U.S. Public Support for Biofuels Tax Credits: Cost Frames, Local Fuel Prices, and the Moderating Influence of Partisanship

Jillian L. Goldfarb*
Department of Biological and Environmental Engineering
Cornell University
jlg459@cornell.edu

Douglas L. Kriner
Department of Government
Cornell University
kriner@cornell.edu

Supplemental Information (for online publication only)

* To whom correspondence should be addressed: goldfarb@cornell.edu; JillianLGoldfarb@gmail.com; 607.255.5789

Appendix 1. Variation in Local Gasoline Prices

The analyses in the paper examine the relationship between the price of gasoline in each survey respondent's home ZIP code and support for a biofuel production tax credit, as well as whether this relationship varies by respondents' political partisanship and whether local gas prices moderate the effect of two experimental cost treatments on support for the credit. Data on daily gas prices for 98,753 gas stations during the week of September 8-14, 2014 were generously shared by Hakan Yilmazkuday.[1] From these, we calculated the average price per gallon during this period for each ZIP code for which data on at least one gas station was available. Figure 1 in the text shows the considerable variation in gasoline prices across the United States during this period. Much of this variation was regional. However, many states also exhibited considerable intra-state variation in gas prices, as shown for Pennsylvania, Florida, Ohio, and Massachusetts in Figure S1, below.

Using the Yilmazkuday data, we were able to calculate a home ZIP code-level average gasoline price for more than 90% of the survey respondents in our sample. In the rare cases where the data did not include a gas station for a subject's home ZIP code, we then identified all ZIP codes within a specified radius (5, 10, 15, or 20 miles) of its centroid and used the mean gas price in those ZIP codes as our measure of local gas prices. For only 3 subjects in our sample did we have to use gas price data from ZIP codes more than 10 miles from the home ZIP code centroid.

[1] Demet Yilmazkuday and Hakan Yilmazkuday. 2019. "Redistributive Effects of Gasoline Prices." *Networks and Spatial Economics* 19: 109-124.

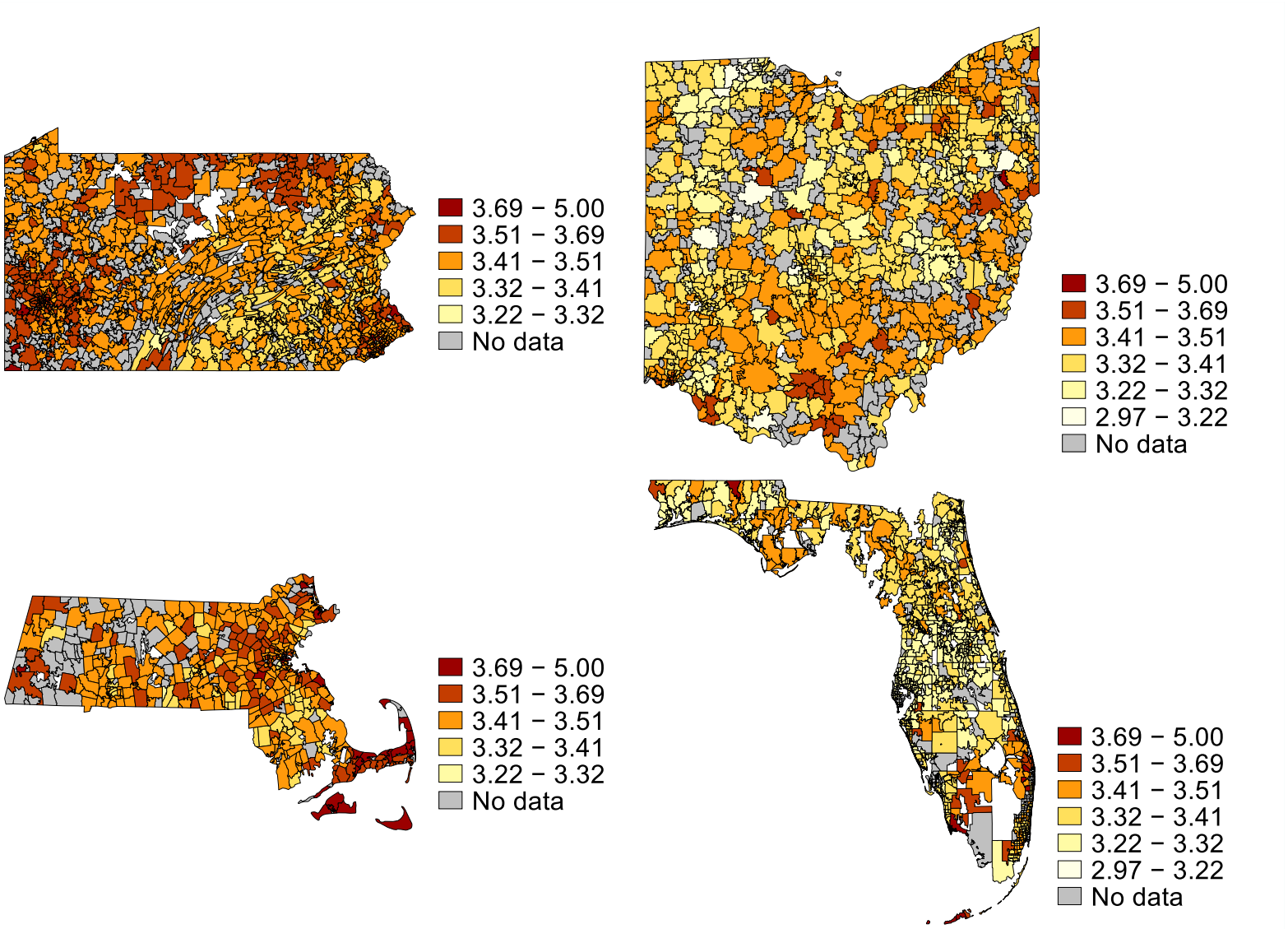


Figure S1: Intra-State Variation in ZIP Code-level Gasoline Prices in Four States (data provided by Hakan Yilmazkuday)

Appendix 2. Logistic Regression Diagnostics

Our survey experiment asked subjects to indicate their support for a biofuels production tax credit on a four-point Likert scale ranging from "strongly oppose" to "strongly support." Because the percentage supporting a proposed policy is the most politically salient quantity of interest – and the figure most often reported in the mass media – we collapsed the "strongly" and "somewhat support" response options to create a binary dependent variable coded 1 for those who supported the tax credit and 0 for those who opposed it. Table 1 in the text presents a series of logistic regressions to test the seven hypotheses about the factors driving public support for the biofuel tax credit.

Like all statistical models, logistic regression involves several assumptions. Perhaps most importantly, it assumes that the true conditional probabilities are a logistic function of the explanatory variables in the model; that there are no extreme outliers; and that there is no multicollinearity among the independent variables. We employ a range of tests of these to assess the validity of these assumptions in the current context. To test for multicollinearity, we calculated the mean variance inflation factor for the Independent Variables in our model. For each variable, the VIF was below 2 and the average for all independent variables in the base model (model 1 of Table 1) was below 1.5, far below any critical values for multicollinearity concerns. To test for the presence of significant outliers, after running our base model (model 1 of Table 1) we calculated standardized Pearson residuals for each observation. Seven observations in our sample yielded a standardized Pearson residual greater than the absolute value of 3, a commonly used rule of thumb for identifying outliers.

To ensure that the inclusion of these observations did not skew the results presented in Table 1, Table S1 (below) re-estimates each model specification excluding these 7 observations. Results are virtually identical, alleviating such concerns.

As a final test of these assumptions, we look for evidence of specification error. We do so in two ways. First, for the baseline model we calculate the Hosmer-Lemeshow goodness of fit test for ten quantiles. The test suggests a strong goodness of fit ($p = .67$; model is rejected if $p < .05$). We also conduct a link test, which estimates a new model of support for the biofuels tax credit that includes two terms, the linear predicted values from our baseline logistic regression model and the linear predicted values squared. The linear predicted values term is highly statistically significant ($p < .001$), as expected if the model fits the data reasonably well. The squared predictions term is also significant ($p < .05$), suggesting there may be a specification issue. However, running the same link test on the baseline model excluding the seven outlier observations (model 1, Table S1) shows no evidence of specification error (the coefficient for the squared prediction term is statistically insignificant, $p = .41$).

Table S1: Replicating Main Analyses Excluding 7 Outliers

	(1)	(2)	(3)	(4)
Decrease costs treatment	0.021	-1.899	0.404	0.025
	(0.198)	(2.868)	(0.406)	(0.198)
Increase costs treatment	-0.440**	1.921	-0.353	-0.434**
	(0.190)	(2.619)	(0.406)	(0.190)
Decrease costs x Local gas price		0.557		
		(0.827)		
Increase costs x Local gas price		-0.677		
		(0.750)		
Decrease costs x Democrat			-0.277	
			(0.530)	
Decrease costs x Republican			-0.700	
			(0.509)	
Increase costs x Democrat			-0.188	
			(0.513)	
Increase costs x Republican			-0.040	
			(0.503)	
Local gas price x Democrat				-0.684
				(0.854)
Local gas price x Republican				-0.746
				(0.849)
Democrat	0.986***	0.989***	1.149***	3.361
	(0.216)	(0.216)	(0.367)	(2.973)
Republican	-0.269	-0.259	-0.031	2.325
	(0.215)	(0.215)	(0.352)	(2.956)
Male	-0.536***	-0.533***	-0.542***	-0.535***
	(0.162)	(0.162)	(0.162)	(0.162)
Education	-0.005	-0.007	-0.003	-0.008
	(0.056)	(0.056)	(0.056)	(0.056)
Age	-0.036***	-0.036***	-0.036***	-0.036***
	(0.006)	(0.006)	(0.006)	(0.006)
White	0.309	0.303	0.289	0.305
	(0.200)	(0.201)	(0.202)	(0.200)
Local gasoline price	0.520	0.626	0.505	1.076
	(0.319)	(0.547)	(0.320)	(0.681)
Constant	0.722	0.354	0.638	-1.176
	(1.174)	(1.942)	(1.207)	(2.362)
Observations	878	878	878	878

Note: Models are logistic regressions and exclude 7 outliers from Table 1 (standardized residuals |3|). Standard errors in parentheses. All significance tests are two-tailed.
*** p<0.01, ** p<0.05, * p<0.10

Appendix 3. Alternate Models Analyzing a Four-Point Ordinal Dependent Variable

The analyses in the text analyze a dichotomized operationalization of the dependent variable as the percentage supporting (i.e. pooling strongly and somewhat) a tax credit is the most politically relevant measure of interest. However, as a robustness check we re-estimate all of the analyses in Table 1 in the text using the four-point ordinal dependent variable. Table 2 presents the results of a series of ordered logistic regressions. Table 3 presents the results of a series of OLS regressions on the same four-point ordinal scale. In both models, the results are substantively similar to those presented in Table 1 in the text.

In the baseline model in both Tables S2 and S3, we find evidence that the increase costs treatment significantly decreased support for the biofuels tax credit; but the decrease costs treatment had no effect. Moreover, SI Figures S2-S4 illustrating the moderating influence of local gas prices and partisanship largely mirror those presented in Figures 3-5 of the text. It is important to note, however, that for the ordered logit models SI Figures S2-S4 present the predicted probability of "strongly supporting" the tax credit. Plotting separate lines and confidence intervals for both the strongly support and somewhat support outcomes for multiple groups simultaneously renders graphical interpretation difficult.

Table S2: Ordered Logit Models Using 4-Point Support Score

	(1)	(2)	(3)	(4)
Decrease costs treatment	0.090	-0.869	0.112	0.097
	(0.155)	(2.208)	(0.328)	(0.156)
Increase costs treatment	-0.272*	1.237	-0.528	-0.265*
	(0.154)	(2.131)	(0.352)	(0.154)
Decrease costs x Local gas price		0.277		
		(0.633)		
Increase costs x Local gas price		-0.433		
		(0.609)		
Decrease costs x Democrat			0.174	
			(0.401)	
Decrease costs x Republican			-0.310	
			(0.426)	
Increase costs x Democrat			0.458	
			(0.420)	
Increase costs x Republican			0.119	
			(0.438)	
Local gas price x Democrat				-0.934
				(0.672)
Local gas price x Republican				-0.764
				(0.708)
Democrat	0.909***	0.908***	0.715**	4.157*
	(0.173)	(0.173)	(0.288)	(2.345)
Republican	-0.370**	-0.362**	-0.291	2.284
	(0.183)	(0.184)	(0.304)	(2.467)
Male	-0.297**	-0.294**	-0.296**	-0.300**
	(0.132)	(0.132)	(0.132)	(0.132)
Education	-0.033	-0.033	-0.033	-0.035
	(0.045)	(0.045)	(0.045)	(0.045)
Age	-0.023***	-0.023***	-0.023***	-0.023***
	(0.004)	(0.004)	(0.004)	(0.004)
White	0.170	0.161	0.146	0.174
	(0.156)	(0.157)	(0.157)	(0.156)
Local gasoline price	0.439*	0.504	0.410	1.121**
	(0.255)	(0.428)	(0.256)	(0.556)
Observations	885	885	885	885

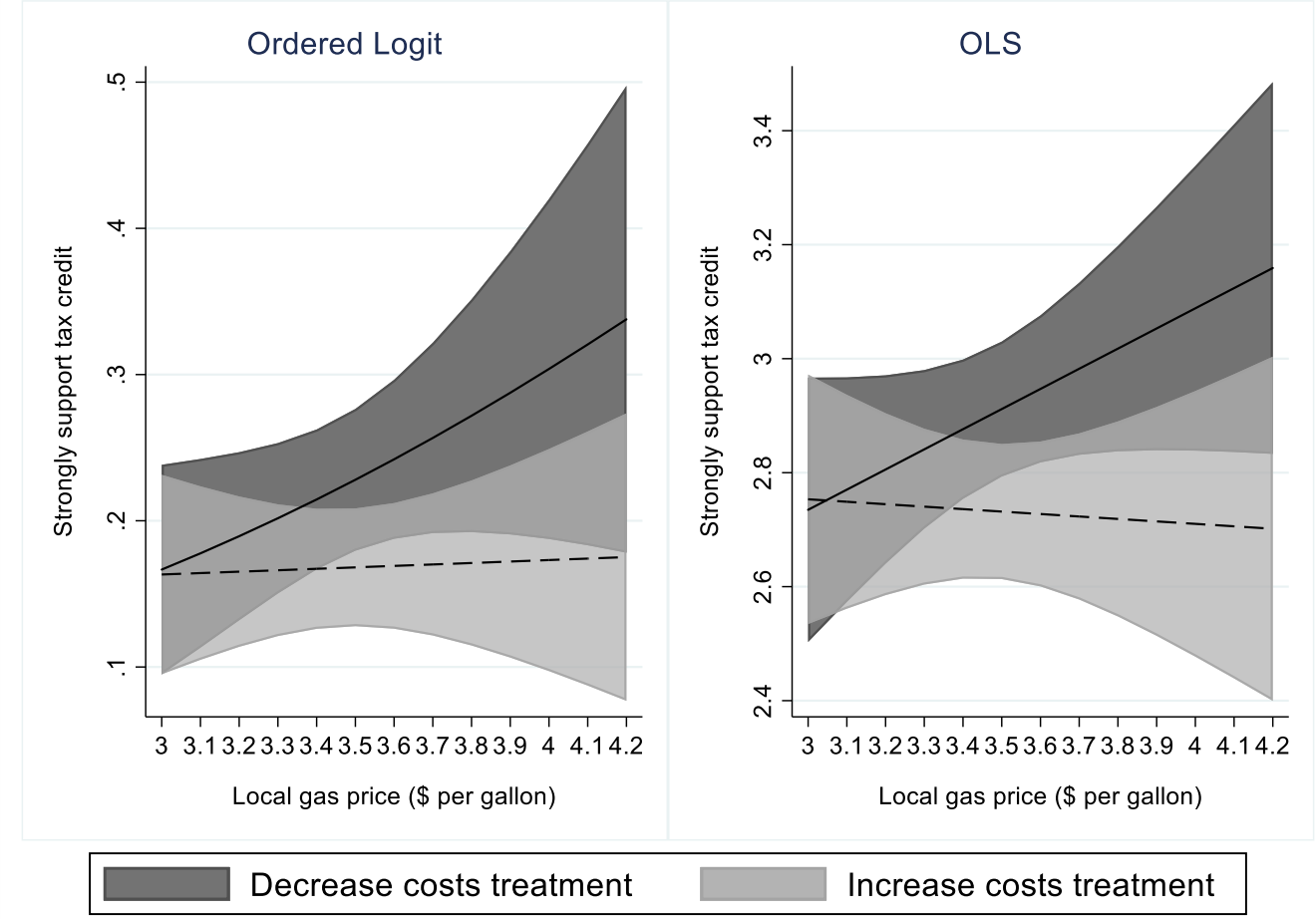
Note: Models are ordered logistic regressions. Standard errors in parentheses. All significance tests are two-tailed.
*** p<0.01, ** p<0.05, * p<0.10

Table S3: OLS Regressions Using 4-Point Support Score

	(1)	(2)	(3)	(4)
Decrease costs treatment	0.032	-0.526	0.044	0.035
	(0.072)	(1.015)	(0.154)	(0.072)
Increase costs treatment	-0.141**	0.682	-0.284*	-0.136*
	(0.070)	(0.971)	(0.163)	(0.070)
Decrease costs x Local gas price		0.161		
		(0.291)		
Increase costs x Local gas price		-0.236		
		(0.277)		
Decrease costs x Democrat			0.069	
			(0.186)	
Decrease costs x Republican			-0.142	
			(0.200)	
Increase costs x Democrat			0.233	
			(0.193)	
Increase costs x Republican			0.093	
			(0.203)	
Local gas price x Democrat				-0.548*
				(0.308)
Local gas price x Republican				-0.379
				(0.333)
Democrat	0.404***	0.405***	0.309**	2.313**
	(0.078)	(0.078)	(0.133)	(1.077)
Republican	-0.189**	-0.185**	-0.167	1.128
	(0.085)	(0.085)	(0.141)	(1.161)
Male	-0.154**	-0.152**	-0.154**	-0.153**
	(0.060)	(0.060)	(0.060)	(0.060)
Education	-0.020	-0.021	-0.020	-0.021
	(0.021)	(0.021)	(0.021)	(0.021)
Age	-0.011***	-0.011***	-0.011***	-0.011***
	(0.002)	(0.002)	(0.002)	(0.002)
White	0.092	0.088	0.079	0.092
	(0.071)	(0.071)	(0.071)	(0.071)
Local gasoline price	0.158	0.193	0.146	0.542**
	(0.117)	(0.197)	(0.117)	(0.259)
Constant	2.730***	2.606***	2.820***	1.401
	(0.428)	(0.701)	(0.441)	(0.906)
Observations	885	885	885	885
R-squared	0.151	0.153	0.154	0.154

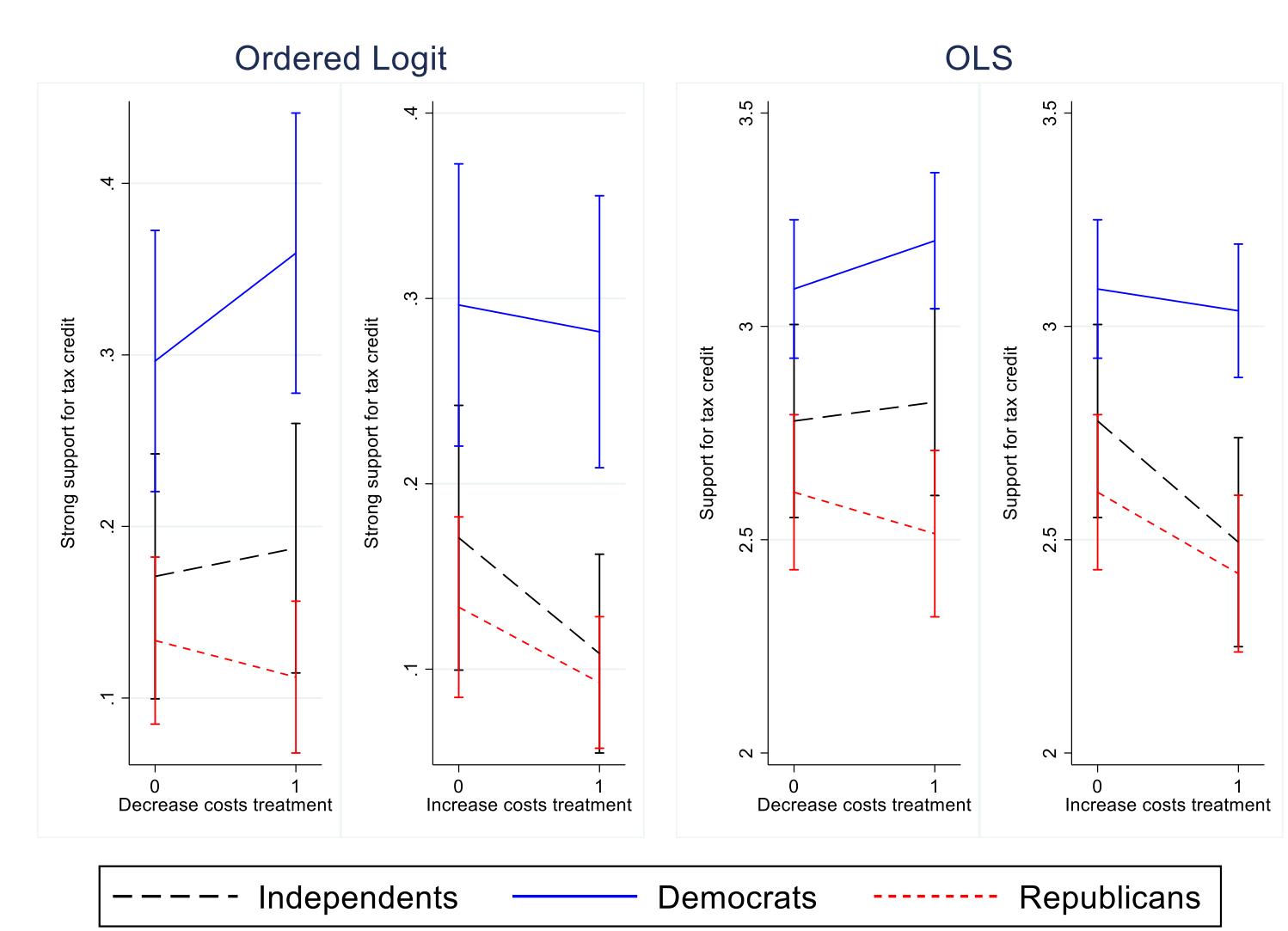
Note: Models are ordered logistic regressions. Standard errors in parentheses. All significance tests are two-tailed.
*** p<0.01, ** p<0.05, * p<0.10

Appendix 4. Supplemental Figures



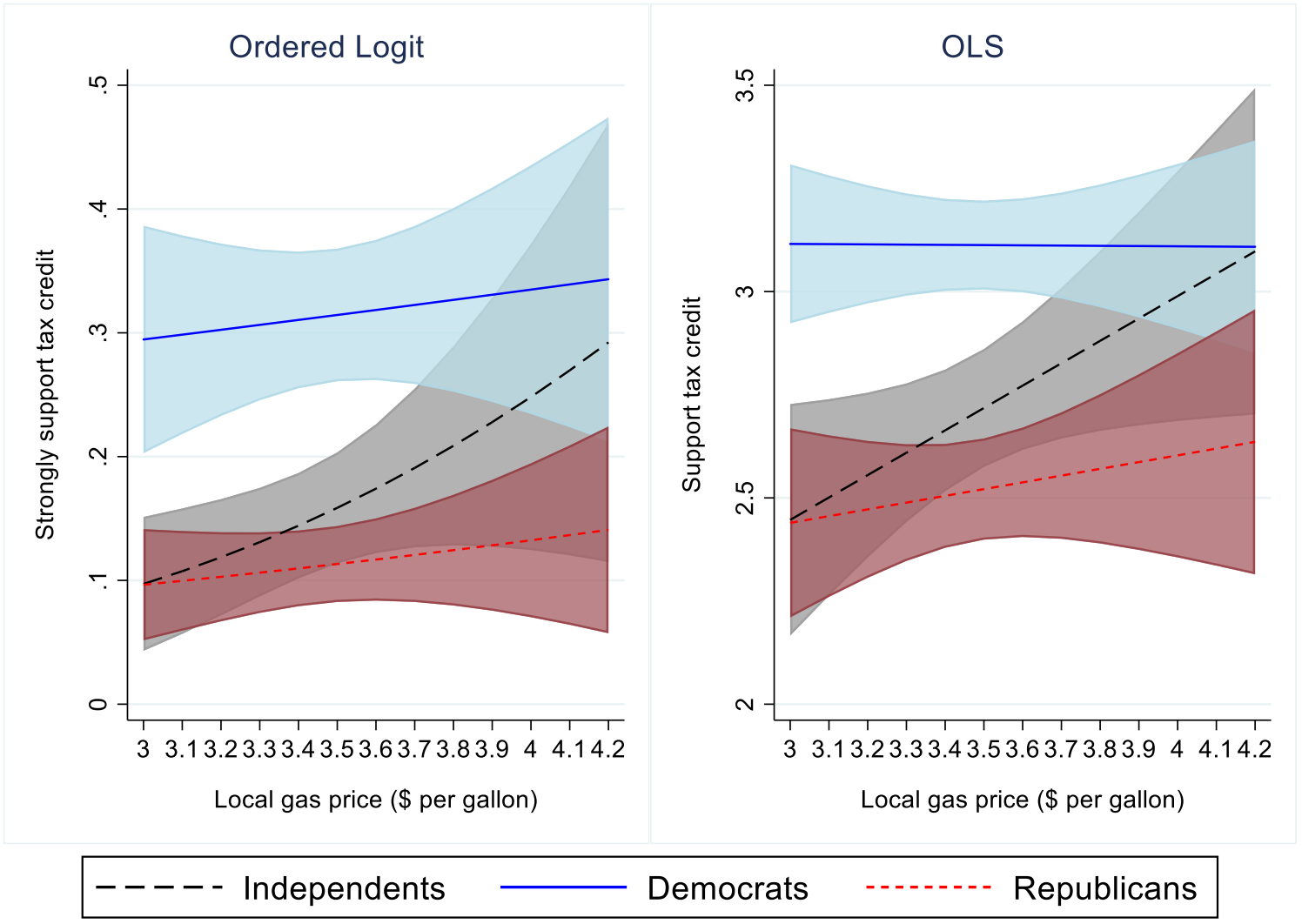
Note: Shaded regions present 95% confidence intervals around point predictions.

Figure S2: Moderation of Treatment Effects by Local Gas Prices, Ordered Logit and OLS Models of Ordinal DV



Note: I-bars present 95% confidence intervals around point predictions.

Figure S3: Partisan Moderation of Treatment Effects, Ordered Logit and OLS Models of Ordinal DV



Note: Shaded regions present 95% confidence intervals around point predictions.

Figure S4: Partisan Moderation of Effect of Local Gas Prices, Ordered Logit and OLS Models of Ordinal DV