Data Appendix

In this appendix we provide additional information on various segments of the data used to construct the lobbyist-level panel used in the paper. We also provide more details on the name matching procedure used to link the databases on lobbying reports and political employment.

Lobbying Reports

We use the LDA data sourced from the Senate Office of Public Records (SOPR) and compiled by the Centre for Responsive Politics (CRP) as part of their 'Open Secrets' database. The CRP is a non-profit and non-partisan organization with a stated mission of collating information on all types of politically related expenditures (i.e. campaign contributions, lobbying expenditures, member personal finances, etc.) and facilitating the availability of this data. We use the lobbying reports provided as part of their bulk data facility. This is the full universe of available LDA-sourced reports (approximately 35,000 per year) that CRP has formatted, cleaned and modified. For example, the CRP reconciles different types of reports (taking account of amendments to original mid-year and end-of-year reports) and constructs lobbyist, firm and client identifiers. We conduct further cleaning and consolidation of the CRP identifiers in cases where the same individuals are split across different identifier codes. The LDA requires the reporting of lobbying spending above a $10,000 threshold and rounded to the nearest $20,000. In the case of self-filing organizations, the LDA requires the reporting of all expenses made on lobbying activities, including payments to outside entities as well as in-house employees.

The lobbying reports data is collapsed to the lobbyist-period level. Revenue is aggregated by lobbyist-period according to the 'unweighted' and 'weighted' measures defined in Section 2.3 of the main paper. As a robustness check we re-estimated the main models and trimmed outliers at the 1 per cent and 2.5 per cent thresholds. This lead to only minor changes in coefficients and the results are available from the authors on request. Note that a manual check of outlier observations indicated that the majority of these high revenues belonged to well-known 'superstar' lobbyists.

Political Employment Database

Columbia Books Lobbyist Directory

Columbia Books publishes a comprehensive directory of Federal lobbyists under its suite of *Lobbyists.info* products. This directory initially began as a hard copy directory (titled *Washington Representatives*) containing contact information on lobbyists and potential clients published in the late 1970s. Since this time Columbia Books has expanded the directory with further information on career histories, biographical information, educational background and areas of expertise. The publisher then consolidated this directory in electronic form in 2006 as *Lobbyists.info* with daily updates and related supplementary databases.

This online version of the directory contains records on approximately 15,000 lobbyists. The career history information in *Lobbyists.info* includes the employer, job title and period of tenure for lobbyists' current and previous jobs, inclusive of private and public sector positions. We extract information on lobbyists who have had periods of political employment (that is, working as congressional staffers, in government agencies or as part of Presidential administrations) which is then matched by name into the CRP Lobbying Reports data.

LegiStorm *Congressional Staffer Salaries*

The second political employment database that we use is the LegiStorm *Congressional Staffer Salaries* (CSS) database. Based on Capitol Hill, LegiStorm is a company that aims to improve the availability of political information on the operations of the US Congress. For example, it provides easy-to-use versions of public data on Congressional remuneration; privately financed travel for Congress members and staff; financial disclosures; foreign gifts to members; and spending earmarks attached to bills.

The CSS database that we use is obtained by LegiStorm from published reports by the Secretary of the Senate and the Clerk of the House of Representatives. These reports are not actually made available in electronic form and LegiStorm takes the step of transferring the information from hard copy into an electronic format. As part of this process, LegiStorm also creates consistent set of identifiers for the staffers, offices and politicians that appear in the database.

LegiStorm's database contains information from late 2000 onwards. The main information provided

is: staffer name; start and end dates for a given employment spell; office of employment within Congress; the job title or position; and the total salary amount for a given job spell. The full staffers database contains information on approximately 90,000 staffers over a nine-year period. This large number of staff is due to the inclusion of non-partisan institutional staff such as Capital Police. Our analysis focuses on the pool of approximately 40,000 staffers working in political or policy related offices over the 2000-2008 period.

Name Matching of Lobbyists

The full list of lobbyists represented in the CRP lobbying reports database is matched with individuals appearing in our two political employment databases, *Lobbyists.info* from Columbia Books and *Congressional Staffer Salaries* from LegiStorm. The name matching is implemented using a string-based algorithm which cleans the raw names for punctuation and shortened names (for example, "JIM" becomes "JAMES" and so on). The same algorithm is applied to each set of names and each political employment database is separately matched with the CRP lobbying reports list. The subsequent matches are then compiled into one list of political employees-turn-lobbyists. Middle names or initials are used as part of the name matching procedure where available.

We then score matches on their accuracy according to two criteria. Firstly, each match receives a 1-4 score based on how often a particular first or last name appears. We call this a 'name frequency' score. Commonly occurring names such as 'SMITH' are given a score of 4 while the least common names are given a score of 1. This process is repeated separately for first and last names to produce a 2-8 score. For example, a name such as 'JOHN SMITH' is given an overall score of 8 since it is comprised of two common names while 'MILLICENT SMITH' receives a score of 5 (+4 for the common last name but +1 for the relatively uncommon first name).

Secondly, we score the matches according to how well the timing of employment transitions links up across the data. Staffers leaving employment in the Congress should appear in the lobbying reports data within 1-2 years of their final employment spell. We construct a 0-1 flag for whether the timing of the transitions is consistent across the data. In the final stage of the matching we then manually check the accuracy of the matches characterised by very common names and/or inconsistent timing. We do this by manually cross-referencing names with online CVs and biographies. This final step of manual

checks is done for all names with a name frequency score above 5.

Congressional Politicians

Our final major dataset contains information on the service and characteristics of politicians in the House and Congress since the beginning of the available lobbying reports in 1998. The specific data used is Stewart and Woon's (2009) compilation focusing on committee membership and updated periodically from the Congressional Record. This membership data here contains periods of service and reasons for exit (retirement, defeat for re-election etc) where applicable. The politicians appearing in the data have been allocated the ICSPR 'member id' that is common across political science studies in this area. We have name matched the list of politicians given in the political employment databases against the Stewart and Woon (2009) list using the same string-based algorithm developed for the lobbyist-level matching.

Table A2: Average Effects of Revolving Door Connections on Lobbying Revenue
Weighted Measure of Revenue

	Dependent Variable: (Log) Weighted Revenue per Lobbyist			
	(1)	**(2) plus Party**	**(3) plus Chamber**	**(4) plus Experience**
# of Senators:	.20** (.08)	.19*** (.07)	.20** (.08)	.24*** (.07)
# of Representatives:	-.01 (.05)	-.03 (.05)	-.03 (.05)	-.01 (.05)
Individual Dummies	Yes	Yes	Yes	Yes
Time	Yes	No	No	No
TimeXParty	No	Yes	No	No
TimeXPartyXChamber	No	No	Yes	Yes
Lobbyist Experience	No	No	No	Yes
Individuals	1,113	1,113	1,113	1,113
Observations	10,418	10,418	10,418	10,418

Note: This table presents the average effects of political connections on ex-staffers lobbying revenue. The dependent variable is the log of the **weighted revenue** generated from all the clients that an individual lobbyist serves in a time (semester) period. The two main independent variables are the number of Senators and Representatives that an individual lobbyist worked for previously to entering the lobbying industry *and are serving in Congress in that time period*. All regressions use a sample containing ex-staffers turned lobbyists and include both individual lobbyist dummies and time effects (i.e. semester dummies). Column (2) allows for different time effects for lobbyists connected to politicians in different parties (i.e. Democrats versus Republicans). Columns (3) and (4) allow for different time effects for lobbyists connected to politicians in different party/chamber combinations (i.e. Democrats in the Senate, etc.). Column (4) includes lobbyist experience (i.e. number of periods that a lobbyist appears in the sample) in quadratic form. Standard errors are clustered by lobbyist.

Table A3: Effects of Party Control

Dependent Variable: (Log) Revenue per Lobbyist		
	(1)	**(2) Decomposition**
# of Senators:	.25*** (.07)	.25*** (.07)
# of Representatives:	.12** (.05)	.11** (.05)
House Dem. Control X Dem.	.18*** (.06)	.19*** (.06)
Senate Dem. Control X Dem.	.14*** (.05)	
2001-2002 X Dem.		.08 (.08)
2007-2008 X Dem.		.18*** (.06)
Individual Dummies	Yes	Yes
Time	Yes	Yes
Lobbyist Experience	Yes	Yes
Individuals	1,113	1,113
Observations	10,418	10,418

Note: This table presents the effects of a party's control of Congress on the revenues of its affiliated lobbyists. One period of Democrat control in the House is considered (2007-2008) along with two periods of Democrat control in the Senate (2001-2002 and 2007-2008). In the second column we include the two periods of Democrat control in the Senate separately. We interact these party control dummies with whether the lobbyist is an ex-staffer for a Democratic politician. We also include the main variables in Table 2 as well as individual dummies, time dummies and lobbyist experience (and its square). Standard errors are clustered by lobbyist.

Table A4: Robustness Tests

Dependent Variable: (Log) Revenue per Lobbyist

	(1) Party in Control X Ind. Dum.	(2) Time Trends	(3) Firm Dummies
# of Senators:	.26** (.12)	.19* (.11)	.28*** (.08)
# of Representatives:	.12 (.09)	.03 (.08)	.04 (.06)
Individual Dummies	Yes	Yes	Yes
TimeXPartyXChamber	Yes	Yes	Yes
Lobbyist Experience	Yes	Yes	Yes
Ind. Dum. X Party in Control	Yes	No	No
Lobbyist Time Trends	No	Yes	No
Firm Dummies	No	No	Yes
Individuals	1,113	1,113	1,113
Observations	10,418	10,418	10,418

Note: This table presents a number of robustness tests on Table 2. The main independent variables are as in Table 2. Every regression contains individual dummies, lobbyist experience in quadratic form, and different time effects for lobbyists connected to politicians in different party/chamber combinations. Column (1) includes the party control dummies used in Table A3. These dummies are interacted with the individual lobbyist dummies. Column (2) introduces 1,113 lobbyist-specific linear time trends. Column (3) includes 726 lobbying firm dummies. Standard errors are clustered by lobbyist.

Table A5: Effects of Entry into Committees

Dependent Variable: (Log) Revenue per Lobbyist

	Senate		House	
	(1)	**(2) plus Party-Chamber and Experience**	**(3)**	**(4) plus Party-Chamber and Experience**
One Senator:				
in Finance	.44*** (.14)	.38*** (.14)		
in Appropriations	.12 (.36)	-.06 (.36)		
One Representative:				
in Ways & Means			.46** (.22)	.44** (.22)
in Appropriations			-.17 (.23)	-.21 (.23)
Dummies 2 Senators	Yes	Yes	No	No
Dummies 2 Representatives	No	No	Yes	Yes
TimeXPartyXChamber	Yes	Yes	Yes	Yes
Lobbyist Experience	No	Yes	No	Yes
Individuals	684	684	818	818
Observations	4,792	4,792	5,149	5,149

Note: This table presents the effects of Congressmen joining important committees on the revenue of their connected lobbyists. The dependent variable is as in Table 2. The Senate regression sample contains only lobbyists connected to serving Senators. The displayed variable Finance takes value one when the connected and serving Senator has joined the Finance committee. We define similarly the variable Appropriations. The omitted group is being connected to **one** serving Senator in neither of these two Committees. For lobbyists connected to two serving Senators we also define and include dummies capturing whether one connected Senator joined the Finance or Appropriations committees. In practice, there are very few such cases and the estimated parameters are not displayed. The House sample and regressions are constructed equivalently. All regressions include individual lobbyists dummies and time effects (i.e. semester dummies). Columns (2) and (4) allow for different time effects for lobbyists connected to politicians in different party/chamber combinations (i.e. Democrats in the Senate, etc.) and also include lobbyist experience (i.e. number of periods that a lobbyist appears in the sample) in quadratic form. Standard errors are clustered by lobbyist.

Table A6: Robustness in Participation Regressions

	Dependent Variable: Participation Dummy	
	(1) Logit	**(2) LPM Balanced Sample**
# of Senators:	1.72*** (.30)	.24*** (.05)
# of Representatives:	.48** (.19)	.07** (.03)
Individual Dummies	Yes	Yes
Time	Yes	No
TimeXPartyXChamber	No	Yes
Lobbyist Experience	Yes	Yes
Individuals	1,113	1,113
Observations	13,473	24,486

Note: This table presents two robustness tests of the participation regressions in Table 4. The dependent variable is as in Table 4. Column (1) displays the results of a logit regression using the same sample as in Table 4. Column (2) displays the results of a OLS regression using a balanced panel, in which every individual appears during 22 periods (i.e. between 1998 and 2008). All regressions include individual dummies, time effects and lobbyist experience in quadratic form. Column (2) allows for different time effects for lobbyists connected to politicians in different party/chamber combinations (i.e. Democrats in the Senate, etc.). Standard errors are clustered by lobbyist.

Table A7: Accounting for Endogenous Participation Bounds Exercise

Dependent Variable: (Log) Revenue per Lobbyist		
Imputation using Last Period Non-Zero Revenue		
	(1) Baseline	**(2) With Imputation**
# of Senators:	.24***(0.07)	.29***(0.06)
# of Representatives:	.10*(.05)	.08*(.05)
Individual Dummies	Yes	Yes
TimeXPartyXChamber	Yes	Yes
Lobbyist Experience	Yes	Yes
Individuals/Observations	1,113/10,418	1,113/13,637

Note: This table provides in Column (2) a bounds exercise to account for endogenous participation in the estimation of the average effects of political connections on ex-staffers lobbying revenue. The dependent variable is as in Table 2. The sample in Column (1) is as in Table 2. The sample in Column (2) comprises every period following an individual's first period in the lobbying industry. For periods in which revenue is zero, the last non-zero revenue is used. The two main independent variables are as in Table 2. All regressions include individual lobbyist dummies, different time effects for lobbyists connected to politicians in different party/chamber combinations and lobbyist experience in quadratic form. Standard errors are clustered by lobbyist.

Table A1: Descriptive Statistics – In-House Sector, 1998-2008.

Panel A - Organization Level	
Mean Number of Lobbyists	1.5
Mean Expenditures	786.1
Total Number of Organizations	3233
Panel B - Types of Lobbyists	
Revolving Door Lobbyist	0.185
Ex-Congressman	0.004
Ex-staffer:	0.137
- of politician serving pre-1998	0.017
- of politician serving post-1998	0.101
- of a congressional committee	0.019
Outside Congress	0.044
Panel C - Mean Expenditure, by Type of Lobbyist	
Weighted	
Revolving Door Lobbyists	292.5
-Ex-Congressmen	368.4
-Ex-staffers	280.5
-Outside Congress	323.1
Other lobbyists	204.9
Unweighted	
Revolving Door Lobbyist	2,319.8
-Ex-Congressmen	2,134.0
-Ex-staffers	2,287.5
-Outside Congress	2,438.3
Other lobbyists	1,752.1
Panel D - Share of Total Industry Expenditure, by Type of Lobbyist	
Revolving Door Lobbyist	0.222
-Ex-Congressmen	0.005
-Ex-staffers	0.159
-Outside Congress	0.060
Other lobbyists	0.778
Average Number of Periods	6.5
Number of Lobbyists	21,374
Number of Lobbyist-Period Observations	127,960

Note: Panel A based on 1998-2008 panel of organizations by period. Panels B, C and D based on 1998-2008 lobbyist-level panel. Length of each period is 6 months. Panel C presents annualised measures of expenditure per lobbyist. Panel D aggregates the weighted expenditures of lobbyists in order to calculate expenditure shares by type. 'Ex-Congressman' denotes former members of the House or Senate who are lobbyists. 'Ex-staffer' represents lobbyists who have employment experience as Congressional staffers. Congressional committee ex-staffers were employed in committee offices but not as personal staff to politicians. 'Outside Congress' lobbyists represents lobbyists who have experience as government employees in workplaces outside of the Congress.