Supporting Information for
Party Brands and Partisanship: Theory with Evidence from a Survey Experiment in Argentina

Noam Lupu

Formal Branding Model of Partisanship

Consider a multiparty democracy with M parties. Voters develop a perception of a party's brand over time on the basis of observations of party behavior. This brand is the perceived position of the party on the real line. For a voter who observes n actions by a party, denote her jth observation by y_j, where $j \in \{1, \ldots, n\}$.[1] Voters observe some subset of the actions taken by parties in each electoral cycle – with some voters of course drawing more observations from this distribution than others – and use these observations to update their beliefs about party brands. I assume that each observation y_j is a random draw from a distribution of observable party actions with unknown mean μ and unknown variance σ^2.[2] For simplicity, I assume a normal distribution of party actions, meaning that $y_j \sim_{iid} N(\mu, \sigma^2)$ for all j. Denote the sample mean of y_j by $\bar{y}$.

During any particular electoral cycle, let k_0 denote the number of observations on which a voter's prior beliefs about the party are based, such that at the end of each electoral cycle $k_n = k_0 + n$. Similarly, let v_0 denote the degrees of freedom in the uncertainty of the voter's prior beliefs, such that $v_n = v_0 + n$.

For convenience, I assume that an individual's prior belief about the mean and variance of a party's position is represented by the normal-inverse-chi-squared distribution, such that the posterior and prior distributions are drawn from the same family of distributions. As a result, prior beliefs about μ and σ^2 are given by $\mu, \sigma^2 | y \sim N - \chi^2(\mu_n, k_n, v_n, \sigma_n^2)$, where $k_0, v_0, \sigma_0^2 > 0$.

Decomposing the joint distribution of μ and σ^2 into $p(\mu, \sigma^2) = p(\mu|\sigma^2)p(\sigma^2)$, we find that the true distribution of party actions given its variance $(\mu|\sigma^2)$ is normally distributed with mean μ_0 and variance $\frac{\sigma^2}{k_0}$. Here $\hat{\mu}_0$ represents the voter's best guess about the position of the party at time 0 – that is, her prior – and $\frac{\sigma^2}{k_0}$ represents the amount of uncertainty about that best guess. That uncertainty is based on the actual variance in the distribution of party actions (σ^2) and the number of observations on which a voter's prior beliefs are based (k_0). With regard to the uncertainty parameter σ^2, we find that it is inverse-chi-squared distributed along a prior scale σ_0^2 with v_0 degrees of freedom. Formally, $\mu|\sigma^2 \sim N(\mu_0, \frac{\sigma^2}{k_0})$ and $\sigma^2 \sim \chi^{-2}(v_0, \sigma_0^2)$. Thus, the prior uncertainty about the party's location is given by $E[\sigma_0^2] = \sigma_0^2 \frac{v_0}{v_0 - 2}$.

I assume that voters obey the axioms of probability theory and Bayes' Rule. Individual voters observe y_j and use this information to update their beliefs about μ and σ^2. Given these assumptions, the voter's updated perception of the party brand combines prior and new information. Formally, the joint posterior distribution of μ and σ^2 is given by $\mu, \sigma^2|y \sim N - \chi^2(\hat{\mu}_n, k_n, v_n, \sigma_n^2)$, where

$$\hat{\mu}_n = \frac{k_0}{k_0+n}\hat{\mu}_0 + \frac{n}{k_0+n}\bar{y} \quad (1)$$

$$k_n = k_0 + n \quad (2)$$

$$v_n = v_0 + n \quad (3)$$

$$v_n\sigma_n^2 = v_0\sigma_0^2 + (n-1)s_n^2 + \frac{k_0 n}{(k_0+n)(\bar{y}-\hat{\mu}_0)^2} \quad (4)$$

$$s_n^2 = \sum \frac{(y_i - \bar{y})^2}{n-1}. \quad (5)$$

The new perceived party position of the party $(\hat{\mu}_n)$ is a weighted average of the prior mean $(\hat{\mu}_0)$ and the average party position of the new observations $(\bar{y})$. The weights of the prior and new observations are determined by the amount of prior information (k_0) and the number of new observations (n). And the new estimate of the party's position has a new estimate of precision (σ_n^2) determined by the precision about the prior mean (σ_0^2), the precision about the new observations (s_n^2), and the combination of old and new degrees of freedom (v_n).[3]

Since the marginal posterior distribution of μ now has a student-t distribution, in expectation a voter's updated perception of the party position is equal to the true party position. That is, $\mu|y \sim t_{vn}(\hat{\mu}_n, \frac{\sigma^2}{k_n})$, meaning that $E_n[\mu] = \hat{\mu}_n$. In other words, the voter should expect that the true party position (μ), after having drawn n new observations, is the weighted average μ_n. The marginal posterior distribution of σ^2 has an inverse-chi-squared distribution, such that a voter's uncertainty about their estimate of the party brand σ^2 is, in expectation, a function of the precision around their old and new observations of the party. Formally, $\sigma^2|y \sim -\chi^{-2}(v_n, \sigma_n^2)$, and thus $E_n[\sigma^2] = \frac{\frac{\sigma_n^2}{v_n}}{v_n - 2}$.

Party brands help determine how voters feel about parties. Voters are assumed to have an ideal point along the continuum on which they evaluate all the parties. Let x_i denote voter i's ideal point along a real line representing this continuum. In multiparty settings, many voters are unaware of the existence of very minor parties. More likely, voters consider some subset of the universe of parties that are particularly visible, by virtue of their electability or regional presence. We could therefore say that some subset of the parties in the system are in an individual's "consideration set" (Wilson 2008). Put formally, each voter i has in mind an $M \times 1$ vector $\mathbf{S}_i = (S_1^i, \cdots S_i^m, \cdots, S_i^M)$ in which each component takes a value of zero or one – one if the voter thinks of party m as a real option and zero otherwise. I assume that at least one element of $\mathbf{S}_i$ is nonzero, so that each voter considers at least one party.

Voters identify with the party that most represents people like them. As in self-categorization theories of identity formation (Turner et al. 1987), voters identify with the party whose prototypical partisan they believe they most closely resemble. Let y^m denote draws of voter types from the distribution of prototypical partisans of party $m \in M$. Let μ^m denote the expected value of these draws from the posterior distribution for party m, and σ^{2m} denote the variance of this distribution. Using this setup, we can model resemblance as a function of proximity and uncertainty:

$$R_i[y^m] = -(x_i - \mu^m)^2 - \sigma^{2m} \text{ for all } m \text{ such that } S_i^m = 1. \quad (6)$$

Substituting the voter's expected values of μ and σ^2 from the updating model, we get:

$$R_i[y^m] = -(x_i - \mu_n^m)^2 - \sigma_n^{2m} \text{ for all } m \text{ such that } S_i^m = 1. \tag{7}$$

Thus, a voter's resemblance with a party is a function of her perceived proximity to the party – the distance between her ideal point and the party brand – as well as her certainty about the party brand. Resemblance is therefore increasing in the voter's perceived proximity to the party and decreasing in her uncertainty about the brand. And voters do not consider parties they think are unelectable: when a party is thought to have no chance of being elected, $S_i^m = 0$.

If partisanship is an identity formed on the basis of self-categorization and resemblance, we can model model partisanship in continuous terms, denoted γ_i^m.[4] As an identity, partisanship is an increasing function of a voter's resemblance with the party's prototypical partisan. But, like other identities, a voter will feel a particularly strong attachment to her party if she resembles its prototypical partisan far more than she does another party's (comparative fit). Conversely, if she perceives that she almost equally resembles her party's prototype and another party's prototype, her attachment will be very weak. Put formally,

$$\gamma_i^m = \min_{q \in M}(R_i[y^m] - R_i[y^q])^2 \text{ for all } q, m \in M \text{ such that } S_i^m = S_i^q = 1. \tag{8}$$

Survey Methodology

The 2009 Argentina survey was conducted in collaboration with Valeria Brusco, Thad Dunning, Marcelo Nazareno, and Susan Stokes. The survey consisted of face-to-face interviews in Spanish of 1,199 eligible voters in the Argentine provinces of Córdoba and Santa Fe. On average, the survey interview lasted 51 minutes. The survey was administered by the Córdoba-based polling firm Consultores en Políticas Públicas between August 28 and October 2, 2009.

A two-stage clustered random sample based on the 2001 national census was generated within each province. Sixty *radios censales*, the smallest available geographic unit in the census, were selected as the primary sampling units (PSUs) in each province, with ten cases conducted in each PSU. PSUs in localities with a population under 1,000 were replaced with the corresponding PSU in a second sample. Interviewers began from a randomly selected corner in the PSU and proceeded in a clockwise direction selecting every fourth household. Within each household, the adult Argentine citizen with the most recent birthday was selected. In cases where no one responded or the selected individual was not home, interviewers either made an appointment to return to the household when the individual would be home or returned later in the same day. If an effective interview could not be conducted upon the second attempt, or if the household/individual refused to complete the survey, the household was replaced. Households were replaced with the adjacent household. Interviewers used the four different questionnaires in sequential order.

Interviewers were recruited from Córdoba universities and were mostly advanced undergraduate and graduate students. Extensive training was conducted for interviewers on selection methodology, the logistics of the survey instrument, and issues of respondent protection such as anonymity and privacy. On a separate sheet from the questionnaire, interviewers recorded the first name only and phone number of each respondent for the purposes of later supervision. Post-sampling verification was conducted on a randomly-selected 30% of the sample by telephone, after which this information was destroyed. If mistakes or omissions were found in one case, further verification was conducted on the interviews conducted by that individual. If a second mistake or omission was then found by the same interviewer, the full set of interviews conducted by that individual was conducted anew by a different interviewer. The response rate for the survey was 19.3%, the cooperation rate 30.7%, the refusal rate 43.5%, and the contact rate 62.9%. The margin of error was 6.7%.

The survey was preceded by a pilot consisting of 100 cases each in the provincial capital cities of Córdoba and Santa Fe. The pilot was administered between August 1 and August 9 using an identical selection methodology. Its goal was to test the survey instrument and to elicit feedback from respondents and interviewers regarding question wording and order.

Table A1. *Survey sample representativeness*

		2001 census	
Variable	**2009 survey**	**Córdoba & Santa Fe**	**National**
Gender			
Male	39.70	48.54	48.70
Female	60.30	51.46	51.30
Age			
18-24	13.36	18.43	18.54
25-34	20.67	19.82	21.01
35-44	18.66	17.55	18.12
45-54	17.48	15.81	15.87
55-64	14.20	12.51	11.57
65+	15.63	16.24	14.89
Education			
No primary	8.71	37.32	38.83
Primary only	39.20	36.87	37.33
Secondary only	36.01	14.07	17.56
Tertiary	16.08	11.74	6.27
Population			
Córdoba	50.04	50.46	8.44
Santa Fe	49.96	49.54	8.29

Table A2. *Comparison of treatment conditions*

Variable	Condition 1	Condition 2	Condition 3	Condition 4
Observations	301	295	302	301
Female	0.628	0.627	0.579	0.579
Age	45.35	44.02	45.40	44.68
Education	2.557	2.626	2.631	2.565
Income	4.801	5.035	5.037	5.095

Notes: Although the treatment conditions were intended to be of equal size ($N = 300$), a clerical error by the polling firm resulted in a slightly less-than-balanced set of groups. This minor difference, however, should not have any effect on the ability to draw inferences from the data.

Table A3. *Average treatment effects*

Treatment	ATE	T-test	Wilcoxon test	F-P test
Partisanship				
Platform information	0.0877	0.029	0.030	0.033
Alliance/switching information	-0.116	0.002	0.002	0.007
All information	-0.0324	0.405	0.405	0.247
Strength of partisanship				
Platform information	0.663	0.026	0.012	0.006
Alliance/switching information	-0.992	0.003	0.002	0.001
All information	-0.100	0.756	0.846	0.425
Polarization				
Platform information	0.451	0.092	0.059	0.031
Alliance/switching information	-0.585	0.022	0.020	0.010
All information	-0.214	0.409	0.436	0.220
Mandate-responsiveness				
Platform information	-0.060	0.543	0.515	0.287
Alliance/switching information	-0.056	0.564	0.996	0.498
All information	-0.062	0.516	0.824	0.423

Notes: Average treatment effects are differences in mean values (proportions in the case of partisanship) between each treatment group and the control group. P-values are reported for difference-of-means test with Welch's approximation (difference-of-proportions in the case of partisanship), the Wilcoxon rank-order test, and the Fligner-Policello rank-order test.

Table A4. *Regression analyses of average treatment effects, without covariates*

	Partisanship	**Strength of partisanship**		**Polarization**	**Mandate responsiveness**	
	Probit	**OLS**	**Ordered probit**	**OLS**	**OLS**	**Ordered probit**
Platform information	0.227**	0.663**	0.338**	0.451*	-0.060	-0.067
	(0.104)	(0.295)	(0.144)	(0.267)	(0.099)	(0.101)
Alliance/switching information	-0.337***	-0.992***	-0.465***	-0.585**	-0.056	-0.015
	(0.108)	(0.322)	(0.143)	(0.254)	(0.097)	(0.097)
All information	-0.088	-0.100	-0.041	-0.214	-0.062	-0.040
	(0.106)	(0.321)	(0.145)	(0.259)	(0.096)	(0.097)
Constant	-0.364***	7.131***		5.543***	1.788***	
	(0.074)	(0.214)		(0.177)	(0.071)	
Observations	1,190	397	397	740	1,168	1,168
Pseudo-R^2	0.019		0.016			0.000
Adjusted R^2		0.060		0.022	0.001	
Wald $\chi^2_{(3)}$	28.076***		29.764***			0.512

Notes: Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5. *Regression analyses of average treatment effects, with covariates*

	Partisanship	**Strength of partisanship**		**Polarization**	**Mandate responsiveness**	
	Probit	**OLS**	**Ordered probit**	**OLS**	**OLS**	**Ordered probit**
Platform information	0.262**	0.763***	0.396***	0.578**	-0.069	-0.078
	(0.106)	(0.292)	(0.144)	(0.272)	(0.098)	(0.101)
Alliance/switching information	-0.355***	-1.089***	-0.516***	-0.549**	-0.068	-0.026
	(0.109)	(0.319)	(0.143)	(0.255)	(0.096)	(0.098)
All information	-0.073	-0.020	-0.001	-0.159	-0.072	-0.044
	(0.108)	(0.325)	(0.148)	(0.261)	(0.096)	(0.099)
Political information	0.102*	0.047	0.033	0.141	0.032	0.057
	(0.060)	(0.158)	(0.078)	(0.149)	(0.052)	(0.054)
Education	-0.007	-0.141	-0.081	-0.148	0.106**	0.122***
	(0.050)	(0.145)	(0.067)	(0.130)	(0.045)	(0.045)
Age	0.006**	0.015**	0.007**	0.007	-0.004*	-0.005**
	(0.002)	(0.007)	(0.003)	(0.006)	(0.002)	(0.002)
Gender	-0.046	0.227	0.146	0.044	-0.114	-0.070
	(0.079)	(0.223)	(0.105)	(0.193)	(0.072)	(0.073)
Córdoba	-0.102	-0.366	-0.171	0.113	0.054	0.060
	(0.076)	(0.226)	(0.105)	(0.190)	(0.068)	(0.070)
Constant	-0.635***	6.736***		5.345***	1.704***	
	(0.204)	(0.669)		(0.483)	(0.196)	
Observations	1,176	394	394	731	1,154	1,154
Pseudo-R^2	0.029		0.024			0.010
Adjusted R^2		0.088		0.030	0.017	
Wald $\chi^2_{(8)}$	45.021***		45.121***			23.720***

Notes: Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6. *Complier average causal effects*

	Partisanship
Platform information	0.192***
	(0.068)
Alliance/switching information	-0.115*
	(0.062)
All information	0.020
	(0.068)
Constant	0.319***
	(0.047)
Observations	1,190
R^2	0.015
Wald $\chi^2_{(3)}$	42.002***

Notes: Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A7. *Average causal mediation effects*

	Polarization	Mandate responsiveness
Platform information	0.023	-0.002
	[0.0004, 0.053]	[-0.012, 0.004]
Alliance/switching information	-0.022	-0.003
	[-0.045, -0.004]	[-0.013, 0.008]
All information	-0.005	-0.002
	[-0.028, 0.016]	[-0.011, 0.005]

Notes: Effects are estimated following Imai et al. (2011) with linear regression for the mediator model and probit for the outcome model. Models include controls for political information, education, age, gender, and province. In brackets are 95% confidence intervals based on 1,000 nonparametric bootstrap simulations.

References

Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. 2011. "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." *American Political Science Review* 105 (4): 765-789.

Turner, John C., Michael A. Hogg, Penelope Oakes, Stephen Reicher, and Margaret Wetherell. 1987. *Rediscovering the Social Group: A Self-Categorization Theory*. New York: Blackwell.

Wilson, Carole J. 2008. "Consideration Sets and Political Choices: A Heterogeneous Model of Vote Choice and Sub-national Party Strength." *Political Behavior* 30 (2): 161-183.