

## Appendix

### 1. *Robustness checks*

As the probit model used to estimate the selection equation of the hurdle model is unable to handle county dummy variables, in the main specifications I employ state dummy variables to control for unobserved heterogeneity between counties.

As a check for how the coefficient on instrumented lobbying changes with state dummies instead of county dummies, I estimated each individual part of the hurdle model separately: the selection part using a linear probability model, and the levels part using an OLS regression for counties receiving at least one earmark.

For the linear probability model, the coefficients on Lobbying were different in magnitude (0.026 for county effects, 0.009 for state effects), but not statistically different from each other.

Similarly, for the OLS regression with counties receiving at least one earmark, the magnitudes of the coefficients on Lobbying were different (0.159 for county effects, 0.12 for state effects), but not statistically different from each other.

I also estimated an additional specification (results available upon request) similar to the first stage equation from Table 5 with an added instrument: a housing prices growth rate squared term. This allows for the Sargan-Hansen test of overidentifying restrictions to be performed, where the null hypothesis is that the instruments are valid instruments (uncorrelated with the error term) and that the instrumental variables are appropriately excluded from the second stage equation. The estimated Sargan statistic was 0.262, with a P-value of 0.61 under the Chi-squared distribution, suggesting that the instruments in my empirical model are valid. See Hayashi (2000) for a discussion of the Sargan-Hansen test.

Table 5: Rate of Returns to Lobbying

Hurdle model with instrumented lobbying

	Selection	Levels	First stage
Lobbying	-0.069	0.17*	
	(0.06)	(0.094)	
HAC	0.866***	0.745***	0.275
	(0.214)	(0.22)	(1.0)
SAC	-0.193*	-0.092	-0.002
	(0.115)	(0.22)	(0.12)
House D.	0.016	-0.006	0.094
	(0.013)	(0.013)	(0.058)
House R.	0.013	0.001	0.219***
	(0.022)	(0.026)	(0.055)
Senate D.	0.001	0.003	-0.011
	(0.005)	(0.01)	(0.008)

Senate R.	-0.008	-0.018	0.013
	(0.008)	(0.016)	(0.012)
Vote-gap	-0.016***	-0.003	-0.009
	(0.003)	(0.005)	(0.009)
Housing Price Index ( % $\Delta$ )			-0.065***
			(0.007)
Obs.	8,122	8,122	8,122
Counties	2,714	2,714	2,714
IV F Statistic			81.69
Marginal effects for Lobbying			
Hurdle part	-0.015		
	(0.013)		
Combined	-0.155		
	(0.167)		
State Effects	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes

Note: \*\*\* denotes 1 percent confidence level, \*\* denotes 5 percent confidence level, \* denotes 10 percent confidence level. Standard errors in parentheses. First stage standard errors clustered at the state level. Specifications also include control variables not reported here: the growth rate in personal income per capita, change in unemployment rate, population growth rate, state-to-local aid, log of population, and log of personal income per capita. The “First stage” column displays results from the first stage of an instrumental variables 2SLS procedure for the second stage hurdle model results displayed in the columns “Selection” and “Levels.”

## 2. *Data Description*

### *Lobbying expenditures*

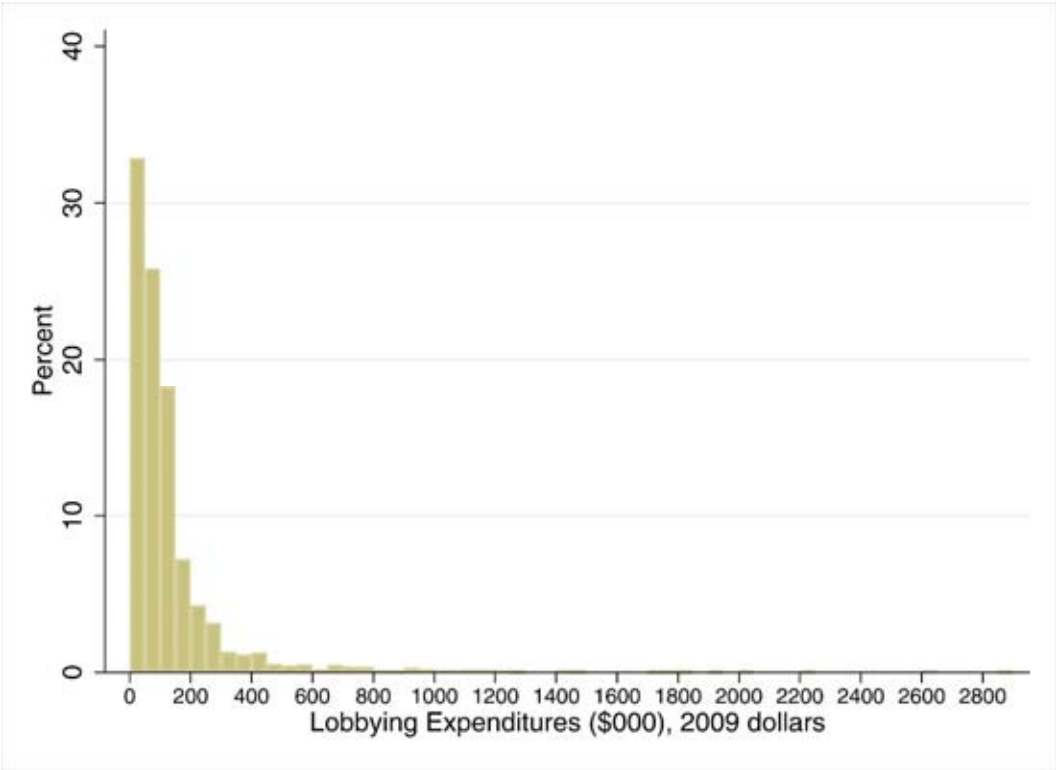
The data on lobbying expenditures come from the SOPR database. The Lobbying Disclosure Act of 1995 (LDA 1995) mandated that federal lobbyists semi-annually report their activities to the Secretary of the Senate and Clerk of the House of Representatives. All lobbying revenues greater than \$5,000 were to be reported by lobbying firms, and lobbying expenditures greater than \$20,000 for organizations with in-house lobbyists (“self-filers”). The law was later amended with the Honest Leadership and Open Government Act of 2007, which altered the monetary threshold and frequency of the reporting requirements. Under the amended law, the threshold was lowered to \$2,500 for lobbying firms (\$10,000 for self-filers), and reports were required quarterly instead of semi-annually. Additionally, reports were required to be submitted through a single, electronic system through the SOPR; prior to 2008, reports could be filed on paper, or through two different electronic reporting systems, one for the House and one for the Senate.

LDA 1995 defines a lobbyist as “any individual who is employed or retained by a client for financial or other compensation for services that include more than one lobbying contact, other than an individual whose lobbying activities constitute less than 20 percent of the time engaged in the services provided by such individual to that client over a six-month period.” Reports must be filed no later than 20 days after the lobbying-client relationship triggers one of several different requirements outlined by the law and must continue, regardless of the size of lobbying expenditures until the relationship is terminated.

The reports are assembled by SOPR staff into a database of records, where the name of the client, time period, and amount, in addition to other information, are recorded. I matched the client name with Census FIPS codes. In some cases, there were clear mistakes in the record that were easy to correct, such as misspelled names. In a few cases, the record was not able to be matched. Records with missing amounts were given the value of zero, although they could technically be any amount less than \$5,000. Clients that were listed as multiple governments were dropped from the analysis.

Figure 8 shows that roughly 30 percent of lobbying expenditures were \$50,000 or less.

Figure 8: Frequency of Lobbying Amounts.



*Earmarks*

In 2007, earmark reform began, ending with a ban on earmarks taking effect fiscal year 2011 (Doyle, 2011b). Before the ban took place, an Obama administration executive order directed the OMB to keep track of congressional earmarks contained in appropriations bills in order to improve transparency (Executive order 13457). This was to “establish a clear benchmark

for measuring progress.”<sup>1</sup> The OMB defined earmarks as “funds provided by the Congress for projects, programs, or grants where the purported congressional direction (whether in statutory text, report language, or other communication) circumvents otherwise applicable merit-based or competitive allocation processes, or specifies the location or recipient, or otherwise curtails the ability of the executive branch to manage its statutory and constitutional responsibilities pertaining to the funds allocation process.”

For the years 2005, 2008, 2009, and 2010, information on the legislation citation, description, and amount of each earmark is available. Most importantly, the recipient of each earmark is also noted. The advantage of the OMB earmark data over other earmark databases, such as those collected by nonprofit groups Taxpayers for Common Sense (TCS) and Citizens Against Government Waste (CAGW), is that the OMB required federal agencies to send in reports detailing their expenditures in relation to each earmark and list the recipients of these funds. The CAGW data do not list recipients, and the TCS data match earmarks to recipients by searching for the *intended* recipient in news releases by each earmark sponsor. For my purposes, the TCS data would be highly misleading, since the intended recipient of each earmark is often only one of many recipients of the actual funds. However, the OMB 2010 data do not include the recipients, which is unfortunate as it prevents these data from being included in my analysis.

Using the name of the recipient as reported in the OMB data, I matched earmarks with county and municipal governments. The decision to include only county and municipal governments was based primarily on the goal of consistency in regard to geographic boundaries of each observation. Special purpose governments and school districts often overlap county lines, and boundaries of the former are not available. Also, any earmark that listed multiple recipients was dropped.

Figure 9 shows that nearly 80 percent of earmarks to county or municipal governments were \$1 million or less, while Figure 10 shows that over half of the counties that received earmarks received 1 or less.

Figure 9: Frequency of Earmark Amounts.

---

<sup>1</sup> OMB, Press Release: New Features Added to Earmark Database; available from: <http://www.whitehouse.gov/sites/default/files/omb/assets/omb/pubpress/2007/07>: accessed 10 July 2007.

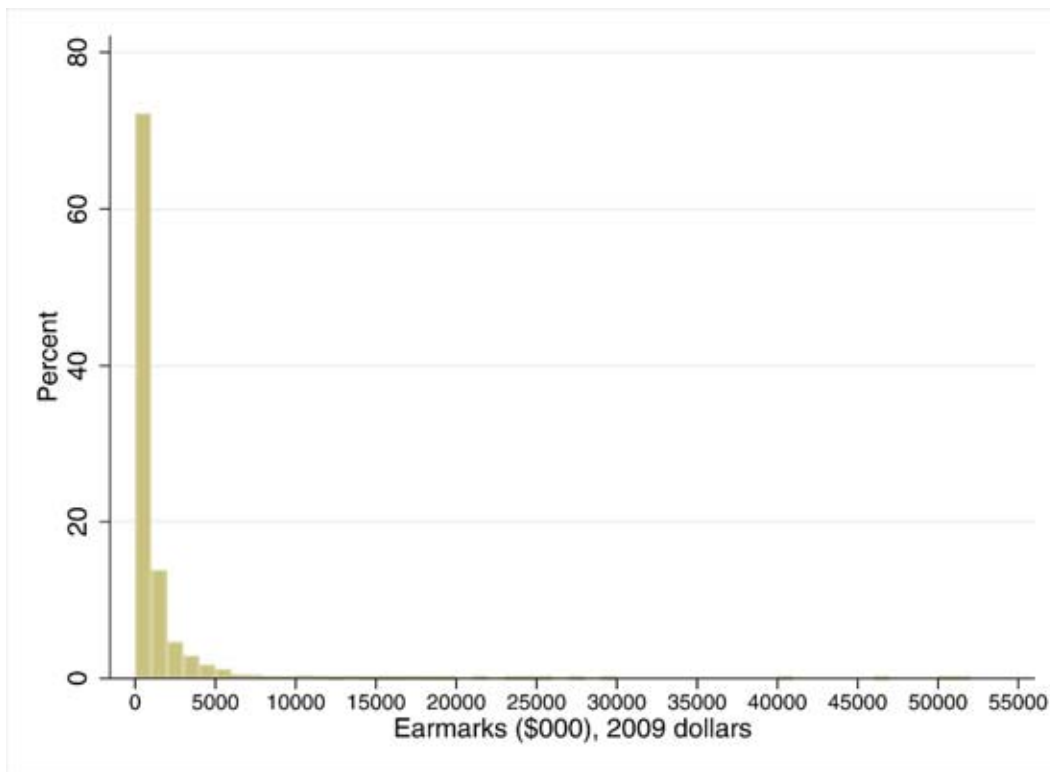
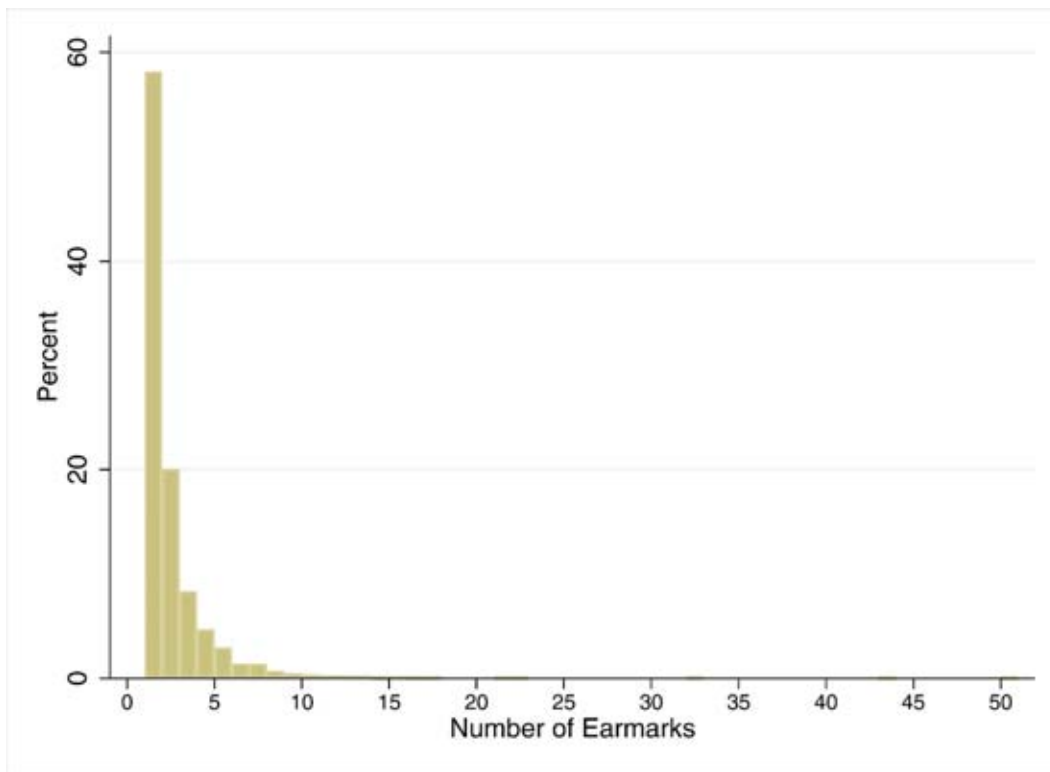


Figure 10: Frequency of Earmark Numbers.



*Congressional variables*

The Congressional variables come from Charles Stewart’s Congressional Data Page.<sup>2</sup> For the 109th Congress (2005-2006 data years), the Republicans outnumbered the Democrats 233 (25 freshmen) seats to 201 (16 freshmen) seats in the House, and 55 (7 freshmen) seats to 44 (2 freshmen) seats in the Senate. This reversed with the 110th Congress (2007-2008 data years)

<sup>2</sup> See: <http://web.mit.edu/cstewart/www/data/codebook.txt>

where Democrats had 233 (41 freshmen) and Republicans had 202 (13 freshmen) seats in the House and a tie at 49 (8 freshman Democrat, 1 freshman Republican, 1 freshman Independent) seats per party in the Senate.<sup>3</sup>

In terms of committee membership, the House Appropriations Committee (HAC) chairman was Jerry Lewis (R, CA) for the 109th Congress, and David R. Obey (D, WI) for the 110th Congress. The HAC was reorganized in 2007, increasing the number of subcommittees to 12, which gave each house an identical committee structure. The shift in majority party in the House was reflected in the party composition of the HAC; for the 109th Congress, Republicans outnumbered Democrats 38 to 29, while during the 110th Congress, Democrats outnumbered Republicans 37 to 30.

For the Senate Appropriations Committee (SAC), the chair during the 109th Congress was Thad Cochran (R, MS), which changed to Robert C. Byrd (D, WV) during the 110th Congress. During the 109th Congress there were 15 Republicans and 13 Democrats, while during the 110th Congress there were 15 Democrats and 14 Republicans.

I construct the vote-gap “political competition” variable in an analogous way to Levitt and Poterba (1999): it is the absolute value of the difference between the percent Democrat vote in each county and the national average in the most recent presidential election. Thus, for years 2004-2007, the number is the absolute difference from 48.3, while for years 2008 and 2009 it is the absolute difference from 52.9.

#### *Weighting procedure for House congressional variables*

Formally, I define a constituent-weighted House congressional variable ( $C_{it}$ ) at the county  $i$  level, based on congressional representative variables ( $y_{jt}$ ) at the congressional district  $j$  level in the following way:

$$C_{it} = \sum_{j=1}^{435} y_{jt} \alpha_{ij}$$

Where  $i$  indexes counties,  $j$  indexes the 435 congressional districts,  $t$  indexes the year, and  $\alpha_{ij}$  represents the constituent weight given to each county. For every congressional district  $j$ , split by  $H$  counties,  $\sum_{i=1}^H \alpha_{ij} = 1$ .  $y_{jt}$  represents each of the House variables as commonly defined: 0 or 1 for HAC representation, or the number of terms served in the House by political party for the tenure variables. The resulting House variables ( $HAC$ ,  $HouseD$ ,  $HouseR$ .) are weighted according to my method. The Senate variables, naturally, do not need to be weighted.

As an example, Kansas’s 4th Congressional District was represented by Republican Todd Tiahrt from 1995 to 2011. The fact that Tiahrt served on the HAC implies that the counties within his congressional district should receive positive values for the HAC variable. However, as Figure 11 shows, the 4th Congressional District varied greatly in terms of its population distribution across the counties it contains. Sedgwick County, which includes the city of Wichita, contains the majority of the district’s population. Thus, to assign a dummy variable equal to 1 for all 11 counties in the 4th district would

---

<sup>3</sup> The lone independent was Bernie Sanders of Vermont.

