

The Effects of Economic Variables on State Political Preferences

Introduction

This project is based on data I obtained via Paul Gronke's US Congress class and some Economic data give to me by Jeff Parker. It examines how state wide political polarization (measured using the mean DW nominate scores of the states congressmen) is effected by a variety of state level economic variables. The data starts in 1974 and goes to 2013, while the political data is mostly complete, but many early data points are missing or incomplete in the economic data.

Data

This project uses a dataset of my own creation generated by averaging the year-by-year ideological preferences of legislators from each state. This data was obtained from DW nominate (it can be downloaded here¹) and has two ideological dimensions, though for the purpose of this report I only use 1. These dimensions extend to -1 for very liberal and 1 for very conservative, though the data I obtained has some observations less than -1 for the second dimension, they could be extremely liberal delegates or potentially just mistakes within the data.

Examining the relationship between state policy and political preference

Since tax rates are in percentages (0 to 100) I created a variable that was first dimension DW nominate score x 100 (dwnom1_100) in order to more easily interpret regression coefficients.

My initial hypothesis for these variables was the more liberal states would have higher tax rates and therefore negative coefficients. Whether this is causation

¹ <http://voteview.com/downloads.asp>

or reverse causation is totally speculative, but the results proved more complicated than this.

```
. reg dwnoml_100 salesrate cigrate gasrate topindmtr rmw
```

Source	SS	df	MS			
Model	83006.4999	5	16601.3	Number of obs =	1702	
Residual	1159999.11	1696	683.961737	F(5, 1696) =	24.27	
Total	1243005.61	1701	730.749915	Prob > F =	0.0000	
				R-squared =	0.0668	
				Adj R-squared =	0.0640	
				Root MSE =	26.153	

dwnoml_100	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
salesrate	-2.040561	.3712812	-5.50	0.000	-2.768779	-1.312343
cigrate	.053351	.0134383	3.97	0.000	.0269936	.0797083
gasrate	.5583716	.1068168	5.23	0.000	.348865	.7678781
topindmtr	-.8685115	.1825208	-4.76	0.000	-1.226501	-.5105219
rmw	-1.274934	.260853	-4.89	0.000	-1.786562	-.7633069
_cons	12.85084	2.676809	4.80	0.000	7.600646	18.10104

This regression shows that while sales tax rate, top individual marginal income tax rate and state minimum wage correspond to more liberal delegations, cigarette tax rate and gas tax rate correspond to more conservative delegations. Cigarette tax rate and gas tax rate are particularly perplexing since cracking down on cigarette use is considered a liberal issue, and conservatives tend to be friendlier to big oil.

I then hypothesized that the gas tax rate may be related to urbanization since gas taxes are more lucrative when people live far away from each other/ have to drive everywhere and I ran this regression.

```
reg dwnoml_100 salesrate cigrate gasrate topindmtr urbanization rmw
```

Source	SS	df	MS	Number of obs = 1026		
Model	80799.8964	6	13466.6494	F(6, 1019) = 17.65		
Residual	777537.169	1019	763.03942	Prob > F = 0.0000		
-----				R-squared = 0.0941		
Total	858337.066	1025	837.402015	Adj R-squared = 0.0888		
-----				Root MSE = 27.623		
dwnoml_100	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
salesrate	-2.861612	.4964908	-5.76	0.000	-3.835873	-1.887351
cigrate	.048442	.0177775	2.72	0.007	.0135572	.0833267
gasrate	.284399	.1576741	1.80	0.072	-.025004	.5938019
topindmtr	-1.25236	.285632	-4.38	0.000	-1.812854	-.6918655
urbanization	.3788639	.0634392	5.97	0.000	.2543775	.5033503
rmw	-1.11697	.3898942	-2.86	0.004	-1.882057	-.3518827
_cons	-.7828085	5.723965	-0.14	0.891	-12.01491	10.4493

As shown in this regression urbanization does in fact render gas tax rate insignificant. However, it is also appears that urbanization makes a congressional delegation more conservative, and while the effect is small, it is significant and it significantly increases the R-squared of the regression (though it is still extremely small). While it is clear that urbanization does have an effect that will have to be looked into further, it is not the end all be all of this regression that I anticipated.

Urbanization in relation to a majority conservative delegation is counter-intuitive but not absurd. Some Republican bulwark states do have large cities in them, such as Texas, and the concentration of liberals in urban centers makes it easy for Republican state legislators to gerrymander around them. Clearly urbanization is not an entirely a liberal force as some people (including myself) tend to think.

Political Ideology and State GDP

Since State GDP has a mean in the 10,000s (of millions) I took the log of it to bring it to scale with dwnom1 using lstate_gdp and lstate_taxrev.

```
reg dwnoml_100 salesrate cigrate gasrate topindmtr rmw urbanization lstate_ta
> xrev lstate_gdp
```

Source	SS	df	MS	Number of obs =	1026
Model	106347.765	8	13293.4707	F(8, 1017) =	17.98
Residual	751989.3	1017	739.419174	Prob > F =	0.0000
Total	858337.066	1025	837.402015	R-squared =	0.1239
				Adj R-squared =	0.1170
				Root MSE =	27.192

dwnoml_100	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
salesrate	-1.529031	.5612253	-2.72	0.007	-2.630323 - .4277392
cigrate	.0677883	.0180289	3.76	0.000	.0324101 .1031664
gasrate	.1837824	.1566167	1.17	0.241	-.1235464 .4911113
topindmtr	-.9453697	.3207891	-2.95	0.003	-1.574854 -.3158855
rmw	-1.199264	.3862761	-3.10	0.002	-1.957253 -.4412742
urbanization	.5538768	.0754925	7.34	0.000	.405738 .7020156
lstate_tax~v	-13.83285	5.424115	-2.55	0.011	-24.47659 -3.189113
lstate_gdp	7.832086	5.293463	1.48	0.139	-2.555272 18.21945
_cons	9.609963	17.53797	0.55	0.584	-24.80478 44.0247

Gas tax rate continues to spiral towards insignificance. Urbanization perplexingly becomes even more significant, and somewhat perversely log of state tax revenue produces a significant liberal coefficient while log of gdp has an almost significant conservative coefficient. Since state revenue and GDP are inherently very collinear, I decided that trying to make sense of these effects separately was too tricky and that it would be more prudent to throw lstate_taxrev out. However, rerunning the regression without lstate_taxrev we obtain this.

```
reg dwnoml_100 salesrate cigrate gasrate topindmtr rmw urbanization lstate_gd
> p
```

Source	SS	df	MS	Number of obs =	1026
Model				F(7, 1018) =	19.51

Model		101538.749	7	14505.5355	Prob > F	=	0.0000
Residual		756798.317	1018	743.416814	R-squared	=	0.1183

Total		858337.066	1025	837.402015	Adj R-squared	=	0.1122
					Root MSE	=	27.266

dwnoml_100		Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
salesrate		-2.14103	.5087012	-4.21	0.000	-3.139253 -1.142807
cigrate		.057667	.0176342	3.27	0.001	.0230635 .0922705
gasrate		.2303197	.1559699	1.48	0.140	-.0757395 .536379
topindmtr		-1.337061	.2823911	-4.73	0.000	-1.891197 -.7829263
rmw		-1.278553	.3860622	-3.31	0.001	-2.036122 -.5209845
urbanization		.5903482	.0743256	7.94	0.000	.4444992 .7361971
lstate_gdp		-5.413483	1.024945	-5.28	0.000	-7.42473 -3.402236
_cons		45.59776	10.44188	4.37	0.000	25.10769 66.08783

This paints a more conventional picture of the relationship between political ideology (as viewed through the DW nominate scale) and GDP. This reinforces the narrative that liberal states tend to be more economically developed (eg. liberal states contribute more to the federal government but conservative states receive more aid).

Political Ideology and Demographics

Running this regression with the inclusion of demographic data such as level of educational attainment, (ed_coll) showed a significant effect in terms of making states more liberal. Union membership (tot_permem) also correlates with more liberal delegations while percentage of public sector workers (pub_per) is insignificant but also appears to have a liberal effect.

```
. reg dwnoml_100 salesrate cigrate gasrate topindmtr rmw urbanization lstate_gdp
ed_coll tot_permem pub_per urate
```

Source		SS	df	MS	Number of obs =	1026	

Model		129358.984	11	11759.9077	F(11, 1014) =	16.36	
Residual		728978.081	1014	718.913295	Prob > F	= 0.0000	
					R-squared	= 0.1507	

```
-----+-----
Total | 858337.066 1025 837.402015
Adj R-squared = 0.1415
Root MSE = 26.813
```

```
-----+-----
dwnoml_100 | Coef. Std. Err. t P>|t| [95% Conf. Interval]
-----+-----
salesrate | -2.519731 .5393412 -4.67 0.000 -3.578084 -1.461378
cigrate | .1135461 .020021 5.67 0.000 .0742588 .1528334
gasrate | .2908168 .1568611 1.85 0.064 -.0169928 .5986264
topindmtr | -1.18666 .2810459 -4.22 0.000 -1.738158 -.6351619
rmw | -.747729 .4237818 -1.76 0.078 -1.579319 .0838608
urbanization | .726255 .0816321 8.90 0.000 .5660678 .8864422
lstate_gdp | -3.932578 1.134833 -3.47 0.001 -6.159468 -1.705688
ed_coll | -1.078714 .2407723 -4.48 0.000 -1.551183 -.6062451
tot_permem | -.3628322 .1821173 -1.99 0.047 -.7202022 -.0054623
pub_per | -.0497574 .2581198 -0.19 0.847 -.5562676 .4567527
urate | -2.668468 .5290497 -5.04 0.000 -3.706625 -1.63031
_cons | 58.14652 14.0157 4.15 0.000 30.64343 85.64962
-----+-----
```

Unemployment rates' (urate) negative coefficient suggests that states with liberal delegations tend to have higher levels of unemployment. This effect is interesting, but may be complicated by co-linearity with state GDP. Running the regression without GDP:

```
. reg dwnoml_100 salesrate cigrate gasrate topindmtr rmw urbanization ed_coll
tot
> _permem pub_per urate
```

```
-----+-----
Source | SS df MS
-----+-----
Model | 120725.876 10 12072.5876
Residual | 737611.19 1015 726.710532
-----+-----
Total | 858337.066 1025 837.402015
Number of obs = 1026
F( 10, 1015) = 16.61
Prob > F = 0.0000
R-squared = 0.1407
Adj R-squared = 0.1322
Root MSE = 26.958
-----+-----
```

dwnom1_100	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
salesrate	-2.775289	.537165	-5.17	0.000	-3.82937	-1.721208
cigrate	.1156347	.0201201	5.75	0.000	.0761529	.1551165
gasrate	.3605057	.156408	2.30	0.021	.0535856	.6674258
topindmtr	-1.159295	.2824543	-4.10	0.000	-1.713556	-.6050338
rmw	-.5732206	.4230549	-1.35	0.176	-1.403383	.2569416
urbanization	.6007298	.0735509	8.17	0.000	.4564006	.7450589
ed_coll	-1.250674	.2368774	-5.28	0.000	-1.715499	-.7858484
tot_permem	-.3548141	.1830875	-1.94	0.053	-.7140874	.0044592
pub_per	.306234	.2380771	1.29	0.199	-.1609455	.7734136
urate	-3.114346	.5159402	-6.04	0.000	-4.126777	-2.101914
_cons	20.10977	8.762611	2.29	0.022	2.914869	37.30468

We see that unemployment rate is negatively correlated with DW nominate score when we do not control for state GDP. In other words, liberal states have higher employment rate, but conservative states have a better ratio of employment to GDP, ie they generate more jobs per amount of wealth.

Significance Tests:

```
. hettest
```

```
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
```

```
Ho: Constant variance
```

```
Variables: fitted values of dwnom1_100
```

```
chi2(1) = 1.27
```

```
Prob > chi2 = 0.2596
```

The heteroskedasticity test does not reject the null that our results are homoscedastic, which is extremely important given that a lot of things I cannot account for change between 1974 and 2013 and these do not seem to effect the variance towards the beginning or end of the sample period more or less.

Problems (Things I would tackle is I had unlimited time):

Endogeneity: With data this messy and inter-related, some amount of endogeneity is probably inevitable. However, I did not have a good intuition for which variables

are most likely to be endogenous and as such could not really test every variable for potential endogeneity.

Differences between the special and time driven trends within the data set: Using panel data I could not help but wonder which regression coefficients were being driven by differences between states and which were being driven by difference between states over time. Unfortunately it was hard enough putting the data set together, and breaking it up into 38 pieces and examining each one would have left me with very few degrees of freedom.

Reverse Causation: Does political preference drive demographics and wealth of a state or vice versa? I do not even know how to begin with this one but it is certainly an interesting problem.

Conclusions:

While the highest R^2 I obtained for any of my regressions is .15, the data pretty clearly shows that more liberal states tend to adopt policies like higher income and sales tax (Oregon is a significant outlier there) and a higher minimum wage. The data also show that states with higher levels of urbanization are more conservative, which will have to be looked into more closely, and that people living in states with liberal delegations tend to be more educated, more unionized, and more likely to be public sector workers than those who live in states with conservative delegations.

Undoubtedly not all of these correlations are statistically rigorous, but as a whole they do show that economic variables can drive political change and vice versa. As a Political Science major, it gives me hope that not all of my attempts at statistical work are in vain.

Bonus Regression:

```
reg state_gdp dwnom1_100 is_jeff_parker
```

Source		SS	df	MS	Number of obs =	1750
--------	--	----	----	----	-----------------	------

-----+-----				F(1, 1748) =	1.67
Model		8.7206e+10	1	8.7206e+10	Prob > F = 0.1960
Residual		9.1086e+13	1748	5.2109e+10	R-squared = 0.0010
-----+-----				Adj R-squared =	0.0004
Total		9.1174e+13	1749	5.2129e+10	Root MSE = 2.3e+05

-----+-----						
state_gdp		Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
-----+-----						
dwnom1		25811.95	19952.74	1.29	0.196	-13321.79 64945.7
is_jeff		1,000,000	not a lot	a lot	0.000	1,000,000 1,000,000
-----+-----						