature, we converted each of these state-level variables into a dyadic form using the weak-link assumption (Oneal and Russett, 1997). In the case of the fertility and youth bulge models, this involves assigning a value to the dyad that corresponds to the higher of the two fertility levels or youth bulge ratios found among the members

of the dyad (as higher values of each of these are thought to increase the probability of conflict). In the case of the life expectancy and median age models, we assigned the lower of the two values.

In order to test for the existence of a possibly non-linear relationship between each of our demographic variables and the probability of conflict, we modeled these relationships using polynomial terms. This involved the inclusion of squared and, in some cases, cubic measures of the fertility and life expectancy measures alongside the lower-order terms. This approach is therefore directly analogous to the treatment that Carter and Signorino (2010) had proposed (discussed above) in the context of capturing temporal dependence in models of conflict. This approach has the advantage of allowing for greater flexibility in terms of the forms that these relationships take without adding unnecessary complexity to the model. However, in order to guard against the possibility of "overfitting" (Beck, King and Zeng, 2000)—i.e., a situation in which the model does a good job of describing idiosyncratic features of the data used to fit the model but fails to generalize beyond that—we employed the four-fold cross-validation protocol suggested by Ward, Greenhill and Bakke (2010). We tested variants of the models in which the relevant demographic variables were specified in linear, quadratic, or cubic forms, and then we chose the model that performed best in terms of its out-of-sample predictive power.

4. Control Variables

Each of our models includes the following battery of control variables that are typically used in studies of interstate conflict:

- *Contiguity* is a dichotomous indicator that is coded as 1 whenever the states in the dyad share a land or river border or are separated by less than 400 miles of water. Data on direct contiguity were obtained from the Correlates of War Project Direct Contiguity Data, Version 3.1.

- *Dist. capital cities* is a measure of the great circle distance (in kilometers) between the capital cities of the two states in the dyad. It provides a more nuanced measure of the extent to which geographical separation deters conflict. These distances were obtained from the database maintained by Kristian Gleditsch.[3]

- *Democracy (low)* is the lower of the Polity 2 democracy/autocracy scores of each member of the dyad. Polity 2 scores range from -10 to +10, representing the most autocratic and

[3]The Gleditsch capital city data was downloaded from http://privatewww.essex.ac.uk/~ksg/data-5.html on February 1, 2012.

democratic regimes, respectively. These scores were obtained from version 2010 of the data released by the Polity IV project (Marshall and Jaggers, 2009).

- *Alliance* indicates whether a formal alliance exists between the members of the dyad. Data on formal alliances were obtained from the Correlates of War Formal Interstate Alliance dataset, version 3.03 (Gibler and Sarkees, 2004). For the purposes of this analysis, the alliance variable is coded as a dummy variable where a value of 1 indicates the presence of a defense pact, neutrality agreement, or entente and a value of 0 indicates the absence of a formal alliance.

- *Bilateral Trade* provides a measure of the extent of bilateral trade within the dyad. It is calculated by taking the log of the total volume of trade (exports plus imports, in millions of US dollars) within the dyad. Data on bilateral trade were obtained from Gleditsch (2002).

- *Capability Ratio* is the log of the ratio of the Composite Index of National Capability (CINC) score of the larger state to that of the smaller state. CINC scores were obtained from the Correlates of War project's National Material Capabilities dataset, version 4.0 (Singer, 1988).

- *GDP per capita (low)* is the lower of the (logged) GDP per capita values of the members of the dyad. More specifically, it is the log of the chain-indexed real GDP per capita converted at Purchasing Power Parity, in 2005 US dollars. Data on this measure were obtained from the Penn World Table, version 7.0 (Heston, Summers and Aten, 2011).[4]

- *Peace years* is the count of the number of years that have elapsed since the last instance of conflict within the dyad.

5. Results

Table 1 shows the results of logit models of conflict that separately include the fertility, life expectancy and youth bulge variables.[5] (Note that the numbers in parentheses below each coefficient estimate represent p-values rather than standard errors.) The cross-validation exercise described above was used to select the non-linear specification of each model that maximizes

[4]This source provides two estimates of GDP for China: an official and an adjusted (unofficial) estimate. We used the unofficial estimate, which is regarded as more accurate.

[5]Due to problems with complete or partial separation, we chose to drop the *Peace Years (cubed)* term from the models of youth bulge and median age.

its out-of-sample predictive power. In the case of the fertility model, the model with the greatest predictive value turned out to be the one with a quadratic term. In the case of the other four dependent variables, the models that included the more complex cubic forms of the polynomial turned out to have the highest out-of-sample predictive power.

Table 1: Results of conflict models. The numbers in parentheses represent p-values, not standard errors.

	Fertility	Life Exp.	Youth Bulge	Median Age
Intercept	-7.30	-4.94	-27.91	-9.54
	(0.00)	(0.44)	(0.00)	(0.09)
Contiguity	1.86	1.71	1.58	1.57
	(0.00)	(0.00)	(0.00)	(0.00)
Dist. capital cities	-0.00	-0.00	-0.00	-0.00
	(0.94)	(0.85)	(0.55)	(0.50)
Democracy (low)	-0.00	-0.01	0.00	0.01
	(0.98)	(0.43)	(0.80)	(0.60)
Alliance	-0.11	-0.04	-0.09	-0.10
	(0.51)	(0.81)	(0.61)	(0.56)
Bilateral Trade	0.03	-0.02	0.00	0.01
	(0.52)	(0.65)	(0.91)	(0.82)
Capability Ratio	-0.20	-0.21	-0.20	-0.20
	(0.00)	(0.00)	(0.00)	(0.00)
GDP per capita (low)	0.18	0.05	0.16	0.17
	(0.08)	(0.61)	(0.14)	(0.13)

(table continues on next page)

Table 1: (continued from the previous page)

	Fertility	Life Exp.	Youth Bulge	Median Age
Fertility	1.04			
	(0.00)			
Fertility (squared)	-0.09			
	(0.00)			
Peace years	-0.19	-0.18	-0.09	-0.09
	(0.00)	(0.00)	(0.00)	(0.00)
Peace years (squared)	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Peace years (cubed)	-0.00	-0.00		
	(0.00)	(0.00)		
Life Exp.		-0.10		
		(0.77)		
Life Exp. (squared)		0.01		
		(0.45)		
Life Exp. (cubed)		-0.00		
		(0.25)		
Youth Bulge			213.03	
			(0.01)	
Youth Bulge (squared)			-636.83	
			(0.01)	

(table continues on next page)

Table 1: (continued from the previous page)

	Fertility	Life Exp.	Youth Bulge	Median Age
Youth Bulge (cubed)			628.42	
			(0.02)	
Median Age				0.57
				(0.45)
Median Age (squared)				-0.02
				(0.56)
Median Age (cubed)				0.00
				(0.78)
N	34118	34092	36011	36011
AIC	3997.1	3967	4559.69	4547.51

Appendix B

The following analysis is based on one experiment embedded within a larger set of experiments conducted in April 2009 by Deborah Jordan Brooks and Benjamin Valentino using YouGov/Polling Point for sample, data collection, and weighting the data back to population percentages. Within their experiment (in which the stakes varied between a conflict in two fictional countries over oil versus humanitarian goals) we included age as an independent variable. For more information on the sampling, methodology, treatment, and question wording, see Brooks and Valentino (2011).

The table below shows the relevant findings. The dependent variable is measured on a continuous scale, with "1" representing "strongly disapprove" and "6" representing "strongly approve" of going to war in the described situation.

	MODEL A: Age - continuous *(no controls)*	MODEL B: Age - continuous *(with controls^)*	MODEL C: Age Dummy Variables *(with controls^)*			
Constant	3.712 ***	3.764 ***	2.555 ***	2.815 ***	3.305 ***	3.187 ***
Condition (Oil)	-.402	-.414	-.249	.136	-.390	-.180
Age (continuous)	-.014 **	-.018 ***				
Condition Oil x Age (contin)	.007	.007				
Age - **18-29** years			.632 *	.373	-.117	*(excluded)*
Age - **30-44** years			.750 **	.490 **	*(excluded)*	.117
Age - **45-64** years			.260	*(excluded)*	-.490 **	-.373
Age - **65+** years			*(excluded)*	-.260	-.750 **	-.632 *
Condition Oil x Age 18-29			.069	-.316	.211	*(excluded)*
Condition Oil x Age 30-44			-.142	-.526	*(excluded)*	-.221
Condition Oil x Age 45-64			.385	*(excluded)*	.526	.316
Condition Oil x Age 65+			*(excluded)*	-.385	.142	-.069

**** p ≤ 01; ** p ≤ .05; * p ≤ .10.*

N = 604 in Model A, and 598 in Models B and C (N=280 in the oil condition, and N=318 in the humanitarian condition with missing values excluded; missing = < 1%)

^ Models B and C include controls for: Parent, Parent of military-aged son, Woman, Married, Race (Black), Ethnicity (Hispanic), College (at least some), Republican/ Republican leaners, Independent (pure), Political interest (high), Region (South), Frequent church-goer (few times month+), Income (median imputed), plus war for oil (versus humanitarian war.) Model A uses no controls.

Model A treats respondent age as a continuous variable. The significant coefficient for age shows that older respondents are significantly less likely to support war than younger respondents. Additionally, the lack of significance for the interaction term (Condition: Oil x Age) suggests that age decreases support for war across both types of conflicts. Model A lacks control variables, whereas Model B includes an array of potentially-relevant controls (see the note below the Table 2 for the controls used.) The results of Model A and Model B are similar, and lead to the same conclusions.

Given our particular interest in the lack of bellicosity among retirement-age individuals (age 65+) as compared to their much-younger counterparts, there are reasons to believe that age might be better handled as a categorical variable for these analyses. The four columns in Model C take this approach, making it possible to see whether different age groups are distinct from one another in their preferences. Without a significant or directionally consistent difference in preferences between 18-29 year-olds versus 30-44 year-olds, there appears to be some non-linearity in play; however, the general pattern is still one in which those in the oldest

age group support war less than their substantially younger counterparts. And again, the lack of significance for any of the age and war stakes interaction terms indicates that the general relationship between age and support for war is not statistically distinct across different kinds of wars.

To be clear, we cannot rule out generational effects with this experiment; however, the many non-experimental studies of Americans conducted across the decades do not lead us to expect a difference on that front. Additionally, this result could vary by country. Survey experiments along these lines beyond the U.S. on such question would be beneficial; however, again, we also do not have theoretical reasons to expect cross-national differences on that front.

References

Beck, Nathaniel, Gary King and Langche Zeng. 2000. "Improving Quantitative Studies of International Conflict: A Conjecture." *American Political Science Review* 94(1):21–35.

Brooks, Deborah Jordan and Benjamin A Valentino. 2011. "A war of one's own: Understanding the gender gap in support for war." *Public Opinion Quarterly* 75(2):270–286.

Carter, David B. and Curtis S. Signorino. 2010. "Back to the Future: Modeling Time Dependence in Binary Data." *Political Analysis* 18(3):271–292.

Gibler, Douglas M. and Meredith Reid Sarkees. 2004. "Measuring Alliances: The Correlates of War Formal Interstate Alliance Dataset, 1816–2000." *Journal of Peace Research* 41(2):211–222.

Gleditsch, Kristian Skrede. 2002. "Expanded Trade and GDP data." *Journal of Conflict Resolution* 46(5):712.

Heston, Alan, Robert Summers and Bettina Aten. 2011. "Penn World Table Version 7.0." *Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania* .

Jones, Daniel M., Stuart A. Bremer and J. David Singer. 1996. "Militarized Interstate Disputes, 1816–1992: Rationale, Coding Rules, and Empirical Patterns." *Conflict Management and Peace Science* 15(2):163–213.

Maoz, Zeev. 2005. "Dyadic MID Dataset, version 2.0." *Downloaded from http://psfaculty.ucdavis.edu/zmaoz/dyadmid.html* .

Marshall, Monty G. and Keith Jaggers. 2009. "Political Regime Characteristics and Transitions 1800–2009." *Polity IV Project* .

Oneal, J.R. and B.M. Russett. 1997. "The classical liberals were right: Democracy, interdependence, and conflict, 1950-1985." *International Studies Quarterly* 41(2):267–293.

Singer, J.D. 1988. "Reconstructing the correlates of war dataset on material capabilities of states, 1816–1985." *International Interactions* 14(2):115–132.

Ward, Michael D., Brian D. Greenhill and Kristin M. Bakke. 2010. "The Perils of Policy by p-Value: Predicting Civil Conflicts." *Journal of Peace Research* 47(4):363–375.