Supplemental Methodological Appendix for

US Infant Mortality and the President's Party

Alternative detrending methods

The statistical approach we took in "US Infant Mortality and the President's Party" was to first detrend the log of the infant mortality rate using median cubic b-splines and then to regress the residuals of the estimated trend on an indicator for the political party of the president. It is natural to wonder whether our results are sensitive to the specific de-trending method we used. Here we show that our results are largely robust to alternative de-trending methods.

One common approach to de-trending is to fit third-, fourth-, fifth-order or higher order linear polynomials. These polynomials use integer power transformations of time as covariates. This traditional fitting method is just a special case of fractional polynomial regression, which offers more flexibility than do linear polynomials.[1] The results we report include results based on both third-, fourth-, and fifth-order linear and best fitting fractional polynomials.[1-3]

Cubic splines represent an alternative, semi-parametric means for removing non-linear trends in time series data.[4-9] Cubic splines represent linear combinations of cubic functions linked at "knots." The cubic functions are constrained in such a way that the resulting curve and its first and second derivatives are continuous at the knots. Typically researchers pre-specify the number of knots, and locate these knots at equal intervals.[10] It is possible that our results would have differed had we used a different number of knots. We therefore experimented with alternative 3- and 5-knot splines (as alternatives for using 4 knots, the originally selected method for detrending). By convention we used equidistant knots,[10] located at minimum, maximum, and median of time range for the 3-knot spline, and at minimum, maximum, and 25, 50, and 75 centiles of time range for the 5-knot spline.

Whether one uses linear or fractional polynomials or cubic splines it is possible to overfit the data. Indeed, polynomials of sufficient order or cubic splines with a sufficient number of kinks will always perfectly fit a finite time series. A common strategy to guard against this possibility is to choose the order or number of knots that maximizes the Akaike Information Criterion (AIC).[10] As it turns out, in our case, the AIC is maximized using forth order linear or fractional polynomials or a spline with four knots.

Table 1 reports results using these various detrending methods. In each case, we first detrend ln(IMR) using either simple polynomials, fractional polynomials or cubic splines. We then regress residuals from trend on a presidential party indicator. Along with point estimates, we report 95% confidence intervals and exact P values for the two-tailed test against the null hypothesis that the coefficient is 0. We also report the AIC for the estimated trend. As the reader can see, estimates are largely insensitive to the detrending method used. The one partial exception occurs when we detrend using a 5-knots cubic b-spline. In this case, the estimated coefficients drop from roughly 0.03 to 0.01. However, the value of the AIC suggests that in this case the cubic spline may be overfitting the data.

Detrending the Presidential Party Indicator

The results we report in the text are based on results that detrend the logarithm of infant mortality but not the presidential party indicator. We chose this approach both because it is standard in political science and it maintains a consistency between the figures and regressions we report. That said, a common alternative would have been to detrend all variables. In column 2 of Table 2 we report results when the presidential party indicator is also detrended. Detrending the presidential party indicator actually serves to increase a bit the magnitude of the coefficients on the presidential party indicators.

Table 1. Parameter estimates for "Republican president" using alternative detrending methods (1965-2010)

Dependent Variable		*Cubic b-splines*			*Linear polynomials*			*Best-fitting fractional polynomials*		
		4-knots	3-knots	5-knots	3^{rd} Order	4^{th} Order	5^{th} Order	3^{rd} Order	4^{th} Order	5^{th} Order
Overall IMR	Coeff.	.031	.032	.012	.032	.031	.032	.031	.031	.032
	95% CI	±.016	±.015	±.011	±.015	±.016	±.015	±.016	±.016	±.015
	p value	[.000]	[.000]	[.041]	[.000]	[.000]	[.000]	[.000]	[.001]	[.000]
Whites IMR	Coeff.	.029	.031	.010	.031	.030	.031	.030	.029	.031
	95% CI	±.013	±.013	±.009	±.013	±.014	±.013	±.014	±.014	±.013
	p value	[.000]	[.000]	[.034]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]
Blacks IMR	Coeff.	.031	.031	.010	.031	.031	.031	.030	.032	.031
	95% CI	±.022	±.021	±.015	±.023	±.022	±.023	±.020	±.021	±.023
	p value	[.007]	[.006]	[.181]	[.010]	[.008]	[.011]	[.005]	[.004]	[.010]
	Trend AIC	−203.5	−203.9	−238.0	−207.1	−205.8	−228.5	−207.4	−206.1	−207.8
	Observations	46	46	46	46	46	46	46	46	46

Note: Reported p values and confidence intervals were computed using the Newey-West estimator. All reported estimates are for non-detrended "Republican president" (=1). The error structure is assumed heteroskedastic and autocorrelated up to a maximum lag length of three years. The p values and 95% confidence intervals were adjusted using fixed-bandwidth asymptotics. The p values represent the two-tailed test against the null hypothesis that the coefficient is equal to "0". The first column reports our baseline model estimates (i.e., estimates reported in the first three columns of the first row in Table 2 of our paper). Please see text for description of detrending methods.

Table 2. Parameter estimates for different model specifications

VARIABLES		1965-2010	1965-2010	1965-2010	1965-2010	1965-2008	1965-2008	1965-2008	1965-2008
		Overall IMR							
Republican	Coeff.	.031		.027		.030		.030	
	95% CI	±.016		±.013		±.017		±.016	
	p value	[.000]		[.000]		[.001]		[.001]	
Det-Republican	Coeff.		.037		.033		.037		.036
	95% CI		±.017		±.013		±.018		±.017
	p value		[.000]		[.000]		[.000]		[.001]
Det-Birth Rate	Coeff.			.011	.009				
	95% CI			±.013	±.013				
	p value			[.100]	[.171]				
Det-Abortion Ratio	Coeff.							−.015	−.014
	95% CI							±.048	±.037
	p value							[.512]	[.439]
		Whites IMR							
Republican	Coeff.	.029		.026		.029		.029	
	95% CI	±.013		±.013		±.014		±.014	
	p value	[.000]		[.000]		[.000]		[.000]	
Det-Republican	Coeff.		.035		.032		.036		.035
	95% CI		±.013		±.013		±.014		±.014
	p value		[.000]		[.000]		[.000]		[.000]
Det-Birth Rate	Coeff.			.008	.007				
	95% CI			±.010	±.009				
	p value			[.089]	[.159]				
Det-Abortion Ratio	Coeff.							−.025	−.024
	95% CI							±.044	±.032
	p value							[.249]	[.136]
		Blacks IMR							
Republican	Coeff.	.031		.025		.029		.029	
	95% CI	±.022		±.016		±.023		±.021	
	p value	[.007]		[.004]		[.014]		[.011]	
Det-Republican	Coeff.		.038		.031		.036		.035
	95% CI		±.023		±.015		±.024		±.023
	p value		[.002]		[.000]		[.005]		[.004]
Det-Birth Rate	Coeff.			.017	.015				
	95% CI			±.015	±.015				
	p value			[.027]	[.050]				
Det-Abortion Ratio	Coeff.							−.028	−.026
	95% CI							±.053	±.049
	p value							[.285]	[.269]
Observations		46	46	46	46	44	44	44	44

Note: Reported p values and confidence intervals were computed using the Newey-West estimator. The error structure is assumed heteroskedastic and autocorrelated up to a maximum lag length of three years. The p values and 95% confidence intervals were adjusted using fixed-bandwidth asymptotics. First row is for "Republican president," while the second is for *detrended* "Republican president." The p values represent the two-tailed test against the null hypothesis that the coefficient is equal to "0". Variables were detrended using a 4-knot cubic b-spline (i.e., the method used in our paper). Data for birth rates are from the National Vital Statistics Reports (defined as number of births per 100,000 population of the reference group, 1965-2010). The data for abortion ratios were compiled by Wm. Robert Johnston from various sources, including the Centers for Disease Control and Prevention and the Government Statistical Service (defined as number of reported abortions obtained within the US by residents per 100,000 live births, 1965-2008).

Selective fertility as an alternative explanation for the association between the party affiliation of the president and infant mortality

As we noted in the text, the party of the president could be associated with infant mortality either because of selective fertility or because the same infant might be more likely to survive when the president is a democrat. With this in mind, we tried entering detrended birth and abortion rates into our second stage regressions. As columns 3 and 4, and 7 and 8 of Table 2 show, doing so has very little effect on the estimated coefficient of the presidential party indicator. Given the crudeness of the variables, we cannot completely rule out the possibility that selective fertility might explain our results, but these results do not lend much positive support to this explanation.

Fourier decomposition of IMR trends

As we point out in the text, detrending removes most of the variation from our variables of interest, especially infant mortality. Here we pursue this issue more thoroughly. A finite Fourier transform[11] can be used to generate a frequency-domain-based decomposition of ln(IMR) trends between 1965 and 2010 into approximately orthogonal components that vary at different frequencies. To accomplish the decomposition we fit ln(IMR):

$$LnIMR = \sum_{k=0}^{K}\left[\beta_k \cos\left(2\pi \frac{k(t-1)}{T}\right) + \alpha_k \sin\left(2\pi \frac{k(t-1)}{T}\right)\right] \quad (1)$$

where K=23 (for k values between “0” and “22”), T denotes the total number of years in the series, the ratio $2\pi k/T$ denotes the frequency (in radians), and β and α are the “Fourier coefficients”. We have a total of T parameters, thus generating a perfect fit to the data. In essence, for each k there is a cosine and a sine component, with K representing the total number of orthogonal components (i.e., frequencies) in which ln(IMR) is decomposed.[12]

By decomposing the trends of our variables of interest we estimate how much of the variation is accounted for at each frequency level. Table 3 reports the results of such decompositions for both ln(IMR) and the presidential party indicator. We present results for the original variables and for the variables once the trend has been removed using a cubic spline with 4 knots.

Table 3 shows that, in the case of ln(IMR), most of the variation is accounted for by low level frequencies, with most of the action happening at the lowest frequency level, as expected. However, about 5% of the variation occurs at the 3rd lowest frequency representing periods of about 16 years (46/3). In contrast, more than 50% of the variation in the presidential party indicator occurs at this frequency. The last two columns of Table 3 report what happens when both series are detrended. Detrending has little impact on the variance components for the presidential party indicator, but once the trend has been removed a majority of the variation in the ln(IMR) occurs at the third lowest frequency. As expected, the de-trending soaks up the low frequency variation. This table provides a good illustration of the variation in the data. After the low frequency variation is removed through detrending there remains significant variation in the ln(IMR) series with most of this occurring at the third lowest frequency which corresponds to a period of just under 16 years. What the figures together with the regressions show is that the medium level frequency in the infant mortality series lines up with the medium level frequency in the presidential party indicator variable. This need not have been true. It is this feature in the data that is responsible for our reported estimates of the association between the presidential party indicator and detrended ln(IMR).

Table 3. Fourier decomposition and variance accounted for by each frequency component

			Component's contribution			
			Original data		Detrended data	
			LnIMR	President's Party (Rep=1)	LnIMR	President's Party (Rep=1)
Frequency Component (*K*)	Frequency (1/years)	Period (years)	R^2	R^2	R^2	R^2
1	.02	46.0	.644	.042	.001	.000
2	.04	23.0	.148	.072	.194	.146
3	.07	15.3	.050	.551	.596	.454
4	.09	11.5	.036	.097	.011	.048
5	.11	9.2	.022	.024	.012	.052
6	.13	7.7	.017	.025	.007	.025
7	.15	6.6	.013	.038	.023	.030
8	.17	5.8	.010	.005	.001	.006
9	.20	5.1	.007	.053	.006	.078
10	.22	4.6	.008	.005	.051	.016
11	.24	4.2	.006	.004	.005	.007
12	.26	3.8	.005	.003	.001	.006
13	.28	3.5	.005	.003	.011	.001
14	.30	3.3	.005	.026	.031	.027
15	.33	3.1	.004	.002	.003	.006
16	.35	2.9	.004	.010	.004	.021
17	.37	2.7	.003	.005	.006	.008
18	.39	2.6	.003	.003	.006	.005
19	.41	2.4	.003	.008	.005	.017
20	.43	2.3	.003	.024	.006	.038
21	.46	2.2	.003	.001	.014	.005
22	.48	2.1	.003	.000	.004	.001
		Observations	46	46	46	46
		R^2	1.0	1.0	1.0	1.0

Note: In the context of linear regression, when *k*=0, sin(0)=0, and thus the term is dropped; the term cos(0) [cos(0)=1] would pick up the value of the constant, but since the model does not include a constant, the term is not included in the model specification. Accordingly, the table includes *K*=22 total frequency components. "Component's contribution" is the variance accounted for by each cosine-sine pair of components that belong to a same frequency level. Because the fit is perfect (R^2=1.0) standard errors are equal to "0". Components denoted as "1" and "22," for example, denote the cosine-sine pair components for *k*=1 and *k*=22 (i.e., lowest and a highest frequency components, respectively). The data are for 1965-2010.

References

1. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. *International journal of epidemiology* 1999;**28**(5): 964-74.

2. Royston P, Altman DG. Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. *Applied Statistics* 1994: 429-67.

3. Royston P, Sauerbrei W. *Multivariable model-building: a pragmatic approach to regression anaylsis based on fractional polynomials for modelling continuous variables*: Wiley.com; 2008.

4. Newson RB. B-splines and Splines Parameterized by Their Values at Reference Points on the X-axis. *Stata Technical Bulletin* 2001;**10**(57).

5. Schoenberg IJ. *Approximations with special emphasis on spline functions*: Academic Press New York; 1969.

6. Ziegler Z. One–sided L1–approximation by splines of an arbitrary degree. *Schoenberg IJ (ed)* 1969.

7. Smith PL. Splines as a useful and convenient statistical tool. *The American Statistician* 1979;**33**(2): 57-62.

8. Wegman EJ, Wright IW. Splines in statistics. *Journal of the American Statistical Association* 1983;**78**(382): 351-65.

9. Durrleman S, Simon R. Flexible regression models with cubic splines. *Statistics in medicine* 1989;**8**(5): 551-61.

10. Harrell FE. *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis*: Springer; 2001.

11. Baker M, Benjamin D, Stanger S. The highs and lows of the minimum wage effect: A time-series cross-section study of the Canadian law. *Journal of Labor Economics* 1999;**17**(2): 318-50.

12. Shumway RH, Stoffer DS. *Time series analysis and its applications: with R examples*: Springer; 2011.