Appendix: Data and Measures

The underlying data for this project comes primarily from two sources: Population, ethnicity, poverty, and occupational categories data comes from Bolivia's Instituto Nacional de Estadística (INE). One exception is data on 2001 poverty categories, which was not readily available from the INE website. This data was provided by Datax, a Bolivian data and consulting company based in Cochabamba. Another exception is population density, which required data on the surface area of each of the municipalities. That data was also not available from any official government databases. After some inquiries, the data was acquired from Hugo César Boero Kavlin, Director of the Programa de Fortalecimiento Institucional. That municipal surface area database was compiled by Oswaldo Boero, a geographer for the government of the city of La Paz, based on data from the Ministerio de Autonomías. All election data comes directly from Bolivia's Órgano Electoral Plurinacional (OEP).

Below are some explanations of the individual variables used:

Population and Population Density (pop/km^2)

Because four municipalities (La Paz, El Alto, Cochabamba, and Santa Cruz) comprise about half of the total population of Bolivia, the data for 339 municipalities was highly skewed. In order to properly control for the effects of population, this variable was transformed using the natural log.

After calculating each municipality's population density (2012 population divided by its surface area), a similar transformation was made (using the natural log) in order to adjust for the effects of a handful of highly urban municipalities.

Ethnicity and Ethnic Diversity

Ethnicity data for 2012 was available for all 339 municipalities. Although the 2012 census data included numbers for all 32 nationally recognized ethnic identities (as well as numerous smaller "other" groups self-identified by individual census respondents), this study specifically looked at three key groups, based on census self-identification: Aymara, Quechua, and Non-Indigenous. All are reported as the fractional share of the number of respondents in each municipality who self-identified as belonging to either group divided by the total municipal population.

An additional aggregate measure of *ethnic diversity* was calculated using the Ethnic Fractionalization Index (EFI) proposed by Alesina et al (2002):

$$EFI = 1 - \sum_{i=1}^{n} (s_i)^2$$

where s_i is the population share of each ethnic category i, and n is the number of ethnic groups present in each municipality. This is an aggregate measure of "cultural diversity," which ranges from 1 (all individuals belong to unique ethnic categories) to 0 (all individuals belong to the same ethnic category). This measure is widely used in the literature because of its underlying relationship (in cross-national studies) to political and economic indicators, including democracy

and level of economic development. Thus, EFI scores were used as a control (like population and population density) for underlying contextual features of individual municipalities.

Socioeconomic Structure

In addition to ethnicity (and especially degree of "cultural diversity"), socioeconomic structure is an important contextual variable at the municipal level. This study measures socioeconomic structure by estimating the occupational heterogeneity index, measured using the Herfindahl-Hirschman Index (HHI):

$$HHI = \sum_{i=1}^{n} (s_i)^2$$

where s_i is the share of economically active population engaged in each occupational category i, and n is the number of occupational categories for each municipality. As the number approaches 1 it signals greater concentration in a single occupational category.

An alternate measure is the Index of Occupational Diversity (IOD) proposed by Tatu Vanhanen (1989). IOD is simply the arithmetic mean of the share of urban population and share of population working in non-agricultural sectors. However, this measure is rather crude, since it merely divides populations into two extremes: rural-agricultural and urban-non-agricultural. Because population and population density were already introduced as control variables, it made more sense to measure socioeconomic structure along a single dimension (rather than combine two dimensions). Uniformly, the municipalities with high occupation HHI scores are those in which agricultural employment is the supermajority. Importantly, across Bolivian municipalities, increase in occupational diversity is not strongly correlated with lower population density. ***

Poverty and Wealth Inequality

Although Bolivia's census does not ask about poverty, the INE uses census data to calculate poverty levels using the method of Necesidades Básicas Insatisfechas (NBI) developed by UN Economic Commission for Latin America and the Caribbean (ECLAC) in the 1980s. The method categorizes households into one of five categories based on whether their basic needs are met. Those whose basic needs are not met include the lower three categories: moderate poverty (*pobreza moderada*), indigent (*indigente*,), and marginal (*marginalidad*). Those whose basic needs are met make up the upper two categories: those at risk of poverty (*umbral de pobreza*) and those who are safely outside of poverty.

Rather than simply use absolute poverty levels (as other studies have done), I constructed a municipal wealth index, which is simply the estimated mean of the values for an index constructed from the five-category grouping used by the Bolivian census.

$$Wealth\ Index = \frac{\sum f x_i}{n}$$

where fx_i is the frequency of the *i*-th poverty category and *x* is the category's ascribed ordinal value, which ranges from 0 ("*indigente*") to 4 ("*necesidades básicas satisfechas*"). Higher scores suggest that wealth is, on the whole, higher than in municipalities with lower wealth index scores.

The wealth index tells us little about the shape of the distribution. Thus, I also calculated the municipal poverty variance using the simple formula to estimate variance from grouped data:

$$Poverty\ Variance = \frac{\sum f(x - \bar{x})^2}{n}$$

Measuring the municipal poverty variance provides rough estimate of the level of economic inequality, regardless of mean poverty index score. For example, if we compare two municipalities with similar poverty index scores, the one with the smaller variance is more equal, since the population is clustered more closely around the mean.

Lastly, I also calculated the change in the wealth index measure for each municipality between the 2001 and 2012 census. While the wealth index provides a snapshot of the "average" level of wealth in the municipality, and the poverty variance tells us about inequality, this latter measure tells us about the relative change in economic fortunes in any municipality during roughly the decade between censuses.

Level of Electoral Support for MAS

Level of electoral support for MAS in each of the four elections (2009, 2010, 2014, and 2015) was estimated using the share of the MAS vote in each municipality as a product of the *valid* vote (all votes cast for a political party).

Opposition Dummy

Because in municipal elections (2010 and 2015) some municipalities saw MAS run without any opposition, an "opposition" dummy variable was inserted. This takes a value of 1 if at least one other party (regardless of its electoral support) campaigned against MAS in that municipality, and 0 if MAS ran unopposed.

Issues Related to Missing Data

A small number of municipalities were created between the 2001 and 2012 censuses, which posed problems for comparative analysis. One solution is deletion of missing municipalities. However, this poses additional problems: Since new municipalities are created from existing ones, simple listwise or pairwise deletion would necessitate dropping from the analysis any municipality that was later split into two or more. This poses potential for selection bias, since it is likely that municipalities that were later split (usually because of within-municipality territorial disputes) shared some characteristics that made them different from other municipalities.

Instead of deletion, a strategy of unit data imputation was used. This strategy is facilitated by the fact that data are for aggregate units, not individuals. Additionally, much of the underlying

census data did not have to be imputed because the INE, which had access to original underlying census data, had already retroactively constructed basic measures for the new municipalities. Thus, the only data that needed imputation was 2001 poverty and wealth data, and some specific election data.

The strategy used to impute missing data was simple: Any municipality for which data was missing was identified, along with its "parent" municipality (the municipality it split from). The aggregate measure for the "parent" municipality was then simply copied directly into the new municipality. For example, the new municipality of Chacarilla was created in La Paz prior to the 2012 census. It was created out of the existing municipality of Papel Pampa (in Gualberto Villarroel province). Because poverty data for Chacarilla was not available for 2001, data for Papel Pampa (the "parent" municipality) was used instead. Overall, data imputation as used for a total of 20 municipalities (this number includes new and "parent" municipalities), which accounts for 5.9 percent of the total sample.

Although there is potential for introducing error with any data imputation strategy, the error should be minimal because any overestimation should be accompanied by underestimation. Nevertheless, in order to reduce the possibility of error due to data imputation, all regressions were estimated with jackknifed standard errors. Additionally, all the regression models were tested for multicollinearity (using both variance inflation factor and variance of the coefficient estimates in Stata) and none of the variables approached the standard thresholds.

References

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg. 2003. "Fractionalization." *Journal of Economic Growth* 8 (2): 155-194.

Vanhanen, Tatu. 1989. "The Level of Democratization Related to Socioeconomic Variables in 147 States in 1980-1985." *Scandinavian Political Studies* 12 (2): 95-127.

Table 1. Determinants of Municipal-Level Support for MAS in the 2009 presidential election

Log population	*** -0.058	*** -0.045	**-0.024	* -0.019	** -0.024
	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)
Log population density	*** 0.073	*** 0.082	*** 0.049	*** 0.045	*** 0.048
	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)
Ethnic fractionalization	* 0.182	* 0.173	0.064	0.100	0.062
	(0.076)	(0.072)	(0.055)	(0.055)	(0.055)
Occupational heterogeneity (HHI)	*** 0.684	*** 0.258	*** 0.252	*** 0.277	*** 0.235
	(0.052)	(0.077)	(0.054)	(0.055)	(0.055)
Non-indigenous population			*** -0.483	*** -0.431	*** -0.474
			(0.048)	(0.051)	(0.048)
Aymara				*** 0.089	
				(0.017)	
Quechua					0.031
					(0.018)
Wealth index (2012)		*** -0.264	-0.026	-0.020	-0.027
		(0.043)	(0.036)	(0.036)	(0.036)
Wealth variance (2012)		0.109	0.113	* 0.125	0.100
		(0.082)	(0.059)	(0.058)	(0.060)
Wealth index change (2001-12)		** 0.118	0.040	0.040	0.041
		(0.038)	(0.031)	(0.031)	(0.031)
Adjusted R^2	0.4330	0.5204	0.6928	0.7052	0.6935

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. N=339. Standard errors in parenthesis. OLS with jackknife estimation of standard errors.

Table 2. Determinants of Municipal-Level Support for MAS in the 2014 presidential election

Log population	*** -0.034	** -0.022	-0.008	-0.009	-0.010
	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)
Log population density	*** 0.050	*** 0.057	*** 0.035	*** 0.036	*** 0.033
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Ethnic fractionalization	* 0.132	* 0.132	0.060	0.054	0.055
	(0.061)	(0.058)	(0.044)	(0.044)	(0.043)
Occupational heterogeneity (HHI)	*** 0.627	*** 0.294	*** 0.290	*** 0.286	*** 0.247
	(0.040)	(0.060)	(0.045)	(0.045)	(0.043)
Non-indigenous population			*** -0.320	*** -0.329	*** -0.299
			(0.037)	(0.038)	(0.036)
Aymara				-0.016	
				(0.015)	
Quechua					*** 0.076
					(0.015)
Wealth index (2012)		*** -0.194	-0.037	-0.038	-0.037
		(0.033)	(0.031)	(0.031)	(0.030)
Wealth variance (2012)		0.011	0.014	0.012	-0.018
		(0.069)	(0.056)	(0.056)	(0.056)
Wealth index change (2001-12)		** 0.091	0.039	0.039	0.041
		(0.030)	(0.027)	(0.027)	(0.026)
Adjusted R^2	0.4669	0.5451	0.6643	0.6640	0.6788

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. N=339. Standard errors in parenthesis. OLS with jackknife estimation of standard errors.

Table 3. Determinants of Municipal-Level Support for MAS in the 2010 municipal election

Log population	-0.015	-0.016	-0.017	** -0.026	* -0.022
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Log population density	0.002	0.004	0.005	0.008	0.001
	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)
Ethnic fractionalization	0.060	0.059	0.069	-0.021	0.055
	(0.052)	(0.052)	(0.053)	(0.050)	(0.050)
Occupational heterogeneity (HHI)	* 0.132	0.105	0.086	0.002	-0.030
	(0.058)	(0.076)	(0.078)	(0.068)	(0.066)
MAS 2009 vote share	*** 0.377	*** 0.360	*** 0.439	*** 0.543	*** 0.406
	(0.044)	(0.051)	(0.059)	(0.062)	(0.056)
Opposition dummy	*** -0.441	*** -0.442	*** 0.443	*** -0.422	*** -0.417
	(0.015)	(0.016)	(0.016)	(0.021)	(0.020)
Non-indigenous population			* 0.106	0.031	*** 0.149
			(0.046)	(0.046)	(0.044)
Aymara				*** -0.207	
				(0.028)	
Quechua					*** 0.221
					(0.027)
Wealth index (2012)		-0.030	-0.061	-0.074	-0.065
		(0.044)	(0.047)	(0.043)	(0.042)
Wealth variance (2012)		0.039	0.030	-0.014	-0.068
		(0.059)	(0.059)	(0.056)	(0.058)
Wealth index change (2001-12)		0.006	0.013	0.011	0.023
		(0.042)	(0.043)	(0.043)	(0.043)
Adjusted R^2	0.6037	0.6010	0.6048	0.6667	0.6802

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. N=339. Standard errors in parenthesis. OLS with jackknife estimation of standard errors.

Table 4. Determinants of Municipal-Level Support for MAS in the 2015 municipal election

Log population	* -0.019	* -0.019	* -0.021	*** -0.034	*** -0.029
	(0.008)	(0.009)	(0.009)	(0.008)	(0.008)
Log population density	0.002	-0.001	0.002	0.013	0.002
	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)
Ethnic fractionalization	0.085	0.084	0.099	0.010	0.090
	(0.064)	(0.065)	(0.066)	(0.060)	(0.062)
Occupational heterogeneity (HHI)	*** 0.245	** 0.275	** 0.245	** 0.191	* 0.148
	(0.069)	(0.087)	(0.088)	(0.074)	(0.073)
MAS 2014 vote share	*** 0.224	*** 0.238	*** 0.344	*** 0.315	* 0.190
	(0.059)	(0.065)	(0.078)	(0.078)	(0.077)
Opposition dummy	*** -0.468	*** -0.468	*** -0.479	*** -0.456	*** -0.445
	(0.016)	(0.017)	(0.019)	(0.024)	(0.023)
Non-indigenous population			** 0.129	-0.015	** 0.145
			(0.048)	(0.049)	(0.048)
Aymara				*** -0.223	
				(0.029)	
Quechua					*** 0.249
					(0.029)
Wealth index (2012)		0.034	-0.008	-0.024	-0.014
		(0.044)	(0.047)	(0.043)	(0.043)
Wealth variance (2012)		-0.022	-0.022	-0.055	* -0.135
		(0.063)	(0.062)	(0.061)	(0.062)
Wealth index change (2001-12)		0.017	0.029	0.028	0.039
		(0.047)	(0.048)	(0.044)	(0.045)
Adjusted R^2	0.5290	0.5258	0.5333	0.6128	0.6302

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. N=339. Standard errors in parenthesis. OLS with jackknife estimation of standard errors.