Unobserved Heterogeneity and Bias in Tests of the Leviathan Hypothesis

Luis Gonzalez and Carlos Lopes*

ABSTRACT. This paper highlights a potential source of bias in empirical tests that attempt to link the level of fiscal centralization to the total amount of public spending in a given jurisdiction. This paper is the first to propose biased estimation as a possible explanation for the mixed results in this line of literature and is also the first to test for this bias by controlling for both time-variant and time-invariant population heterogeneity. We begin by establishing a link between government centralization and the level of heterogeneity of preferences among constituents. We then combine this link with results from recent studies that link heterogeneity to lower levels of public provision, which highlights the source of our proposed bias. Our model uses ethnic fractionalization as well as income inequality as measures of time-variant heterogeneity of preferences. We find that the omission of time-invariant heterogeneity has an impact on the estimated relationship between centralization and public spending. Depending on the specification, the resulting bias can be severe enough to change the estimated direction of the relationship between centralization and spending. (H77, H73)

> *Total government intrusion into the economy should be smaller, ceteris paribus, the greater the extent to which taxes and expenditures are decentralized, the more homogeneous are the separate units, the smaller the jurisdictions, and the lower the net locational rents.*
>
> -Brennan and Buchanan, *The Power to Tax: Analytical Foundations of a Fiscal Constitution*, p. 185

I. Introduction

In 1980 Brennan and Buchanan popularized the "Leviathan Hypothesis" in their book *The Power to Tax*. One implication of the Leviathan Hypothesis is that we should observe a positive relationship between fiscal centralization and the size of government as a share of the economy. In other words, states with a larger share of government expenditures occurring at the state (relative to the local) level are expected to have a higher level of overall government spending as a share

*Luis Gonzalez, Valdosta State University; e-mail: Lgonzalez@valdosta.edu. Carlos Lopes, Eastern Kentucky University; e-mail: carlos.lopes@eku.edu.

of the economy. In order to test for this, many studies empirically examine the relationship between centralization of public spending and the government's share of the economy. Empirical tests of the Leviathan Hypothesis have produced mixed results (See Feld, Kirchgässner, and Schaltegger, (2010) for a review). In addition to conducting their own test, Feld, Kirchgässner, and Schaltegger (2010) review the empirical literature. However, some of the reviewed studies test other implications of the Leviathan Hypothesis and therefore may not be subject to the bias highlighted in this paper. We have hypothesized that by neglecting the effect of heterogeneity, some of these studies may obtain biased estimates. However, we find that only the omission of time-invariant heterogeneity influences the estimated impact of centralization. Our study is the first to test the relationship between centralization and the size of government as a share of the economy at the state level while also explicitly controlling for both time-variant and time-invariant heterogeneity of the different states.

Several papers show heterogeneity to be inversely correlated with centralization. As one example, Panizza (1999) finds that ethnically heterogeneous countries are less likely to be centralized than ethnically homogeneous countries. As a second example, Shelton (2007) examines a panel of countries from 1970 through 2000 and finds that population heterogeneity leads to decentralization. In another line of research, numerous authors find a relationship between the level of heterogeneity and government spending. Habyarimana, Humphreys, Posner, and Weinstein (2007) go so far as to say "the negative association between ethnic heterogeneity and public goods provision is widely accepted." Given that heterogeneity is likely to be correlated with government spending, which is our dependent variable, as well as with the level of centralization, which is our key explanatory variable, the estimated effects of centralization on government spending as a share of the economy should be biased whenever heterogeneity is omitted from these models. However, we find that this omission may only partially explain the mixed results commonly found in empirical examinations of the Leviathan Hypothesis.

The remaining portion of the paper is broken into five parts. Section II reviews the empirical literature. Section III provides a brief description of the potential bias of the estimated impact of centralization on government expenditures when heterogeneity is overlooked. Section IV provides an overview of the empirical identification strategy and data. Section V presents the results, and section VI concludes.

II. Literature Review

A. TESTS OF THE LEVIATHAN HYPOTHESIS

Empirical tests of the relationship between centralization and government size have a lengthy history, dating back at least as far as Adams (1965) who predates the Leviathan Hypothesis and finds a positive link between centralization and government spending. Though the Leviathan Hypothesis suggests that a centralized government may overspend, Adams interprets his findings under a different normative light, suggesting that a decentralized system will lead to the underprovision of public services. In any event, Adams was among the first in a long line of empirical tests that considered the impact of centralization on government spending. A comprehensive review of the literature since Adams (1965) is given by Feld et al. (2010), who document mixed results in this line of literature.

A distinctive feature of the Leviathan Hypothesis literature, which is highlighted in Feld et al. (2010), is that the category of government spending matters. Oates (1985) finds no evidence of the Leviathan Hypothesis in a test of the relationship between aggregate expenditures and centralization. However, Nelson (1987) theorizes that since education expenditures are such a large part of public expenditures, and since they typically occur at the local level, it is conceivable that a negative correlation exists between centralization and total government spending simply due to differences in preferences for education expenditures. This does not necessarily contradict the predictions of the Leviathan Hypothesis because there may still be a positive relationship between centralization and education expenditures when preferences for education are held constant. In other words, a Leviathan relationship should exist on a category-by-category basis. Raimondo (1989) tests Leviathan on a category-by-category basis and confirms the Leviathan Hypothesis for welfare expenditures.

B. HETEROGENEITY OF PREFERENCES AND LEVELS OF GOVERNMENT SPENDING

A sizable literature has documented inverse relationships between the heterogeneity of preferences and levels of government spending. One of the most common proxies for differences in preferences is differences in

ethnicity. Ethnic fractionalization is a popular measure of the heterogeneity of preferences constructed from differences in ethnicity. Ethnic fractionalization is a Herfindahl-style index defined as the probability that two individuals randomly selected from the population will be from different groups and will therefore have different preferences.[1] When using ethnic fractionalization as a proxy for preferences McCarty (1993) finds an inverse relationship between heterogeneity and government transfers. Alesina, Baqir, and Easterly (1999) use the same proxy in a more comprehensive study of U.S. local jurisdictions. They compare ethnic fractionalization to various categories of government spending and find that total spending rises as ethnic fractionalization increases.[2,3] However, they find that the share of government spending on what they define as "productive" categories tends to decrease as heterogeneity increases. The authors define productive categories as education, roads, sewers, and trash pickup. Miguel and Gugerty (2005) find that ethnic diversity is inversely related to education and water provision in Kenya. In his review of the literature, Putnam (2007) points out that ethnic diversity has been shown to have a negative effect on social trust and public good provision in developed countries. Luttmer (2001) finds that support for redistribution tends to fall in the presence of ethnic diversity.

Using income inequality as measured by the Gini coefficient as a secondary measure of heterogeneity, McCarty (1993) finds an inverse relationship between heterogeneity and levels of public spending in a small international cross-section. In this paper, we use both ethnic fractionalization and income inequality as measures of two separate types of time-variant constituent heterogeneity. Because we observe states over time, we use fixed effects estimation to address time-invariant heterogeneity.

C. HETEROGENEITY OF PREFERENCES AND THE CENTRALIZATION OF GOVERNMENT

Aside from being related to levels of government spending, heterogeneity of preferences is also linked to government centralization. Empirically, Wallis and Oates (1988) test the relationship between ethnic heterogeneity and centralization of not only government expenditures but also of government revenues for a panel of U.S. states from 1902 through 1982. The authors note that their results are "extremely sensitive to the

specification of the equation." Panizza (1999) finds evidence of an inverse relationship between heterogeneity and government centralization in a cross-section of 60 countries. However, he also notes that the effect of ethnic heterogeneity is "highly dependent upon the sample used."

Alesina and Wacziarg (1999) point out a possible endogenous relationship between the heterogeneity of preferences and centralization, which could partially explain the mixed empirical results found by Wallis and Oates (1988) and Panizza (1999). Strumpf and Oberholzer-Gee (2002) address this endogeneity and find evidence that heterogeneous states in the U.S. are more likely to decentralize the control of alcoholic beverages. Similarly, when using a large panel of 23 OECD countries, Stegarescu (2005) finds a positive relationship between heterogeneity and both tax and revenue decentralization.

In summary, if the Leviathan Hypothesis holds, then centralization has a positive direct effect on government spending. However, as stated earlier, empirical studies have documented an inverse relationship between centralization and heterogeneity. Also as stated earlier, heterogeneity is negatively related to preferences for public spending. Because centralization is linked to heterogeneity, and heterogeneity is linked to preferences for spending, tests of the effect of centralization on spending that do not control for heterogeneity could produce biased results. If heterogeneity is indeed negatively related to both centralization and spending, then the omission of heterogeneity should negatively bias the estimated impact of centralization. However, the inverse relationship documented between heterogeneity and centralization is empirical and may not hold universally. Therefore, we say the direction of the bias is *a priori* unclear.

III. Empirical Identification Strategy and Data

We use the following model to estimate the impact of centralization on government spending.

$$S_{ijt} = \beta_0 + \beta_1 C_{ijt} + \beta_2 H_{jt} + \beta_3 X_{jt} + \varepsilon_{ijt} \quad (1)$$

$$\varepsilon_{ijt} = \mu_{ij} + \nu_{ijt} \quad (2)$$

S_{ijt} represents the amount of government spending as a share of the economy in a particular category i in state j in year t. We measure

spending as a share of the economy by using spending in various categories as a share of state personal income. The categories of spending we consider are total government expenditures, education expenditures, welfare expenditures, health and hospital expenditures, highway expenditures, public safety expenditures, and utility expenditures. Public safety expenditures are the sum of police, fire, and corrections expenditures. C_{ijt} measures centralization as the state share of total state and local government expenditures in spending category i in state j in year t. H_{jt} is a proxy for the time-variant heterogeneity of preferences. We use two different measures of time-variant heterogeneity in addition to estimating models that exclude any measure of time-variant heterogeneity.

We follow the lead of McCarty (1993) by using ethnic fractionalization as one of our time-variant measures of the heterogeneity of preferences. Also following McCarty, we use the Gini coefficient as a secondary measure of time-variant heterogeneity of preferences. Annual state unemployment rates, poverty rates, and population estimates are contained in X_{jt}. The disturbance term in our model is composed of a fixed effect and an idiosyncratic error. We control for time-invariant heterogeneity by using state-level fixed effects. To investigate and determine the source of the proposed bias in the estimated impact of centralization, we also estimate a non-fixed effects model for each of the spending categories considered.

The parameters in Equation 1 are estimated using a 26-year panel from 1981 through 2006. Most of our variables come from the U.S. Census Bureau's internal database on historical finances of state and local governments known as "Rex-Dac".[4] We construct our measure of ethnic fractionalization using Census population estimates of state population by ethnicity. The state-level Gini coefficient data, which are used as our second measure of time-variant heterogeneity of preferences, are constructed from the Current Population Survey. Data on the state unemployment and poverty rates are taken from the national welfare dataset made available by the University of Kentucky Center for Poverty Research.

IV. Results

To investigate the extent to which bias is present in the existing literature, we begin by replicating the analysis of Oates (1985). Oates' seminal

article uses data from the 1977 Census of Governments to examine the effect of fiscal centralization on the fiscal size of government in order to "search for Leviathan." Oates measures total government spending, his dependent variable, as the sum of state and local tax receipts within a state as a share of state personal income. Since this measure always lies between 0 and 1, he also takes the logistic transformation to allow the dependent variable to range over the whole set of real numbers. Oates uses two fiscal centralization ratios to measure centralization, the key explanatory variable in his analysis. One of the ratios he uses is the ratio of state-level revenues to the sum of state and local revenues, and the other ratio is the ratio of state-level expenditures to the sum of state and local expenditures.

Oates fails to find support for the Leviathan Hypothesis regardless of which of the two fiscal centralization ratios he uses to measure centralization. He finds that as fiscal centralization increases, the size of government decreases. This is the opposite of what the Leviathan Hypothesis predicts. In both cases, the coefficients on Oates' measures of fiscal centralization are not statistically significant. However, Oates does not control for any type of heterogeneity of constituent preferences, whether time-variant or time-invariant.

Using the same cross-section of data from the 1977 Census of Governments, we replicate Oates' results below in Table 1, and we also repeat his analysis and include our two measures of heterogeneity of preferences as additional explanatory variables.[5] Consistent with our hypothesis, the estimated effect of centralization of general revenues changes from negative to positive when ethnic fractionalization is included in the model that uses states' share of revenues to measure fiscal centralization. When states' share of expenditures is used to measure fiscal centralization, the inclusion of the Gini coefficient reduces the magnitude of the negatively estimated effect of fiscal centralization, while the inclusion of ethnic fractionalization does not. Possibly due to the use of only a single cross-section of 48 states, none of the estimates of the effect of fiscal centralization in Table 1 is found to be statistically significant.

TABLE 1–Replication of Oates (1985) Using 1977
Census of Governments Data

Dependent Variable is Logistic Transformation of Total State and Local Tax Receipts as a Share of State Personal Income	Using Oates' First Centralization Ratio			Using Oates' Second Centralization Ratio		
	Baseline Replication of Oates (1985)	Including Ethnic Fractionalization as a Measure of Heterogeneity	Including Gini Coefficient as a Measure of Heterogeneity	Baseline Replication of Oates (1985)	Including Ethnic Fractionalization as a Measure of Heterogeneity	Including Gini Coefficient as a Measure of Heterogeneity
State Share of General Revenues	-0.053 (-0.17)	0.066 (0.22)	0.065 (0.20)	-	-	-
State Share of Total Expenditures	-	-	-	-0.289 (-0.78)	-0.317 (-0.89)	-0.227 (-0.374)
Personal Income Per Capita (thousands)	0.075 (2.01)**	0.054 (1.47)	0.086 (2.27)**	0.067 (1.84)*	0.040 (1.08)	0.075 (2.02)**
Population (millions)	0.88 (0.16)	6.66 (1.14)	2.41 (0.43)	-1.43 (-0.23)	3.71 (0.58)	0.14 (0.02)
Heterogeneity	-	0.439 (2.25)**	1.292 (1.31)	-	0.437 (2.29)**	1.156 (1.21)
Adjusted R^2						

Note: Observations = 48. t-statistics in parentheses. *significant at 10%; **significant at 5%; *** significant at 1%. Also included in the models but not shown in the table is a constant term.

Proceeding with our primary analysis, Tables 2a through 2g summarize the results from our estimation of Equation 1, where each of our sets of estimates is associated with a different category of government spending used to construct our dependent variable. We measure our dependent variable, the amount of government spending as a share of the economy, as the amount of state-level government spending in various categories as a share of state personal income. The categories of spending that we consider are total government expenditures, education expenditures, welfare expenditures, health and hospital expenditures, highway expenditures, public safety expenditures, and utility expenditures.[6] Following the lead of Oates (1985), Marlow (1988), and Joulfaian and Marlow (1990), our dependent variable has also undergone a logistic transformation. Furthermore, for each of the spending categories that we consider we estimate a non-fixed effects model, a fixed effects model, as well as fixed effects models that include one of our two measures of heterogeneity of preferences.

Since the nature of our investigation is to determine whether or not bias is present in tests of the Leviathan Hypothesis as a result of omitting measures of heterogeneity of preferences, we propose the following tests. We propose testing across models for the equality of the estimated coefficients on the centralization, personal income per capita, population, unemployment rate, and poverty rate variables. This test tells us whether the omission of unobserved time-invariant state-level factors biases estimates in tests of the Leviathan Hypothesis. We also perform the single test of the equality of the coefficients on the centralization variable across the two models. We then repeat our joint test and our single test of whether the estimated values of the coefficients from the fixed effects model and those from the fixed effects models that include state-level measures of ethnic fractionalization and the Gini coefficient are equal. We interpret rejection of the null hypotheses as evidence of bias. We perform all of these tests for models where our dependent variable derives from total government expenditures, education expenditures, welfare expenditures, health and hospital expenditures, highway expenditures, public safety expenditures, and utility expenditures. The p-values from our joint and single tests are located at the bottom of Tables 2a through 2g.[7]

We find consistent evidence of bias when time-invariant heterogeneity is overlooked. For example, in our joint tests we reject the null hypotheses that the coefficients on each of the centralization, personal income per capita, population, unemployment rate, and poverty rate variables are equal across the non-fixed effects and fixed effects models for each of the following spending categories: total government

expenditures, education expenditures, welfare expenditures, health and hospital expenditures, highway expenditures, public safety expenditures, and utility expenditures. Furthermore, all of the joint test null hypotheses are rejected at the 1 percent level of significance.

When performing our single tests, we reject the null hypotheses that the impact of centralization is identical when time-invariant heterogeneity is addressed in the estimation procedure versus when it is not addressed for each of the following spending categories: total government expenditures, education expenditures, public safety expenditures, and utility expenditures. The null hypotheses for the single tests for education and public safety expenditures are rejected at the 1 percent level of significance. The null hypotheses for the single tests for total government expenditures and utility expenditures are rejected at the 5 percent level of significance.

We find only limited evidence of bias when time-variant heterogeneity is omitted from tests of the Leviathan Hypothesis. In the case of welfare expenditures, the joint test that the coefficients on each of the centralization, personal income per capita, population, unemployment rate, and poverty rate variables are equal across the fixed effects and fixed effects plus ethnic fractionalization models is rejected at the 10 percent level of significance.

V. Conclusion

Empirical tests of the relationship between the level of government spending and centralization of public expenditures have produced mixed results. By omitting some measure of heterogeneity of preferences, previous studies may have produced biased results. This study produces new estimates of the effect of centralization while controlling for both time-invariant and time-variant forms of heterogeneity of preferences. We find consistent evidence of bias with respect to the omission from the empirical specification of time-invariant forms of heterogeneity. We find weak evidence of bias with respect to the omission of time-variant heterogeneity as measured by ethnic fractionalization. This weak evidence is present only for the spending category of welfare expenditures. We find no evidence of bias with respect to the omission of time-variant heterogeneity as measured by the Gini coefficient. A partial explanation is the fact that neither the ethnic fractionalization nor the Gini coefficient measure of time-variant heterogeneity varies substantially within a particular state over time, though there is considerable variation between states.

TABLE 2a–Analysis of the Impact of Centralization on Total Government Expenditures

Dependent Variable is Logistic Transformation of Total Government Expenditures as a Share of State Personal Income	No Measures of Time-Invariant Heterogeneity		Using State-Level Fixed Effects to Address Time-Invariant Heterogeneity	
	No Measures of Time-Variant Heterogeneity	No Measure of Time-Variant Heterogeneity	Including Ethnic Fractionalization as a Measure of Time-Variant Heterogeneity	Including Gini Coefficient as a Measure of Time-Variant Heterogeneity
Explanatory Variables	Model A	Model B	Model C	Model D
State Share of Total Expenditures	0.089 (0.52)	-0.114 (-1.60)	-0.124 (-1.77)*	-0.105 (-1.49)
Personal Income Per Capita (thousands)	0.002 (1.41)	-0.014 (-8.12)	-0.010 (-5.37)***	-0.016 (-8.56)***
Population (millions)	-0.007 (-0.77)	-0.002 (-0.65)	-0.001 (-0.36)	-0.002 (-0.70)
Unemployment Rate	-0.009 (-1.78)*	0.017 (7.08)***	0.020 (8.20)***	0.017 (6.88)***
Poverty Rate	0.014 (3.86)***	0.000 (0.35)	0.002 (1.74)*	-0.000 (-0.14)
Heterogeneity	-	-	-0.470 (-6.57)***	0.483 (2.75)***
Adjusted R^2	0.456	0.920	0.922	0.920
p-value for joint test of A vs B	-	0.0000**	-	-
p-value for single test of A vs B	-	0.0392**	-	-
p-values for joint test of B vs C and B vs D	-	-	0.1500	0.5645
p-values for single test of B vs C and B vs D	-	-	0.8388	0.4161

Note: Observations = 1,248. t-statistics in parentheses. * significant at 10%; ** significant at 5%;*** significant at 1%. Also included in the models but not shown in the table are a constant term, year indicators, and state-level fixed effects (in Models B, C, and D). p-values from joint tests indicate the lowest level of statistical significance at which the simultaneous equality across the models listed (e.g., across models A and B) of the estimated coefficients on each of the explanatory variables in the table is rejected. p-values from single tests indicate the lowest level of statistical significance at which the equality across the models listed of the estimated coefficients on the centralization variable is rejected.

TABLE 2b–Analysis of the Impact of Centralization on Education Expenditures

Dependent Variable is Logistic Transformation of Total Government Expenditures as a Share of State Personal Income	No Measures of Time-Invariant Heterogeneity		Using State-Level Fixed Effects to Address Time-Invariant Heterogeneity	
	No Measures of Time-Variant Heterogeneity	No Measure of Time-Variant Heterogeneity	Including Ethnic Fractionalization as a Measure of Time-Variant Heterogeneity	Including Gini Coefficient as a Measure of Time-Variant Heterogeneity
Explanatory Variables	Model A	Model B	Model C	Model D
State Share of Total Expenditures	0.128 (1.52)	0.295 (10.26)***	0.303 (10.60)***	0.291 (10.11)***
Personal Income Per Capita (thousands)	0.001 (0.75)	-0.004 (-2.33)***	-0.001 (-0.76)	-0.005 (-2.69)***
Population (millions)	-0.009 (-0.94)	0.002 (0.71)	0.003 (0.95)	0.002 (0.66)
Unemployment Rate	-0.017 (-3.02)***	0.012 (4.81)***	0.013 (5.44)***	0.011 (4.67)***
Poverty Rate	0.013 (3.2)***	-0.003 (-2.46)***	-0.002 (-1.57)	-0.004 (-2.73)***
Heterogeneity	-	-	-0.289 (-3.98)***	0.313 (1.78)*
Adjusted R^2	0.421	0.934	0.935	0.935
p-value for joint test of A vs B	-	0.0000**	-	-
p-value for single test of A vs B	-	0.0028**	-	-
p-values for joint test of B vs C and B vs D	-	-	0.6413	0.9158
p-values for single test of B vs C and B vs D	-	-	0.4665	0.3631

Note: Observations = 1,248. t-statistics in parentheses. * significant at 10%; ** significant at 5%;*** significant at 1%. Also included in the models but not shown in the table are a constant term, year indicators, and state-level fixed effects (in Models B, C, and D). p-values from joint tests indicate the lowest level of statistical significance at which the simultaneous equality across the models listed (e.g., across models A and B) of the estimated coefficients on each of the explanatory variables in the table is rejected. p-values from single tests indicate the lowest level of statistical significance at which the equality across the models listed of the estimated coefficients on the centralization variable is rejected.

TABLE 2c–Analysis of the Impact of Centralization on Welfare Expenditures

Dependent Variable is Logistic Transformation of Total Government Expenditures as a Share of State Personal Income	No Measures of Time-Invariant Heterogeneity		Using State-Level Fixed Effects to Address Time-Invariant Heterogeneity	
	No Measures of Time-Variant Heterogeneity	No Measure of Time-Variant Heterogeneity	Including Ethnic Fractionalization as a Measure of Time-Variant Heterogeneity	Including Gini Coefficient as a Measure of Time-Variant Heterogeneity
Explanatory Variables	Model A	Model B	Model C	Model D
State Share of Total Expenditures	-0.014 (-1.80)*	-0.008 (-1.52)	-0.008 (-1.49)	-0.008 (-1.52)
Personal Income Per Capita (thousands)	0.001 (24.68)***	-0.002 (-15.61)***	-0.002 (-12.72)***	-0.002 (-15.08)***
Population (millions)	0.0000 (0.03)	-0.0005 (-1.79)*	-0.0004 (-1.54)	-0.0005 (-1.79)*
Unemployment Rate	0.000 (0.83)	0.001 (3.31)***	0.001 (4.35)***	0.001 (3.33)***
Poverty Rate	0.000 (-0.96)	0.000 (-0.97)	0.000 (0.37)	0000 (-0.88)
Heterogeneity	-	-	-0.036 (-6.25)***	-0.006 (-0.43)
Adjusted R^2	0.766	0.891	0.895	0.891
p-value for joint test of A vs B	-	0.0000**	-	-
p-value for single test of A vs B	-	0.6934	-	-
p-values for joint test of B vs C and B vs D	-	-	0.0959*	0.9998
p-values for single test of B vs C and B vs D	-	-	0.8880	0.9346

Note: Observations = 1,248. t-statistics in parentheses. * significant at 10%; ** significant at 5%;*** significant at 1%. Also included in the models but not shown in the table are a constant term, year indicators, and state-level fixed effects (in Models B, C, and D). p-values from joint tests indicate the lowest level of statistical significance at which the simultaneous equality across the models listed (e.g., across models A and B) of the estimated coefficients on each of the explanatory variables in the table is rejected. p-values from single tests indicate the lowest level of statistical significance at which the equality across the models listed of the estimated coefficients on the centralization variable is rejected.

TABLE 2d–Analysis of the Impact of Centralization on Health and Hospital Expenditures

Dependent Variable is Logistic Transformation of Total Government Expenditures as a Share of State Personal Income	No Measures of Time-Invariant Heterogeneity		Using State-Level Fixed Effects to Address Time-Invariant Heterogeneity	
	No Measures of Time-Variant Heterogeneity	No Measure of Time-Variant Heterogeneity	Including Ethnic Fractionalization as a Measure of Time-Variant Heterogeneity	Including Gini Coefficient as a Measure of Time-Variant Heterogeneity
Explanatory Variables	Model A	Model B	Model C	Model D
State Share of Total Expenditures	-0.001 (-0.34)	0.001 (0.65)	0.002 (0.90)	0001 (0.70)
Personal Income Per Capita (thousands)	-0.0001 (-2.08)**	-0.0057 (-3.95)***	-0.0046 (-2.93)***	-0.0065 (-4.35)***
Population (millions)	0.0000 (0.10)	0.0001 (0.59)	0.0002 (0.67)	0.0001 (0.56)
Unemployment Rate	-0.001 (-3.92)***	0.000 (-0.36)	0.000 (0.01)	0.000 (-0.5)
Poverty Rate	0.001 (4.97)***	0.000 (2.47)***	0.000 (2.86)***	0.000 (2.05)**
Heterogeneity	-	-	-0.012 (-2.08)**	0.030 (2.11)**
Adjusted R^2	0.738	0.851	0.852	0.852
p-value for joint test of A vs B	-	0.0000**	-	-
p-value for single test of A vs B	-	0.1170	-	-
p-values for joint test of B vs C and B vs D	-	-	0.9548	0.8883
p-values for single test of B vs C and B vs D	-	-	0.5101	0.7089

Note: Observations = 1,248. t-statistics in parentheses. * significant at 10%; ** significant at 5%;*** significant at 1%. Also included in the models but not shown in the table are a constant term, year indicators, and state-level fixed effects (in Models B, C, and D). p-values from joint tests indicate the lowest level of statistical significance at which the simultaneous equality across the models listed (e.g., across models A and B) of the estimated coefficients on each of the explanatory variables in the table is rejected. p-values from single tests indicate the lowest level of statistical significance at which the equality across the models listed of the estimated coefficients on the centralization variable is rejected.

TABLE 2e–Analysis of the Impact of Centralization on Highway Expenditures

Dependent Variable is Logistic Transformation of Total Government Expenditures as a Share of State Personal Income	No Measures of Time-Invariant Heterogeneity		Using State-Level Fixed Effects to Address Time-Invariant Heterogeneity	
	No Measures of Time-Variant Heterogeneity	No Measure of Time-Variant Heterogeneity	Including Ethnic Fractionalization as a Measure of Time-Variant Heterogeneity	Including Gini Coefficient as a Measure of Time-Variant Heterogeneity
Explanatory Variables	Model A	Model B	Model C	Model D
State Share of Total Expenditures	0.030 (10.67)***	0.030 (13.25)***	0.030 (13.20)***	0.030 (13.17)***
Personal Income Per Capita (thousands)	-0.0004 (-12.87)***	-0.0002 (-1.86)*	-0.0002 (-1.47)	-0.0002 (-1.89)*
Population (millions)	0.001 (2.24)**	0.001 (3.43)***	0.001 (3.46)***	0.001 (3.43)***
Unemployment Rate	-0.0003 (-2.02)**	0.0002 (1.35)	0.0002 (1.45)	0.0002 (1.32)
Poverty Rate	0.0002 (2.39)***	0.0000 (0.30)	0.0000 (0.45)	0.000 (0.23)
Heterogeneity	-	-	-0.003 (-0.72)	0.004 (0.36)
Adjusted R^2	0.835	0.894	0.894	0.894
p-value for joint test of A vs B	-	0.0065**	-	-
p-value for single test of A vs B	-	0.7332	-	-
p-values for joint test of B vs C and B vs D	-	-	0.9998	0.9999
p-values for single test of B vs C and B vs D	-	-	0.7590	0.8180

Note: Observations = 1,248. t-statistics in parentheses. * significant at 10%; ** significant at 5%;*** significant at 1%. Also included in the models but not shown in the table are a constant term, year indicators, and state-level fixed effects (in Models B, C, and D). p-values from joint tests indicate the lowest level of statistical significance at which the simultaneous equality across the models listed (e.g., across models A and B) of the estimated coefficients on each of the explanatory variables in the table is rejected. p-values from single tests indicate the lowest level of statistical significance at which the equality across the models listed of the estimated coefficients on the centralization variable is rejected.

TABLE 2f–Analysis of the Impact of Centralization on Public Safety Expenditures

Dependent Variable is Logistic Transformation of Total Government Expenditures as a Share of State Personal Income	No Measures of Time-Invariant Heterogeneity		Using State-Level Fixed Effects to Address Time-Invariant Heterogeneity	
	No Measures of Time-Variant Heterogeneity	No Measure of Time-Variant Heterogeneity	Including Ethnic Fractionalization as a Measure of Time-Variant Heterogeneity	Including Gini Coefficient as a Measure of Time-Variant Heterogeneity
Explanatory Variables	Model A	Model B	Model C	Model D
State Share of Total Expenditures	-0.147 (-3.57)***	0.274 (13.67)***	0.279 (13.98)***	0.273 (13.62)***
Personal Income Per Capita (thousands)	0.006 (6.28)***	-0.026 (-14.53)***	-0.023 (-12.57)***	-0.026 (-14.27)***
Population (millions)	-.009 (-1.32)	-0.005 (-1.63)	-0.005 (-1.46)	-0.005 (-1.64)
Unemployment Rate	-0.021 (-4.98)***	0.014 (5.86)***	0.015 (6.41)***	0.014 (5.80)***
Poverty Rate	0.008 (2.82)***	-0.001 (-0.92)	0.000 (-0.17)	-0.001 (-1.00)
Heterogeneity	-	-	-0.245 (-3.51)***	0.088 (0.52)
Adjusted R^2	0.584	0.922	0.922	0.922
p-value for joint test of A vs B	-	0.0000**	-	-
p-value for single test of A vs B	-	0.0000**	-	-
p-values for joint test of B vs C and B vs D	-	-	0.3679	0.9990
p-values for single test of B vs C and B vs D	-	-	0.3114	0.6707

Note: Observations = 1,248. t-statistics in parentheses. * significant at 10%; ** significant at 5%;*** significant at 1%. Also included in the models but not shown in the table are a constant term, year indicators, and state-level fixed effects (in Models B, C, and D). p-values from joint tests indicate the lowest level of statistical significance at which the simultaneous equality across the models listed (e.g., across models A and B) of the estimated coefficients on each of the explanatory variables in the table is rejected. p-values from single tests indicate the lowest level of statistical significance at which the equality across the models listed of the estimated coefficients on the centralization variable is rejected.

TABLE 2g–Analysis of the Impact of Centralization on Utility Expenditures

Dependent Variable is Logistic Transformation of Total Government Expenditures as a Share of State Personal Income	No Measures of Time-Invariant Heterogeneity		Using State-Level Fixed Effects to Address Time-Invariant Heterogeneity	
	No Measures of Time-Variant Heterogeneity	No Measure of Time-Variant Heterogeneity	Including Ethnic Fractionalization as a Measure of Time-Variant Heterogeneity	Including Gini Coefficient as a Measure of Time-Variant Heterogeneity
Explanatory Variables	Model A	Model B	Model C	Model D
State Share of Total Expenditures	0.001 (0.39)	0.001 (-1.93)*	-0.001 (-1.72)*	-0.002 (-1.94)*
Personal Income Per Capita (thousands)	0.0004 (-2.55)***	0.0008 (1.77)*	0.0006 (1.23)	0.0009 (1.86)*
Population (millions)	-0.001 (-1.04)	-0.001 (0.75)	-0.001 (0.62)	0.001 (0.77)
Unemployment Rate	-0.001 (-1.37)	0.001 (0.59)	0.000 (0.48)	0.001 (0.62)
Poverty Rate	0.001 (2.16)**	-0.000 (-0.24)	-0.000 (-0.68)	-0.000 (-0.13)
Heterogeneity	-	-	0.030 (1.47)	-0.030 (0.51)
Adjusted R^2	0.439	0.789	0.789	0.789
p-value for joint test of A vs B	-	0.0011**	-	-
p-value for single test of A vs B	-	0.0243**	-	-
p-values for joint test of B vs C and B vs D	-	-	0.7768	0.9988
p-values for single test of B vs C and B vs D	-	-	0.4365	0.6663

Note: Observations = 1,248. t-statistics in parentheses. * significant at 10%; ** significant at 5%;*** significant at 1%. Also included in the models but not shown in the table are a constant term, year indicators, and state-level fixed effects (in Models B, C, and D). p-values from joint tests indicate the lowest level of statistical significance at which the simultaneous equality across the models listed (e.g., across models A and B) of the estimated coefficients on each of the explanatory variables in the table is rejected. p-values from single tests indicate the lowest level of statistical significance at which the equality across the models listed of the estimated coefficients on the centralization variable is rejected.

References

Adams, Robert F. 1965. "On the variation in the consumption of public services." *The Review of Economics and Statistics*, 47 (4): 400-405.

Alesina, Alberto, Reza Baqir, and William Easterly. 1999. "Public Goods and Ethnic Divisions." *The Quarterly Journal of Economics*, 114 (4): 1243-1284.

Alesina, Alberto, and Romain Wacziarg. 1999. "Is Europe going too far?" In *Carnegie-Rochester Conference Series on Public Policy* (Vol. 51, pp. 1-42). North-Holland.

Brennan, Geoffrey, and James M. Buchanan. 1980 *The Power to Tax: Analytical Foundations of a Fiscal Constitution*. Cambridge: Cambridge University Press.

Feld, Lars P., Gebhard Kirchgässner, and Christoph A. Schaltegger. 2010. "Decentralized taxation and the size of government: Evidence from Swiss state and local governments." *Southern Economic Journal* 77 (1): 27-48.

Habyarimana, James, Macartan Humphreys, Daniel N. Posner, and Jeremy M. Weinstein. 2007. "Why does ethnic diversity undermine public goods provision?" *American Political Science Review*, 101 (04): 709-725.

Joulfaian, David. and Michael L. Marlow. 1990, "Government Size and Centralization: Evidence from Disaggregated Data." *Southern Economic Journal* 56: 1094-1102.

Luttmer, Erzo F. 2001. "Group loyalty and the taste for redistribution." *Journal of Political Economy*, 109 (3): 500-528.

Marlow, Michael L. 1988, "Fiscal Decentralization and Government Size." *Public Choice* 56: 259-269.

McCarty, Therese A. 1993. "Demographic diversity and the size of the public sector." *Kyklos*, 46 (2): 225-240.

Miguel, Edward, and Mary K. Gugerty. 2005. "Ethnic diversity, social sanctions, and public goods in Kenya." *Journal of public Economics*, 89 (11): 2325-2368.

Nelson, Michael A. 1987, "Searching for Leviathan: Comment and Extension." *American Economic Review* 77: 198-204.

Oates, Wallace E. 1985, "Searching for Leviathan: An Empirical Study." *American Economic Review* 75: 748-757.

Panizza, Ugo. 1999. "On the determinants of fiscal centralization: Theory and evidence." *Journal of public economics*, 74 (1): 97-139.

Putnam, Robert D. 2007. "E pluribus unum: Diversity and community in the twenty-first century the 2006 Johan Skytte Prize Lecture." *Scandinavian political studies*, 30 (2): 137-174.

Raimondo, Henry J. 1989. "Leviathan and federalism in the United States." *Public Finance Review*, 17 (2): 204-215.

Shelton, Cameron A. 2007. "The size and composition of government expenditure." *Journal of Public Economics*, 91 (11): 2230-2260.

Stegarescu, Dan. 2005. "Public Sector Decentralisation: Measurement Concepts and Recent International Trends." *Fiscal studies*, 26 (3): 301-333.

Strumpf, Koleman. S., and Felix Oberholzer-Gee. 2002. "Endogenous policy decentralization: Testing the central tenet of economic federalism." *Journal of Political Economy*, 110 (1): 1-36.

Wallis, John. J., and Oates, Wallace E. 1988. "Decentralization in the public sector: An empirical study of state and local government." In *Fiscal federalism: Quantitative studies*: 5-32. University of Chicago Press.

Endnotes

1. Ethnic fractionalization is measured as $1-\sum_{i=1}^{k} r_i^2$, where k represents the number of ethnic categories and r_i represents the fraction of the overall population considered to be of ethnicity i.
2. See also Annett (2001), who finds that governments increase expenditures intended to promote political stability as ethnic fractionalization increases.
3. Using the same proxy for heterogeneity of preferences, Kuijs (2000) finds that health and education outcomes are lower in countries with higher levels of ethnic fractionalization, holding expenditures constant. This suggests that the quality of spending is also inversely related to the heterogeneity of preferences.
4. "Rex-Dac" contains complete information on revenues, expenditures, debts, and cash for state and local governments. The database can be found at: http://www2.census.gov/pub/outgoing/govs/special60/rex-dac.zip
5. Though our descriptive statistics are identical, our estimated coefficients are slightly different than those reported in Oates (1985). However, the sign and statistical significance of these estimated coefficients is identical.
6. Due to imperfections in the data on utility expenditures, our sample size is smaller for this category of expenditures than for the others. The panel is also unbalanced for this category of expenditures.
7. As mentioned in the notes to each of Tables 2a through 2g, the p-values from the joint tests indicate the lowest level of statistical significance at which the simultaneous equality across the models (e.g., across the non-fixed effects and fixed-effects models) of the estimated coefficients on each of the explanatory variables is rejected. The p-values from the single tests indicate the lowest level of statistical significance at which the equality across the models of the estimated coefficients on the centralization variable is rejected.