Online Appendix to Political Alignment and Bureaucratic Pay

Jon H. Fiva* Benny Geys† Tom-Reiel Heggedal‡ Rune Sørensen§

November 13, 2020

*BI Norwegian Business School, Department of Economics, Nydalsveien 37, N-0484 Oslo, Norway. E-mail: jon.h.fiva@bi.no

†BI Norwegian Business School, Department of Economics, Kong Christian Frederiks plass 5, N-5006 Bergen, Norway. E-mail: benny.geys@bi.no

‡BI Norwegian Business School, Department of Economics, Nydalsveien 37, N-0484 Oslo, Norway. E-mail: tom-reiel.heggedal@bi.no

§BI Norwegian Business School, Department of Economics, Nydalsveien 37, N-0484 Oslo, Norway. E-mail: rune.sorensen@bi.no

Table of contents

A Principal-agent model

In this section, we formally analyze how political preference alignment between a principal (politician) and an agent (CMO) affects the agent's expected pay. Preference alignment is thereby understood as a similarity along relevant preference dimensions between principal and agent (see also below). We focus on two potential underlying mechanisms. First, preference alignment gives policy-motivated agents a direct stake in achieving the public output desired by the political principal (the political mission). This is equivalent to the assumptions on policy motivation by Besley and Ghatak (2005). Second, preference alignment works to streamline communication and facilitates cooperation between contracting partners, and thereby improves the productivity of a match. This notion of productivity in a match is central to the literature on the *ally principle* (Bendor, Glazer and Hammond 2001; Huber and Shipan 2008; Dahlström and Holmgren 2019), and its micro-foundations – including improved communication, cooperation and control – have been extensively debated in the foregoing literature (Peters and Pierre 2004; Kopecky et al. 2016; Toral 2019).

Although our empirical setting is characterized by the absence of one-off bonus payments, the model developed below nonetheless builds on an incentive payment scheme. This is done to maintain comparability with previous work on performance contracts (e.g., Besley and Ghatak 2005), but it naturally implies that the model represents an imperfect analogy in our setting. To tie the model closer to the empirical application, however, one could think of the wage contract as consisting of a permanent wage plus an incentive-based, performance-related increase of this wage lasting over the entire employment spell. The net present value of this permanent wage increase would correspond to the value of the bonus, leaving the per-period incentives in the model unchanged.[1] The model thus can give us a sense of expected pay *levels* under different conditions, as well as providing additional clarity on how our two mechanisms push bureaucratic pay

[1] Note also that a performance contract with an output-related bonus may equivalently be set up as a fixed wage contract with a dismissal probability related to output. By convention, we discuss mechanisms using the former, while appreciating that the latter may better fit with our empirical setting.

in opposite directions. In the main model, we focus on the moral hazard problem in the agency relation and incentive pay. However, we also show that our main theoretical propositions persist in a framework without information problems.

A.1 Matching and political alignment

Assume that a principal P and an agent B are randomly matched and choose whether to produce together. We denote the principal's and agent's party-types by P_i and B_i, respectively, where i is either L (left) or R (right). One can think of the policy placement of both players as being measured on the unit interval with L-types placed at 0 and R-types placed at 1. Since preference alignment implies that principal and agent are similar along relevant preference dimensions, we say that political preferences are aligned when $A = 1-|P_i-B_i| = 1$, while political preferences are not aligned when $A = 1-|P_i-B_i| = 0$.

A.2 Technology and preferences

If the principal and agent in a match decide not to produce together, they receive their outside options (represented by u). If they do produce together, let output be given by $f = (a + A\lambda)e + \varepsilon$ where $a > 0$ is a productivity parameter and $\lambda \geq 0$ is the match specific productivity parameter that interacts with preference alignment A. The agent's effort is e, and ε is a stochastic element (noise). Agents' cost of effort is given by $c = \frac{e^2}{2}$.[2]

An agent's utility U is increasing in the expected financial value of the contract w (pay), and decreasing in the risk associated with the contract as well as cost c.

$$U = E[w] - 0.5rVar[w] - c + A\theta E[f],$$

where $r > 0$ measures the degree of risk aversion. The agent's utility may also be directly and positively affected by output f, depending on her potential intrinsic motivation for

[2]Note that our results go through if we let alignment affect productivity in a match through the cost of effort. For instance, assuming cost of effort is given by $c = \frac{e^2}{2t}$, where the parameter $t \in (0, \infty)$ represents agents' cost-type, provides similar inferences to those reported below.

achieving output f (reflected in parameter $\theta \in [0, 1)$).

The principal is risk neutral with utility

$$\pi = E[f] - E[w].$$

Lastly, we assume that the principal can observe the agent's party-type, as these are mapped by party affiliation in the empirical application.

A.3 Optimal performance contract

In this section, we assume that the principal cannot observe effort and, hence, effort is not contractible (note that the next subsection solves the model assuming that the principal can observe effort). Restricting the analysis to linear contracts, let a contract w be given by

$$w = \tau + kf,$$

where τ is a fixed transfer and k is a fraction of output (the incentive part – or 'power' – of the contract). We analyze the optimal contract in two cases: in the first case agents are not motivated by policies (i.e., $\theta = 0$), while in the second case agents have the same productivity in all matches (i.e., $\lambda = 0$).

Looking first at the case without policy motivation, agents' utility of a contract is given by

$$U = E[w] - 0.5rVar[w] - c.$$

Inserting for w and f we get

$$U = \tau + k(a + A\lambda)e - 0.5rk^2Var[\varepsilon] - c.$$

The agent maximizes U with respect to effort e. This gives rise to the incentive compatibility constraint facing the principal

$$k(a + A\lambda) = c'.$$

This equality implies that for a given k, agents in an aligned match ($A = 1$) will put in more effort than in a non-aligned match ($A = 0$) that has lower productivity. However, it is not optimal for the principal to give the same incentives k to agents of different alignment-types. In fact, it is straightforward to show that the incentive part of the optimal contract following from the principal's maximization problem (taking the incentive compatibility and participation constraints as given) is[3]

$$k = \frac{(a + A\lambda)^2}{(a + A\lambda)^2 + rVar(\varepsilon)}.$$

The optimal output-related pay k thus increases in agents' productivity in a match (which is higher for agents in aligned matches). The intuition is that the principal wants to incentivize the high productivity agents more than other agents, and these agents need to be compensated for taking on more risk (and suffering from the induced larger effort). Thus, agents in aligned matches demand higher expected pay to participate. The following proposition summarizes this result.

Proposition 1. *Suppose $\theta = 0$ and $\lambda > 0$. Then politician-bureaucrat preference alignment increases bureaucrats' expected pay, i.e., $E[w|A = 1] > E[w|A = 0]$.*

Now, what happens when we allow for policy-motivated agents? In this case we let $\theta > 0$ while $\lambda = 0$. The agent's utility of a contract w is then be given by

$$U = \tau + (A\theta + k)ae - 0.5rk^2Var[\varepsilon] - c.$$

[3]The participation constraint simply states that the value of the contract to the agent must at least equal her outside option, i.e. it must satisfy $U = E[w] - 0.5rVar[w] - c = u$.

It follows that the agent's first order condition is

$$(A\theta + k)a = c'.$$

Thus, for a given k, the agent puts in more effort when the match is aligned ($A = 1$). Equivalently, the same level of effort can also be achieved with a lower k, although this does not constitute an optimal contract. In fact, it turns out that the optimal k is the same to agents of different alignment-types. This result stems from the fact that the change in effort induced by changing k is the same for all effort levels when c'' is constant (i.e., $de/dk = a$ is invariant to θ). The principal's trade-off when increasing k between the marginal gain in production (through effort) and the marginal cost of risk shifting, is then the same for all agents regardless of the match's (mis-)alignment.

Turning to expected pay, note first that – for given transfer τ – the surplus is larger for agents in aligned matches. Both aligned and misaligned agents face the same risk and get the same performance pay for a given output level. However, the aligned agents also have a direct stake in output. Moreover, these agents work harder creating additional surplus for themselves. That this indeed is a surplus, follows from that these agents could choose the same effort level as the misaligned ones, but optimally choose higher effort for given k. Next note that the participation constraint is given by

$$\tau + (A\theta + k)ae - 0.5rk^2Var[\varepsilon] - c = u.$$

The principal extracts the aforementioned surplus by lowering the fixed transfer τ for the aligned agents until the participation constraint binds. Thus, τ is lower for these agents than for the misaligned ones. It is then straightforward to show that also total expected pay – i.e., $\tau + kae$ – is lower.[4] That is, the reduction in τ is larger than the additional pay to the aligned agents through the performance part of the contract (due to higher effort). The intuition is that the cost accrued from working harder is more than covered by the

[4]For given k, a sufficient condition for the result is that c'' is non-decreasing.

benefit obtained from their stake in the output. The effect of being policy-motivated thus is to lower agents' pay. The following proposition summarizes this result.

Proposition 2. *Suppose $\theta > 0$ and $\lambda = 0$. Then politician-bureaucrat preference alignment decreases bureaucrats' expected pay, i.e., $E[w|A = 1] < E[w|A = 0]$.*

A.4 No moral hazard

In this section, we show that our main results on the effects of preference alignment can be replicated without the moral hazard problem. In this situation, the principal can observe effort, and in the optimal contract the principal bears all risk associated with the stochasticity of output ε. The optimal contract consists of the fixed transfer τ and a given effort level $\bar{e}$, where the principal chooses $\bar{e}$ to maximize total surplus:

$$\pi + U = E[f] - c + A\theta E[f],$$

where $f = (a + A\lambda)e + \varepsilon$ as in the previous sections.

Looking first again at the case without policy motivation, we get the following first-order condition:

$$(a + A\lambda) = c'.$$

This condition implies that the optimal effort level in an aligned match ($A = 1$) is higher than in a non-aligned match ($A = 0$) that has lower productivity. Agents in aligned matches need to be compensated for this higher effort in order to participate. Thus, τ is larger when $A = 1$, replicating the result on productivity from the previous section.

Allowing for policy motivation (i.e., setting $\lambda = 0$ and $\theta = 1$), we get the following first-order condition:

$$(A\theta + 1)a = c'.$$

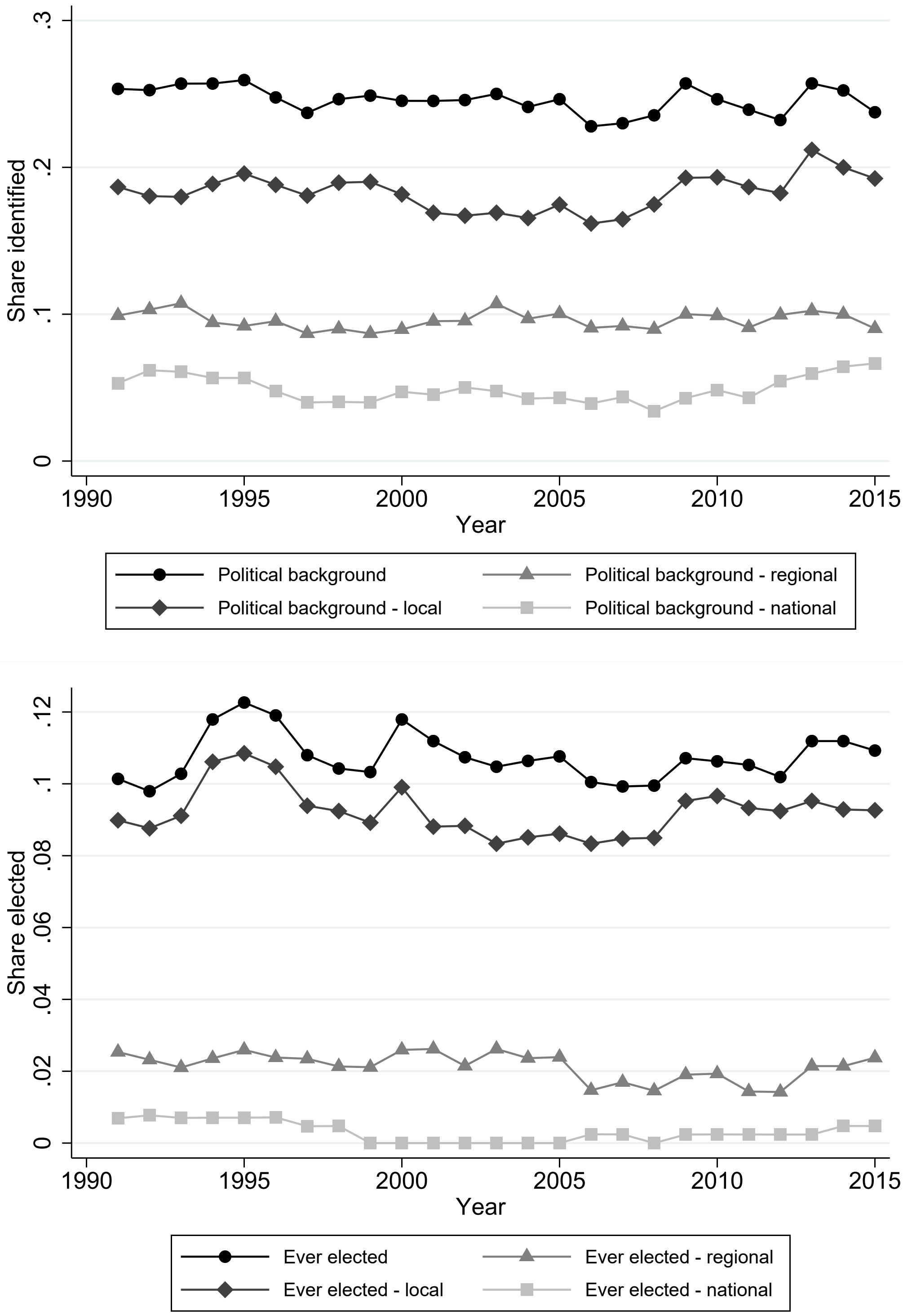
This condition again implies that the optimal effort level in an aligned match ($A = 1$) is higher than in a non-aligned match ($A = 0$). Turning to the wage, note that the participation constraint is given by:

$$\tau + (A\theta)ae - c = u.$$

As in the previous section, the surplus is larger for agents in aligned matches for a given τ. That is, the benefit of these agent's stake in the output is larger than the additional cost following from the larger optimal effort level. This surplus is extracted by the principal. Thus, τ is smaller when $A = 1$, replicating the result on policy motivation from the model with moral hazard in the previous section.

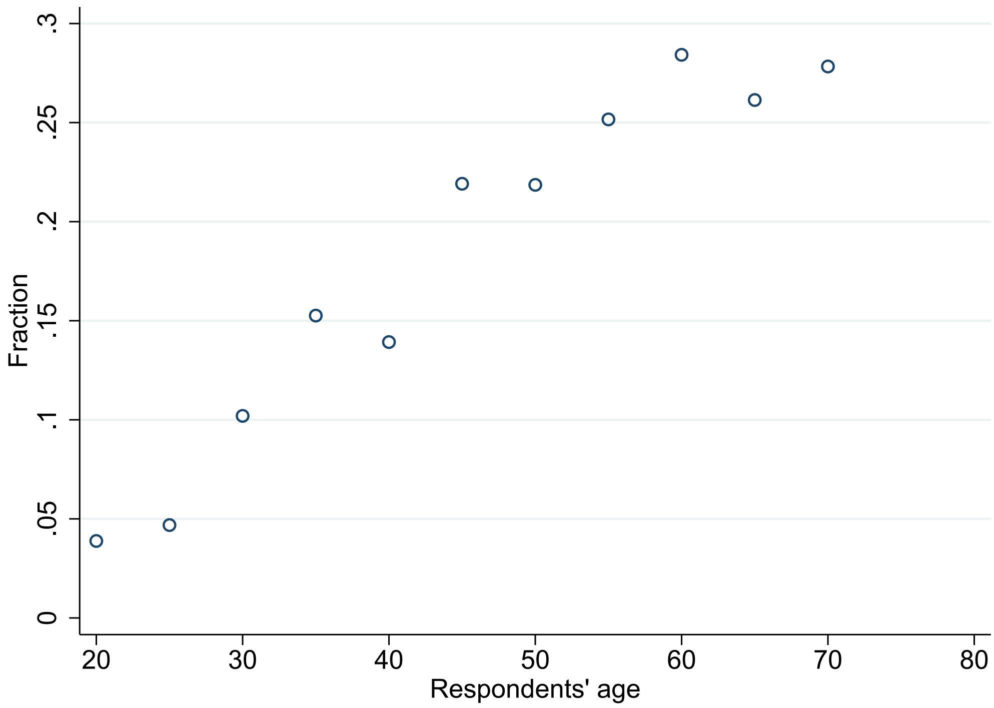
B Supplementary figures

Figure B.1: Fraction of CMOs with background in politics, 1991-2015.



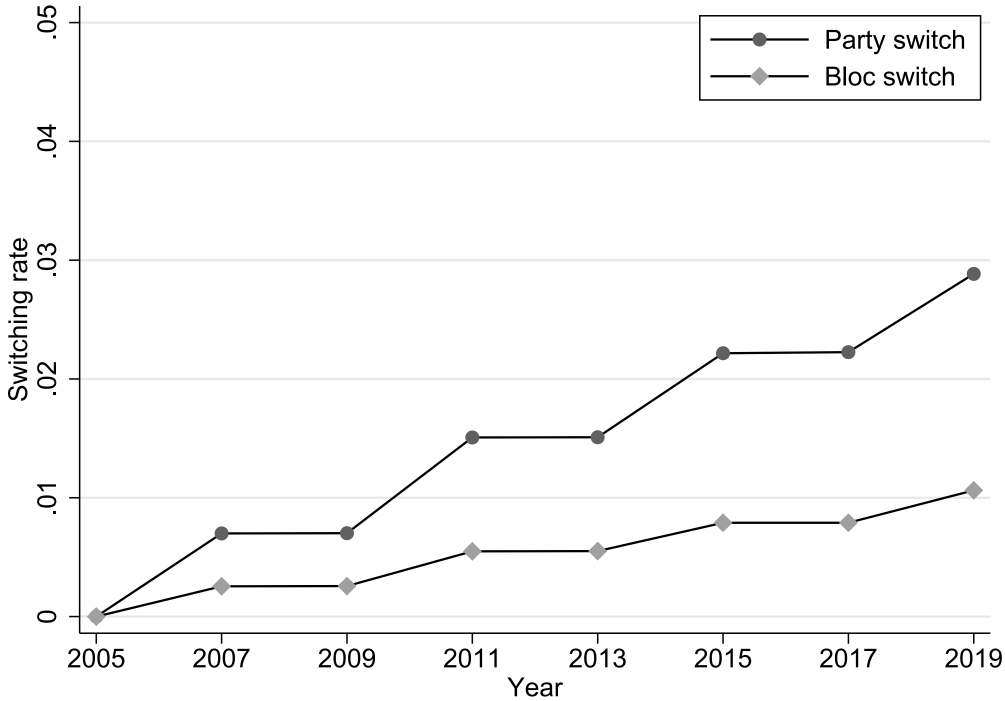
Note: The top panel displays, for each year in the sample, the fraction of CMOs where we can establish the partisan leaning from electoral lists. The bottom panel plots, for each year in the sample, the share of CMOs that have won political office in the election data we have available.

Figure B.2: Fraction of survey respondents previously running for local office by respondents' age



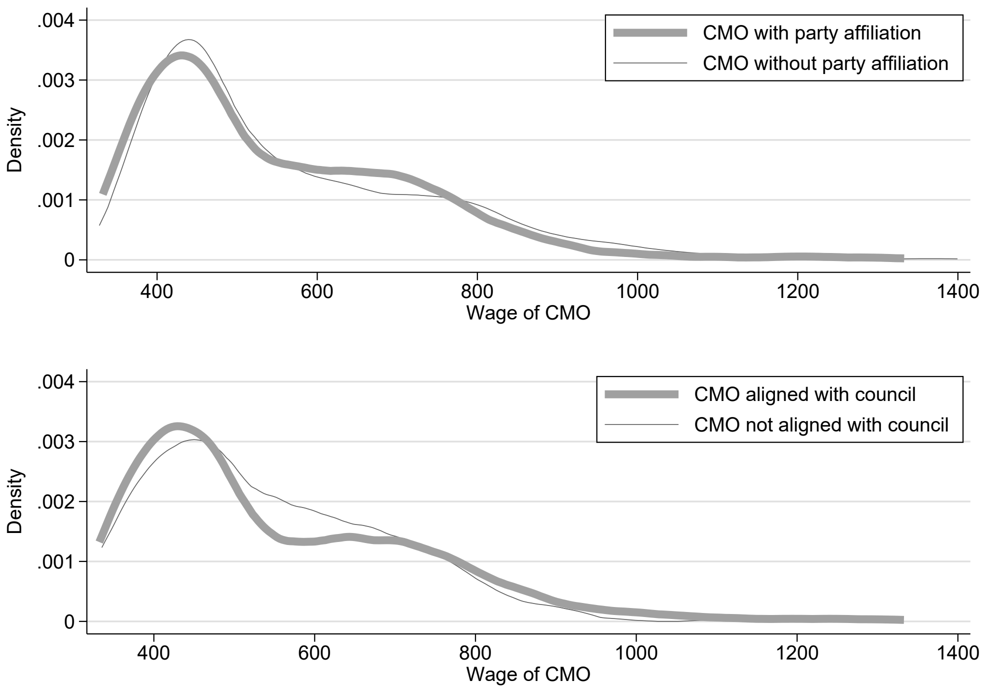
Note: This figure plots the fraction of survey respondents that have previously run for local office against the respondent's age. Data from the 1999-2011 *Local Election Surveys* (N=10,319).

Figure B.3: Future party switching for local candidates running in 2003



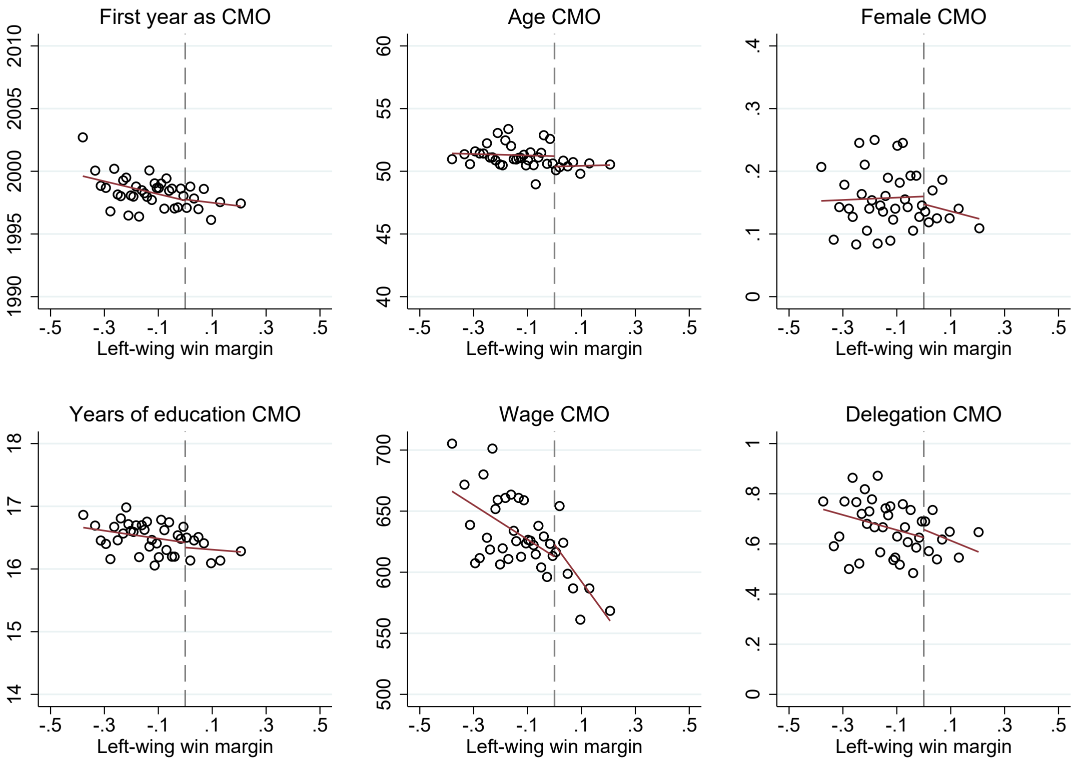
Note: The sample used to construct this figure is the 55,708 candidates running for local office in the 2003 election. We plot the accumulated fraction of candidates switching to another party (bloc) by each following election year. Local elections are held in 2007, 2011, 2015, and 2019. National elections are held in 2005, 2009, 2013, and 2017.

Figure B.4: CMO wage distributions



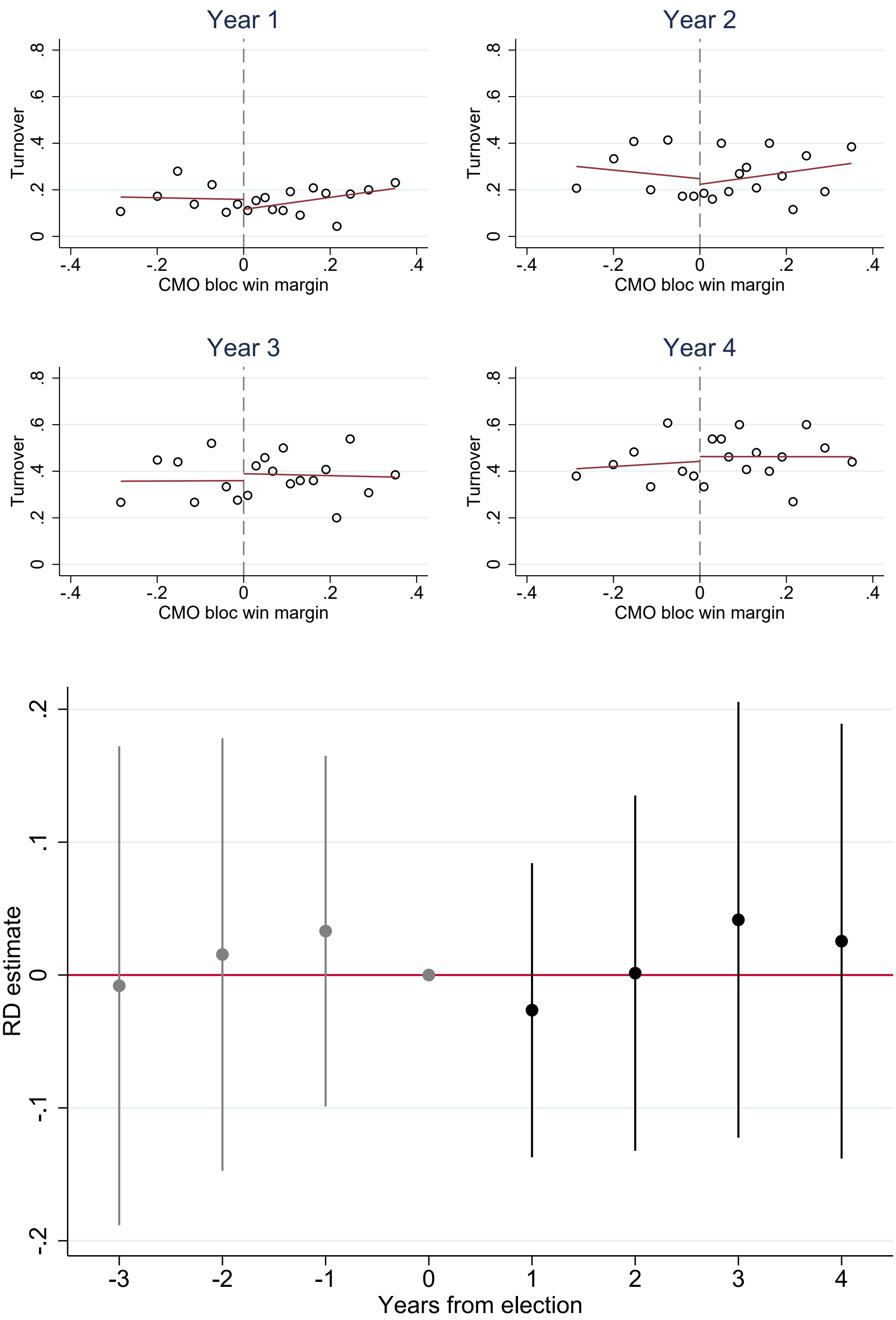
Note: This figure shows four different wage distributions based on real annual gross salary (in 2011 NOK) using an Epanechnikov kernel with optimal bandwidth. The top panel shows kernel density plots for wage levels of CMOs with party affiliation (thick line) and without party affiliation (thin line). The bottom panel shows kernel density plots for wage levels of CMOs aligned with council majority (thick line) and not aligned with council majority (thin line).

Figure B.5: Balance on CMO-level covariates by left-wing win margin



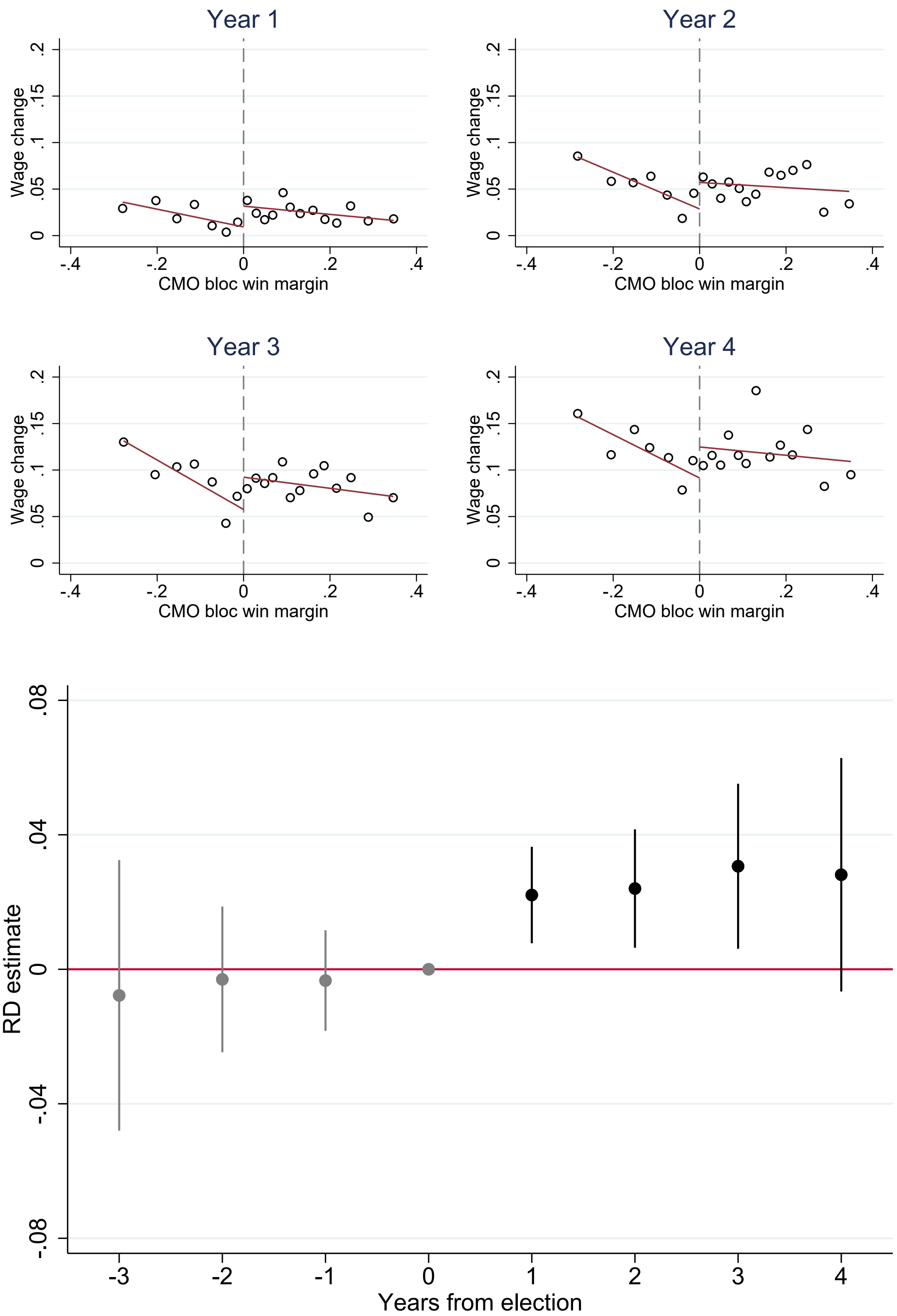
Note: RD plots showing covariate balance for six different CMO-level variables (given in the title of each panel) at the time of election by the left-wing win margin. Separate linear lines are estimated below and above the discontinuity using the underlying data, not the binned scatter points. Each scatter point includes about the same number of observations.

Figure B.6: Council-bureaucrat alignment and bureaucrat turnover



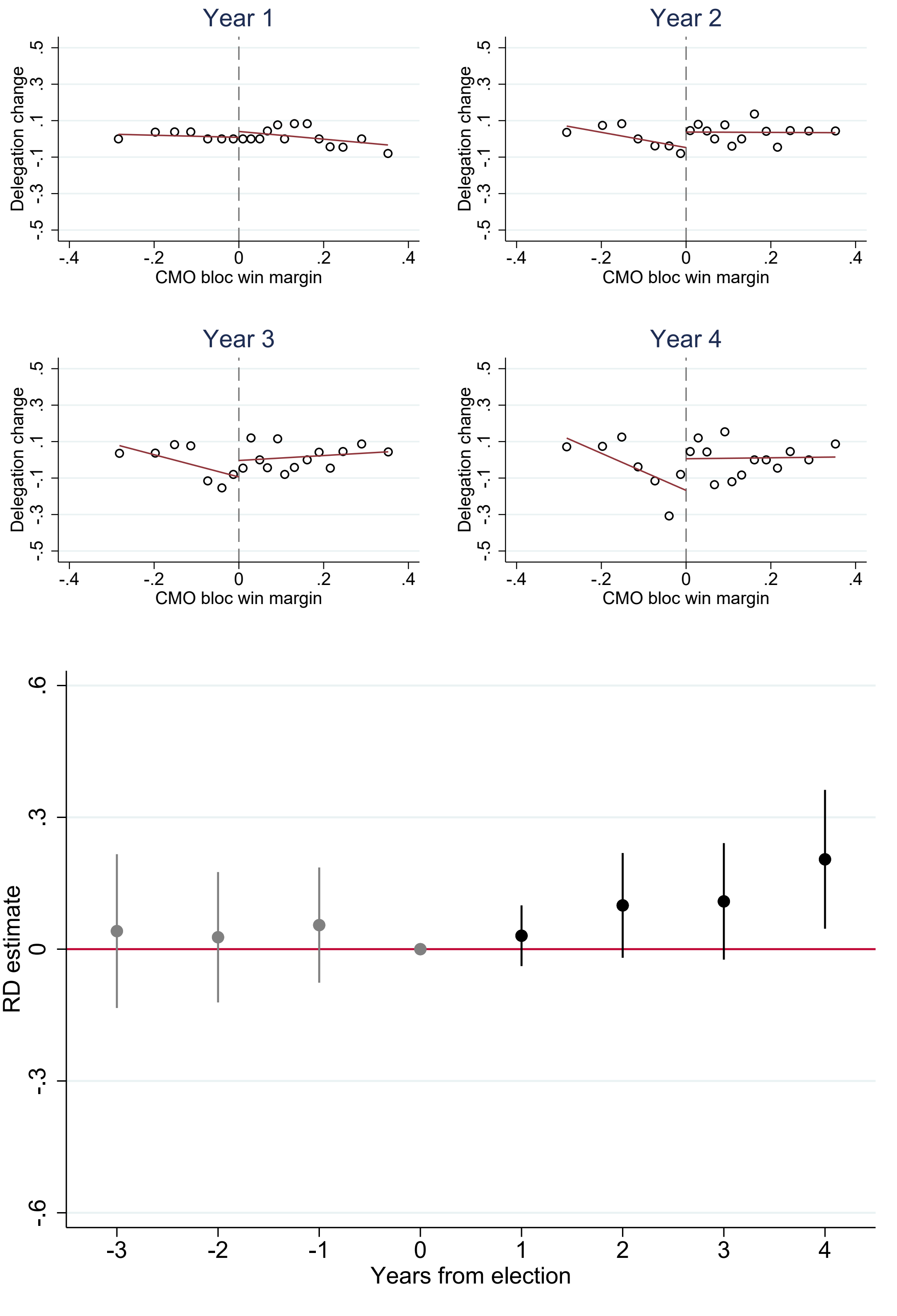
Note: The top panel displays RD plots showing how changes in bureaucratic turnover, from year 0 to year 1, 2, 3, and 4, depends on council-bureaucratic alignment. Separate linear lines are estimated below and above the discontinuity using the underlying data, not the binned scatter points. Each scatter point includes about the same number of observations. The bottom panel shows RD estimates along with 95 % confidence intervals using the full bandwidth and a triangular kernel. Gray bars are based on pre-election years, black bars are based on post-election years.

Figure B.7: Council-bureaucrat alignment and bureaucrat remuneration; Sample limited to municipalities without CMO turnover



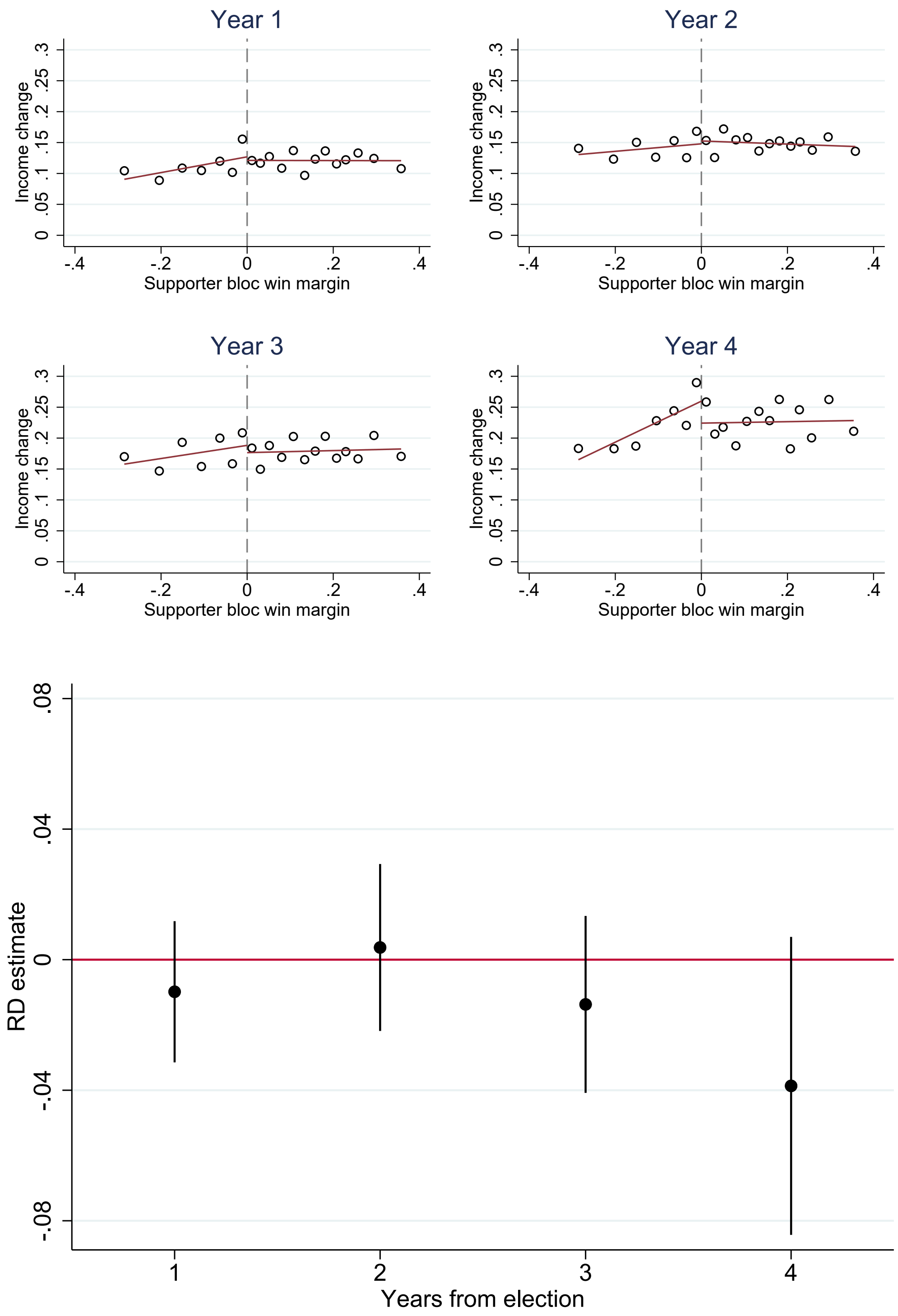
Note: The top panel displays RD plots showing how changes in bureaucratic remuneration, from year 0 to year 1, 2, 3, and 4, depends on council-bureaucratic alignment. Separate linear lines are estimated below and above the discontinuity using the underlying data, not the binned scatter points. Each scatter point includes about the same number of observations. The bottom panel shows RD estimates along with 95 % confidence intervals using the full bandwidth and a triangular kernel. Gray bars are based on pre-election years, black bars are based on post-election years. The sample is limited to municipalities without CMO turnover.

Figure B.8: Council-bureaucrat alignment and delegation (including interpolated delegation data)



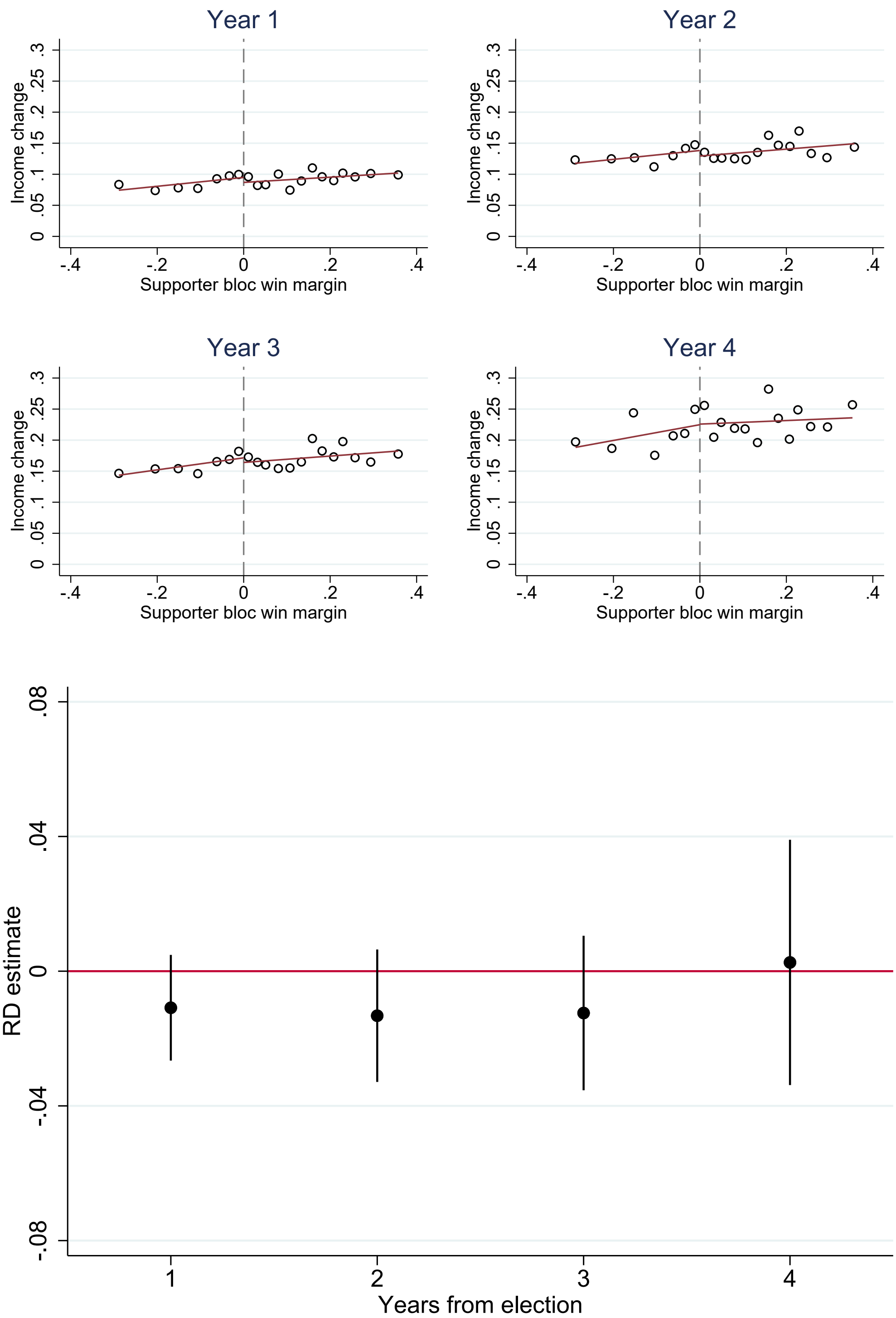
Note: The top panel displays RD plots showing how changes in delegation, from year 0 to year 1, 2, 3, and 4, depends on council-bureaucratic alignment. Separate linear lines are estimated below and above the discontinuity using the underlying data, not the binned scatter points. Each scatter point includes about the same number of observations. The bottom panel shows RD estimates along with 95 % confidence intervals using the full bandwidth and a triangular kernel. Gray bars are based on pre-election years, black bars are based on post-election years. Data on delegation have been interpolated to get a more complete time-series (see Appendix C for details on the interpolation process.

Figure B.9: Council-supporter alignment and supporter income: Elected candidates



Note: The top panel displays RD plots showing how changes in supporter (real) income, from year 0 to year 1, 2, 3, and 4, depends on council-supporter alignment. Income changes are winsorized at the 1% level. Separate linear lines are estimated below and above the discontinuity using the underlying data, not the binned scatter points. Each scatter point includes about the same number of observations. The bottom panel shows RD estimates along with 95 % confidence intervals using the full bandwidth and a triangular kernel. Standard errors are clustered at the municipality-year level.

Figure B.10: Council-supporter alignment and supporter income: Candidates not elected



Note: The top panel displays RD plots showing how changes in supporter (real) income, from year 0 to year 1, 2, 3, and 4, depends on council-supporter alignment. Income changes are winsorized at the 1% level. Separate linear lines are estimated below and above the discontinuity using the underlying data, not the binned scatter points. Each scatter point includes about the same number of observations. The bottom panel shows RD estimates along with 95 % confidence intervals using the full bandwidth and a triangular kernel. Standard errors are clustered at the municipality-year level.

Table B.1: RD analyses of covariate balance on municipality characteristics: Part I

Outcome variable: Population (log)

	(1)	(2)	(3)	(4)	(5)
RD estimate	0.059	-0.212	-0.491	-0.066	-0.580
	(0.228)	(0.293)	(0.339)	(0.251)	(0.335)
Bandwidth	0.500	0.500	0.106	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	207	207	100	57	24
N right of cut-off	341	341	144	68	34

Outcome variable: Children (% of population)

	(1)	(2)	(3)	(4)	(5)
RD estimate	-0.185	-0.351	-0.139	-0.339	-0.267
	(0.292)	(0.381)	(0.469)	(0.319)	(0.439)
Bandwidth	0.500	0.500	0.101	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	207	207	95	57	24
N right of cut-off	341	341	137	68	34

Outcome variable: Young (% of population)

	(1)	(2)	(3)	(4)	(5)
RD estimate	0.190	0.513	0.071	0.281	-0.075
	(0.331)	(0.438)	(0.558)	(0.356)	(0.521)
Bandwidth	0.500	0.500	0.096	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	207	207	91	57	24
N right of cut-off	341	341	128	68	34

Outcome variable: Elderly (% of population)

	(1)	(2)	(3)	(4)	(5)
RD estimate	0.175	0.871	1.128	0.523	1.160
	(0.761)	(0.966)	(1.081)	(0.881)	(0.954)
Bandwidth	0.500	0.500	0.107	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	207	207	100	57	24
N right of cut-off	341	341	145	68	34

Outcome variable: Unemployment (% of lab. force)

	(1)	(2)	(3)	(4)	(5)
RD estimate	-0.013	0.125	0.321	0.029	0.370
	(0.220)	(0.287)	(0.338)	(0.252)	(0.331)
Bandwidth	0.500	0.500	0.102	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	207	207	97	57	24
N right of cut-off	341	341	138	68	34

Note: See Table B.2.

Table B.2: RD analyses of covariate balance on municipality characteristics: Part II

Outcome variable: Women (% of population)

	(1)	(2)	(3)	(4)	(5)
RD estimate	0.059	-0.212	-0.491	-0.066	-0.580
	(0.228)	(0.293)	(0.339)	(0.251)	(0.335)
Bandwidth	0.500	0.500	0.106	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	207	207	100	57	24
N right of cut-off	341	341	144	68	34

Outcome variable: Latitude

	(1)	(2)	(3)	(4)	(5)
RD estimate	-0.266	0.203	0.746	0.236	0.453
	(0.710)	(0.895)	(1.105)	(0.740)	(1.045)
Bandwidth	0.500	0.500	0.104	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	207	207	98	57	24
N right of cut-off	340	340	141	68	34

Outcome variable: Longitude

	(1)	(2)	(3)	(4)	(5)
RD estimate	-0.702	-0.543	0.065	-0.103	-0.435
	(0.872)	(1.151)	(1.429)	(0.892)	(1.402)
Bandwidth	0.500	0.500	0.114	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	207	207	102	57	24
N right of cut-off	340	340	154	68	34

Outcome variable: Year of election

	(1)	(2)	(3)	(4)	(5)
RD estimate	0.588	1.042	1.097	1.376	1.500
	(1.162)	(1.577)	(2.001)	(1.225)	(1.875)
Bandwidth	0.500	0.500	0.103	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	207	207	98	57	24
N right of cut-off	341	341	140	68	34

Note: The reported RD estimates in column (1) correspond to β from Equation (1) using the covariate reported in the panel heading as the outcome variable. In column (2), a second-order polynomial in the forcing variable is included on each side of the discontinuity. In column (3) we use a linear control function and apply the bandwidth suggested by the Calonico et al. (2017) method. In column (4) and (5), we drop the control function and compare differences in means close to the cut-off. Standard errors clustered at the CMO level in parentheses.

Table B.3: RD analyses of covariate balance on CMO characteristics: Part I

Outcome variable: First year as CMO

	(1)	(2)	(3)	(4)	(5)
RD estimate	1.194	2.420	1.803	3.074	1.000
	(1.280)	(1.702)	(2.203)	(1.359)	(1.997)
Bandwidth	0.500	0.500	0.096	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	202	202	89	57	24
N right of cut-off	335	335	126	68	34

Outcome variable: Age CMO (years)

	(1)	(2)	(3)	(4)	(5)
RD estimate	0.855	0.922	0.298	0.579	0.123
	(1.212)	(1.560)	(1.748)	(1.299)	(1.781)
Bandwidth	0.500	0.500	0.118	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	202	202	102	57	24
N right of cut-off	335	335	157	68	34

Outcome variable: Female CMO (dummy)

	(1)	(2)	(3)	(4)	(5)
RD estimate	0.002	-0.002	0.076	0.007	0.076
	(0.059)	(0.072)	(0.081)	(0.065)	(0.070)
Bandwidth	0.500	0.500	0.092	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	202	202	85	57	24
N right of cut-off	335	335	116	68	34

Outcome variable: Education CMO (years)

	(1)	(2)	(3)	(4)	(5)
RD estimate	-0.249	-0.234	-1.408	-0.096	-0.713
	(0.450)	(0.585)	(0.677)	(0.552)	(0.614)
Bandwidth	0.500	0.500	0.070	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	193	193	68	57	24
N right of cut-off	319	319	93	67	34

Note: The reported RD estimates in column (1) correspond to β from Equation (1) using the covariate reported in the panel heading as the outcome variable. In column (2), a second-order polynomial in the forcing variable is included on each side of the discontinuity. In column (3) we use a linear control function and apply the bandwidth suggested by the Calonico et al. (2017) method. In column (4) and (5), we drop the control function and compare differences in means close to the cut-off. Standard errors clustered at the CMO level in parentheses.

Table B.4: RD analyses of covariate balance on CMO characteristics: Part II

Outcome variable: Wage CMO (NOK 1000)

	(1)	(2)	(3)	(4)	(5)
RD estimate	27.117	5.924	-4.374	23.288	0.760
	(29.676)	(40.365)	(52.878)	(31.007)	(52.614)
Bandwidth	0.500	0.500	0.093	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	202	202	85	57	24
N right of cut-off	335	335	121	68	34

Outcome variable: Delegation CMO

	(1)	(2)	(3)	(4)	(5)
RD estimate	0.147	0.127	0.052	0.167	-0.026
	(0.108)	(0.150)	(0.195)	(0.117)	(0.195)
Bandwidth	0.500	0.500	0.102	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	112	112	52	33	13
N right of cut-off	183	183	71	35	12

Note: The reported RD estimates in column (1) correspond to β from Equation (1) using the covariate reported in the panel heading as the outcome variable. In column (2), a second-order polynomial in the forcing variable is included on each side of the discontinuity. In column (3) we use a linear control function and apply the bandwidth suggested by the Calonico et al. (2017) method. In column (4) and (5), we drop the control function and compare differences in means close to the cut-off. Standard errors clustered at the CMO level in parentheses.

Table B.5: RD estimates of council-bureaucrat alignment on bureaucratic turnover

Panel A: One year after the election

	(1)	(2)	(3)	(4)	(5)
RD estimate	-0.026	0.034	0.020	0.012	0.022
	(0.056)	(0.075)	(0.087)	(0.058)	(0.092)
Bandwidth	0.500	0.500	0.097	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	196	196	90	56	24
N right of cut-off	325	325	127	67	34

Panel B: Two years after the election

	(1)	(2)	(3)	(4)	(5)
RD estimate	0.001	0.071	0.045	0.069	0.045
	(0.068)	(0.089)	(0.094)	(0.071)	(0.106)
Bandwidth	0.500	0.500	0.113	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	203	203	100	56	24
N right of cut-off	336	336	149	66	33

Panel C: Three years after the election

	(1)	(2)	(3)	(4)	(5)
RD estimate	0.042	0.074	0.033	0.102	0.061
	(0.084)	(0.111)	(0.126)	(0.089)	(0.127)
Bandwidth	0.500	0.500	0.122	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	198	198	103	57	24
N right of cut-off	334	334	157	68	34

Panel D: Four years after the election

	(1)	(2)	(3)	(4)	(5)
RD estimate	0.025	0.039	-0.002	0.087	-0.034
	(0.084)	(0.114)	(0.136)	(0.094)	(0.135)
Bandwidth	0.500	0.500	0.112	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	203	203	100	57	24
N right of cut-off	335	335	149	68	34

Note: The reported RD estimates in column (1) correspond to β from Equation (1), which are shown in the bottom panel of Figure B.6. In column (2), a second-order polynomial in the forcing variable is included on each side of the discontinuity. In column (3) we use a linear control function and apply the bandwidth suggested by the Calonico et al. (2017) method. In column (4) and (5), we drop the control function and compare differences in means close to the cut-off. Standard errors clustered at the CMO level in parentheses.

C Data sources and measurement

C.1 CMO compensation and turnover

The Norwegian Association of Local and Regional Authorities (KS) is the employers' organization of local government authorities and operates a register of all these authorities' employees (the PAI-register). We use data from this register covering the period 1991-2015. The register has information on the name, birthdate and wages of the CMO employed in each municipality on December 1st of every year. The register provides information about both gross regular monthly salary as well as various supplementary compensations. The latter derive from, for instance, allowances for evening and night shifts or work on Saturdays and Sundays (accounting for approximately 1% of total wage level).

Access to the PAI-register allows us to characterize the complete set and length of employment spells (measured in years) for all CMOs. We have performed extensive quality checking on the data on CMO turnover, and excluded observations where substitute CMOs held temporary positions.

C.2 CMOs' party affiliation and political alignment

Data on the political composition of the municipal council as well as the party of the mayor are obtained from the *Norwegian Centre for Research Data* (NSD), as organized by Fiva et al. (2017). We establish CMOs' party affiliation by searching for matches in data sets covering candidates running for local, regional and national office in Norway. For this purpose we rely on candidate names, birth years, and municipalities of residence. For candidate names we use a fuzzy-matching method to account for occasional spelling errors, typos or differences in the treatment of middle names. We subsequently do extensive quality checks of our resulting matches. This allows us to reduce any danger of misclassification. As we have access to administrative data on all CMOs, we can also account for (female) CMOs changing their name after marriage.

For the local and regional level of government, we rely on candidate data as organized by Fiva et al. (2020). At the local government level, we have data on all candidates running in the last five municipal elections in, respectively, 419, 354, 228, 428 and 356 municipalities (in total 299,926 candidate-year observations). We supplement these data with additional information on mayors (3,600 mayor-year observations) from 1971-2019, obtained from the Norwegian Center for Research Data (NSD). At the regional government level, we collected data on all elected candidates in the 1975-2019 elections, all non-elected candidates in the 2003-2019 period, and about half of non-elected candidates in the 1975-1999 period (in total 75,756 candidate-year observations). For the national level, we rely on the Fiva and Smith (2017) data set which covers the universe of candidates running in the 1906-2017 period. In our search for the party affiliation of CMOs active in the 1991-2015 period, we rely on candidates running for national office in the 1961-2017 period (47,559 candidate-year observations).

These data allow us to establish CMOs' party affiliation based on a comprehensive search at all levels of government in Norway. We find 109 matches for national elections, 182 matches for regional elections, and 340 matches for local elections. This includes overlapping matches whereby the same individual stands for election at different levels of government. In total, we identify the partisan leaning of 446 unique CMOs (though we drop two of these from our final sample due to their unstable party bloc affiliation over time). As documented in the main text, the identified partisan leaning predates the CMO's spell in office in the majority of cases (254 out of 446). For the remaining CMOs, we are forced to rely on party affiliations observed after their CMO spell. Using this information to approximate the partisan affiliation for these CMOs during their spell in office is feasible due to the extreme stability of political attachments over time in Norway (O'Brien and Shomer 2013; Cirone, Fiva and Cox 2020). Indeed, we show that party bloc switching occurs for only two of the 56 CMOs in our sample for whom we observe partisan affiliations at multiple points in time. More generally, we can document that only 1% of all 56,000 candidates running in the 2003 local elections have switched party bloc even after 16 years (for details, see Section 3.3 in the main text). Such temporal stability of political attachments allows us to include in our sample all CMOs with an identified political leaning.

The CMO is classified as aligned if the CMO's party affiliation matches that of the council majority. This means that the CMO is defined as aligned if (s)he is affiliated with the majority party bloc. Party blocs are defined as follows: Right-wing bloc: Progress Party, Conservative Party, Liberal Party, Christian Democratic Party, Center Party and other right-wing lists. Left-wing bloc: Red Party (before 2007, Red Electoral Alliance), Socialist Left Party, Labor Party and other left-wing lists. Following Fiva et al. (2017), we classify independent local lists (23 matches) as belonging to the right-wing bloc.

C.3 Delegation of budgetary powers

The Local Government Organizational Database provides extensive information about the internal organization of Norwegian local authorities, including the extent of budgetary delegation. The data has originally been collected by means of repeated survey questionnaires to local authorities. The complete database is through the Norwegian Center for Research Data (NSD).[5]

The key variable of interest for our analysis relates to the delegation of budgetary powers to the CMO. In Norway, local governments have various ways to organize the preparatory stages of the budget process before the local executive board submits the final proposal to the local council for formal approval. Crucially, the budgetary preparations can thereby involve the CMO to different degrees. In effect, three main approahces are available:

- A: The bottom up process: The administrative agencies and standing political committees draft budget proposals, which are subsequently processed by the CMO. The CMO submits a revised budget proposal for the executive board.

[5] The database is "Kommunal- og moderniseringsdepartementets Organisasjonsdatabase for kommuner og fylkeskommuner". For detailed documentation, see https://nsd.no/nsddata/serier/kommunalorganisering.html.

- B: The centralized administrative process: The CMO presents a coherent budget proposal for treatment in the standing committees. The executive board prepares its proposal on basis of CMO and committee proposals.
- C: The centralized political process: The executive board initiates and controls the budgetary process, collecting information from the standing committees and the CMO. Using this information, the executive board submits its proposal to the local council.

Delegation to the CMO in our analysis is set to 1 when the municipality employs either the bottom up process (A) or the centralized administrative process (B), and zero otherwise.

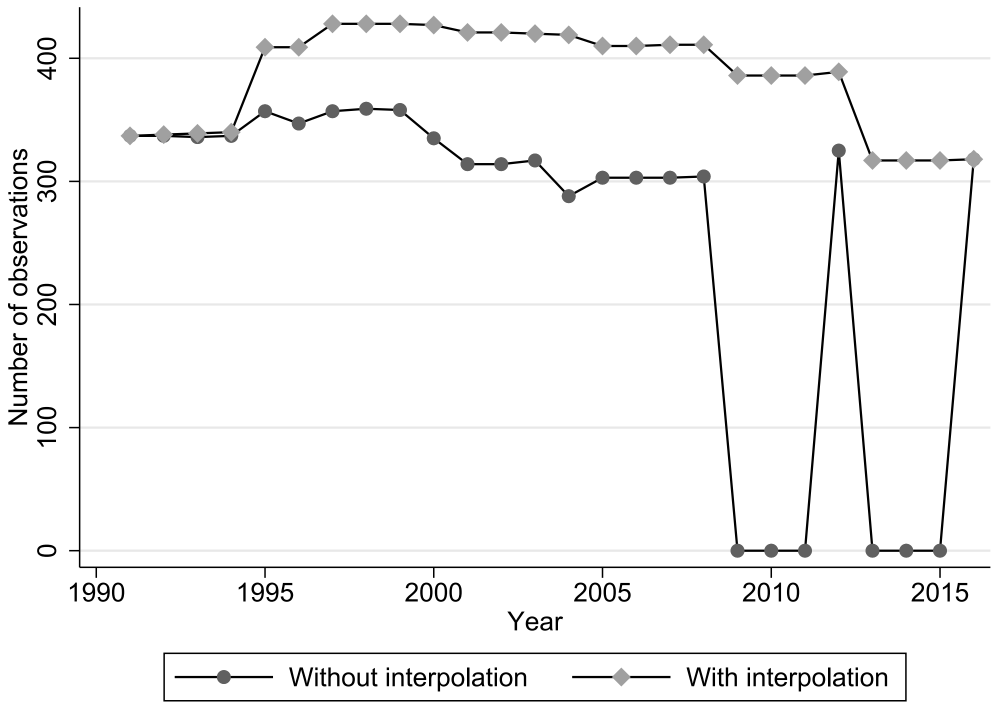
Annual data on the budgetary process is available for the period 1991-2008, and subsequently also for the years 2012 and 2016. The aggregate statistics of our dataset correspond exactly to those in the documentation reports. This typology has been used in previous research, notably in Hagen and Vabo (2005).

Interpolation

The dataset on budgetary delegation has missing observations deriving both from years when no surveys were fielded to collect the data, and from some municipalities not answering the survey in certain years. While our main analysis relies only on the available data, we also engaged in robustness checks where we interpolated the data to get more complete time-series. In cases where we miss an observation for a particular local authority in a particular year, we interpolate by inserting the subsequent observation. For example, if data on delegation is missing for 1997, but not for 1996 or 1998, we replace the missing observation with the one from 1998.

Data collection did not include two three-year periods, i.e. 2009-2011 and 2013-2015. We use the Stata module nnipolate, and apply nearest neighbour interpolation for the delegation indicator. When we have municipality-level delegation data for the start- and end-points, we apply the procedure to fill in missing values using the previous or next known value of delegation, depending on which is closer in time. When the previous and next values are equally distant (i.e. in the years 2010 and 2014), we use the next observation (i.e. delegation observations for 2012 and 2016). Figure C.1 presents the original and interpolated delegation indicators for the 1991-2016 period.

Figure C.1: Delegation data with and without interpolation



Note: The figure presents data on the share of Norwegian municipalities for which we have information about the level of delegation in the budget process. The dark grey dots cover only the raw data, while the light grey squares include interpolated data to correct for missing years.

C.4 Income data for local election candidates

From the administrative registers of Statistics Norway, we have access to income data for the 2007-2014 period. For candidates participating in the 2007 local elections (N=62,755), we can measure their change in income relative to the election year (2007) throughout the subsequent four-year election period (2008-2011). We proceed similarly for candidates participating in the 2011 local elections (N=59,486), but for this sample lack data for the fourth year of the election period (i.e. 2015). All income measures are expressed in 2015 Norwegian kroner.

As in the main analysis, we classify candidates as belonging to either the left-wing or right-wing bloc (see section C.2) based on their party affiliation. We focus on candidates aged 25–54 years in the election year, who for the most part have finished their formal schooling and are not on the verge of retirement during the election period (N=70,138). After excluding candidates we are unable to merge with administrative data and candidates with missing income data in any year, we are left with 62,795 candidate-election observations. 22% of these candidates were elected to the local council (N=13,953). We winsorize the income change data at the 1% level. Figure C.2 display the frequency of observations by income levels (top panel) and income changes (bottom panel).

References

Cirone, Alexandra, Jon H. Fiva and Gary W. Cox. 2020. "Seniority-based Nominations and Political Careers." *American Political Science Review* forthcoming.

Fiva, Jon H., Askill H. Halse and Gisle J. Natvik. 2017. "Local Government Dataset." Available at `www.jon.fiva.no/data.htm`.

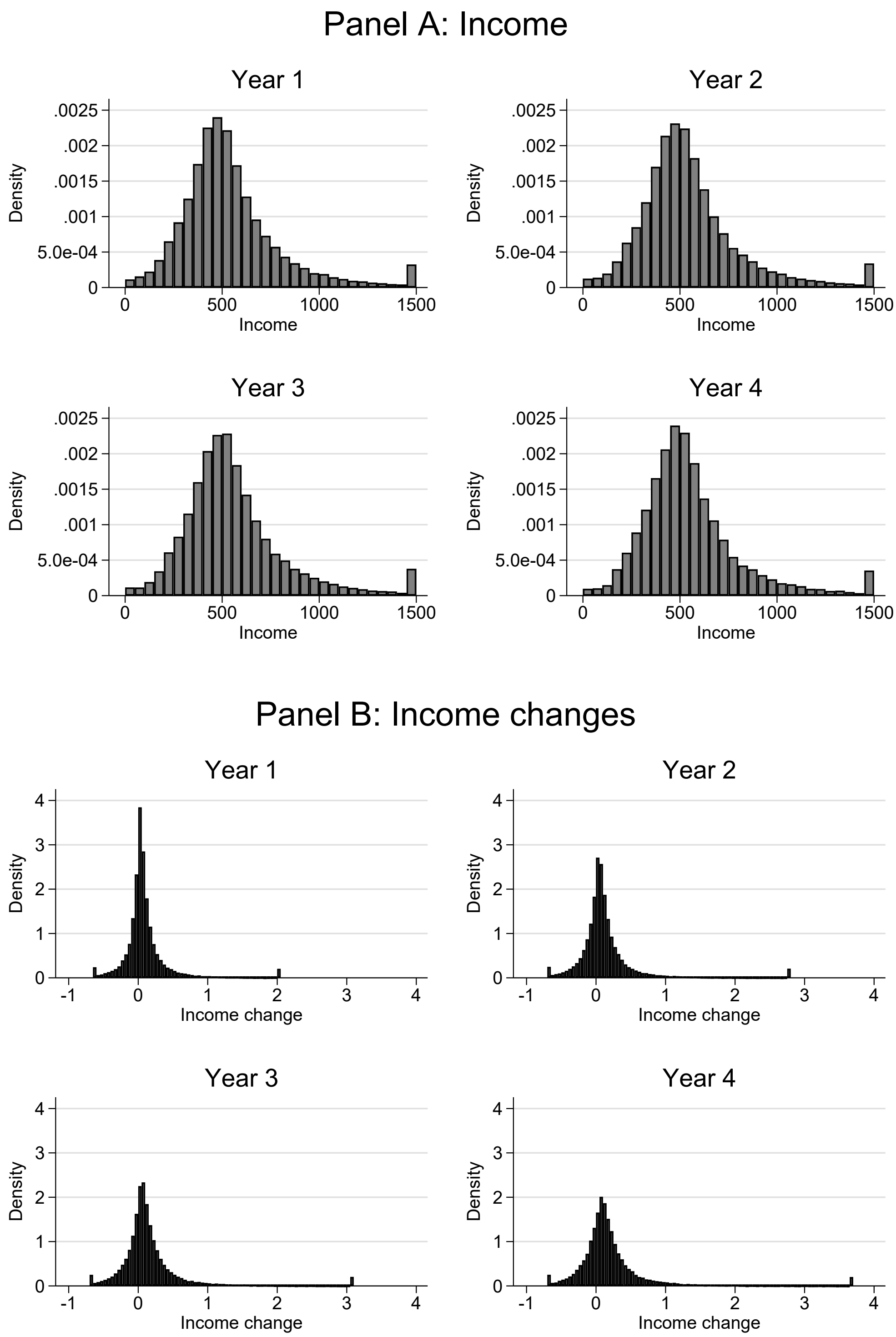
Fiva, Jon H. and Daniel M. Smith. 2017. "Norwegian Parliamentary Elections, 1906-2013: Representation and Turnout Across Four Electoral Systems." *West European Politics* 40(6):1373–1391.

Fiva, Jon H., Rune J. Sørensen and Reidar Vøllo. 2020. “Local Candidate Dataset.” Available at www.jon.fiva.no/data.htm.

Hagen, Terje P. and Signy I. Vabo. 2005. “Political Characteristics, Institutional Procedures and Fiscal Performance: Panel Data Analyses of Norwegian Local Governments, 1991–1998.” *European Journal of Political Research* 44:43–64.

O'Brien, Diana Z. and Yael Shomer. 2013. “A Cross–National Analysis of Party Switching.” *Legislative Studies Quarterly* 38(1):111–141.

Figure C.2: Income data for local election candidates



Note: The top panel of the figure presents year-by-year income data (censored at NOK 1 500 000). The bottom panel shows the winsorized income change data used for the analysis of Section 6.2 in the main text.

D RD analysis of changes in council majority

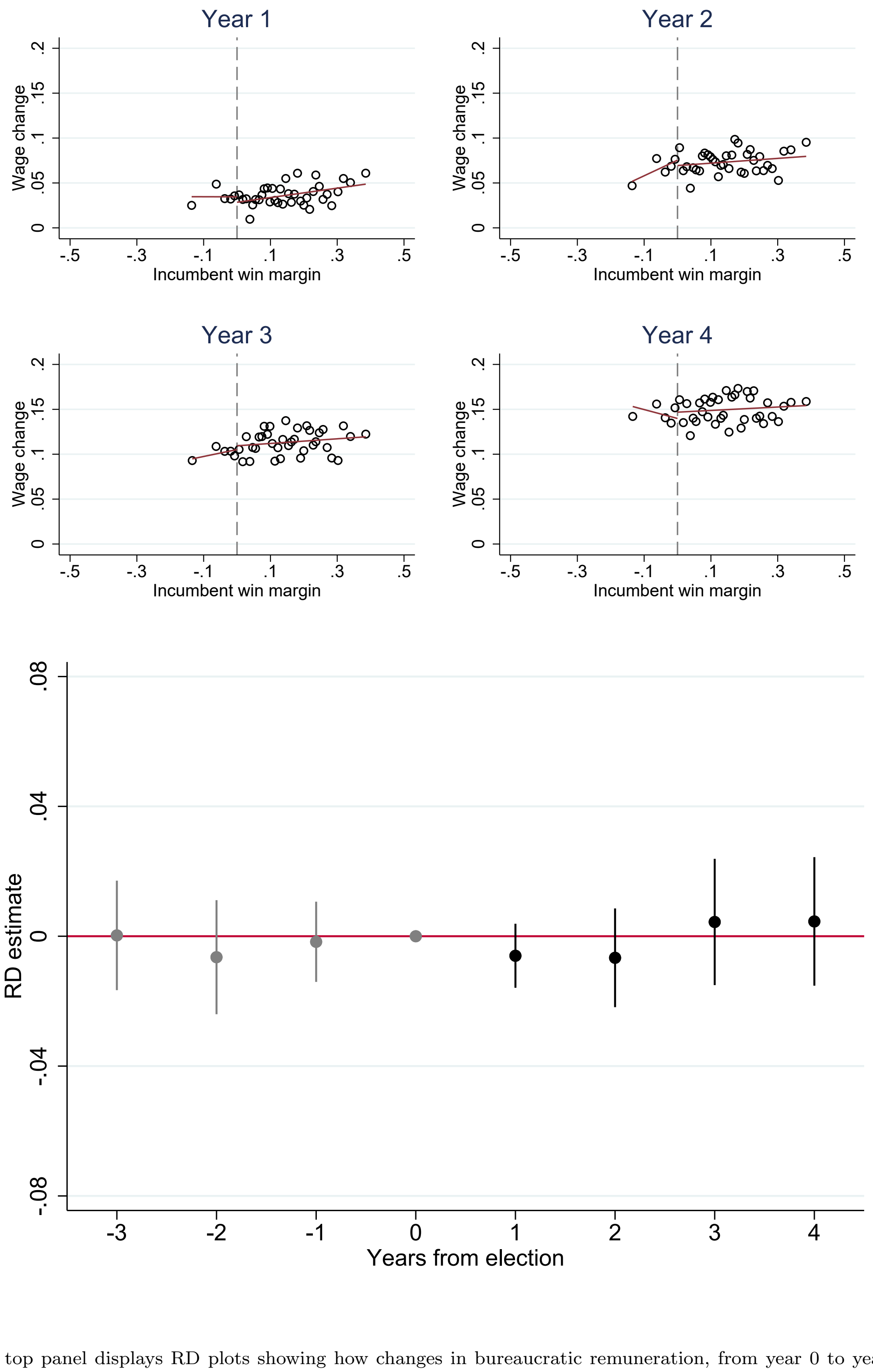
In this section, we look at the impact on bureaucratic pay and turnover from changing the council majority. This has been the common empirical approach in the literature assessing how partisan (mis)alignment affects bureaucratic turnover, even though it rests on the dubious assumption that politician-bureaucrat preference alignment falls with a shift in government. The forcing variables in this RD analysis is the win margin of the incumbent bloc (defined as the political bloc with a seat majority *before* the relevant election). The incumbent bloc win margin ($\widetilde{Margin}$) equals the *left-wing win margin* if the left-wing bloc holds a seat-majority prior to the election. If the right-wing bloc holds a seat majority, $\widetilde{Margin} = (-1)\cdot$ *left-wing win margin*. Naturally, this variable has most of its density to the right of zero, as displayed in the bottom-left panel of Figure 1 in the main text. The regression model takes the following form:

$$Y_i^t = \widetilde{\alpha} + \widetilde{\beta} IncumbentWin_i + \widetilde{\gamma}_1 \widetilde{Margin}_i + \widetilde{\gamma}_2 \widetilde{Margin}_i \cdot IncumbentWin_i + \widetilde{\epsilon}_i \qquad (1)$$

where $IncumbentWin$ is an indicator variable equal to 1 when the incumbent political bloc retains a council seat majority, and 0 otherwise. The key dependent variables (Y_i^t) are changes in bureaucratic pay (i.e. CMO gross salary of municipality i from the last year before the election to year t) and bureaucratic turnover (1 if the CMO of municipality i in place before the election is replaced by year t, 0 otherwise). The coefficient of interest is $\widetilde{\beta}$, which reflects the causal effect of having the same bloc majority both before and after the election.

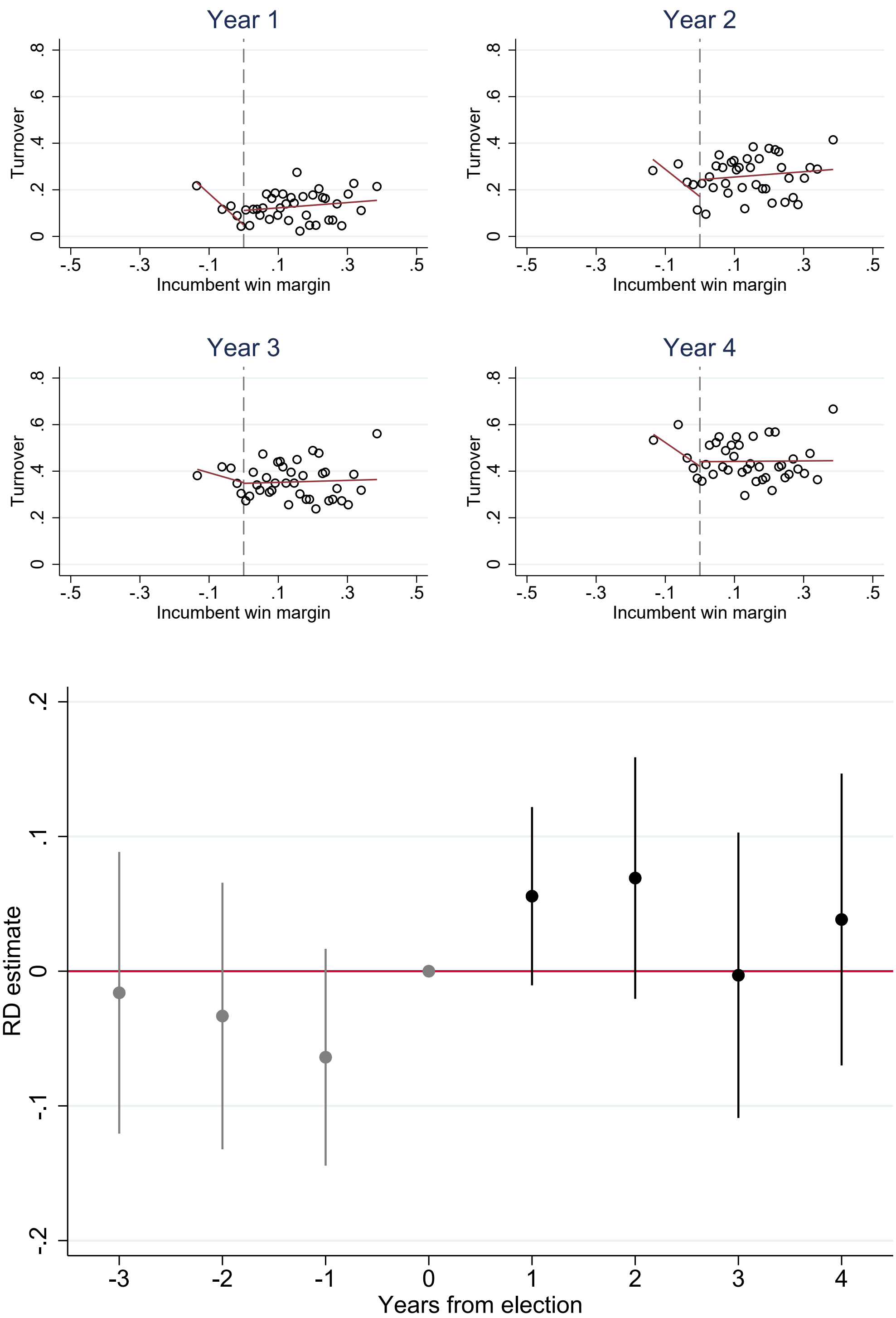
The results are graphically presented in Figure D.1. The top panel shows four RD plots relating contemporaneous shifts in council majorities (year $t = 0$) to changes in bureaucratic pay over the election period (year $t = 1$, 2, 3 or 4). The bottom panel of each figure plots the RD estimates with corresponding 95% confidence intervals. The central observation in Figure D.1 is that a change in the council majority in itself has no clear effect on bureaucrats' wage in subsequent years. This null finding is independent of the number of years we allow to elapse in the election period. Furthermore, as reported in Table D.1, it is equally persistent when changing the bandwidth and polynomial used for implementing the RD. The point estimates never approach statistical significance at conventional levels, and in effect are equally likely to be positive or negative. We also investigate whether a change in the council majority affects the probability that the CMO leaves her position in the years following an election. We find no evidence that this is the case, see Figure D.2. If anything, the results for years 1 and 2 suggest that a shift in council majority initially works to weakly *reduce* bureaucratic turnover. However, these effects are not statistically significant at conventional levels, and are quite imprecisely estimated.

Figure D.1: Incumbent re-election and bureaucrat remuneration



Note: The top panel displays RD plots showing how changes in bureaucratic remuneration, from year 0 to year 1, 2, 3, and 4, depends on incumbent re-election. Separate linear lines are estimated below and above the discontinuity using the underlying data, not the binned scatter points. Each scatter point includes about the same number of observations. The bottom panel shows RD estimates along with 95 % confidence intervals using the full bandwidth and a triangular kernel. Gray bars are based on pre-election years, black bars are based on post-election years.

Figure D.2: Incumbent re-election and bureaucratic turnover



Note: The top panel displays RD plots showing how bureaucratic turnover, from year 0 to year 1, 2, 3, and 4, depends on incumbent re-election. Separate linear lines are estimated below and above the discontinuity using the underlying data, not the binned scatter points. Each scatter point includes about the same number of observations). The bottom panel shows RD estimates along with 95 % confidence intervals using the full bandwidth and a triangular kernel. Gray bars are based on pre-election years, black bars are based on post-election years.

Table D.1: RD estimates of incumbent re-election on bureaucratic remuneration

Panel A: One year after the election

	(1)	(2)	(3)	(4)	(5)
RD estimate	-0.006	-0.006	0.006	-0.007	-0.002
	(0.005)	(0.007)	(0.010)	(0.005)	(0.006)
Bandwidth	0.500	0.500	0.060	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	226	226	157	140	90
N right of cut-off	1504	1504	257	213	94

Panel B: Two years after the election

	(1)	(2)	(3)	(4)	(5)
RD estimate	-0.007	-0.008	0.007	-0.006	0.001
	(0.008)	(0.010)	(0.016)	(0.008)	(0.011)
Bandwidth	0.500	0.500	0.056	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	223	223	146	135	88
N right of cut-off	1507	1507	235	211	93

Panel C: Three years after the election

	(1)	(2)	(3)	(4)	(5)
RD estimate	0.004	0.005	0.003	0.002	-0.004
	(0.010)	(0.013)	(0.019)	(0.009)	(0.012)
Bandwidth	0.500	0.500	0.058	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	223	223	157	141	91
N right of cut-off	1494	1494	245	212	92

Panel D: Four years after the election

	(1)	(2)	(3)	(4)	(5)
RD estimate	0.005	-0.004	0.014	-0.001	0.002
	(0.010)	(0.013)	(0.020)	(0.009)	(0.013)
Bandwidth	0.500	0.500	0.061	0.050	0.025
Order of polynomial	1	2	1	0	0
N left of cut-off	228	228	159	141	91
N right of cut-off	1498	1498	257	211	91

Note: The reported RD estimates in column (1) correspond to $\widetilde{\beta}$ from Equation (2), which are shown in the bottom panel of Figure D.1. In column (2), a second-order polynomial in the forcing variable is included on each side of the discontinuity. In column (3) we use a linear control function and apply the bandwidth suggested by the Calonico et al. (2017) method. In column (4) and (5), we drop the control function and compare differences in means close to the cut-off. Standard errors clustered at the CMO level in parentheses.

References

Bendor, Jonathan, Amihai Glazer and Thomas Hammond. 2001. "Theories of Delegation." *Annual Review of Political Science* 4(1):235–269.

Besley, Timothy and Maitreesh Ghatak. 2005. "Competition and Incentives with Motivated Agents." *American Economic Review* 95(3):616–636.

Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell and Rocıo Titiunik. 2017. "rdrobust: Software for Regression Discontinuity Designs." *Stata Journal* 17(2):372–404.

Cirone, Alexandra, Jon H. Fiva and Gary W. Cox. 2020. "Seniority-based Nominations and Political Careers." *American Political Science Review* forthcoming.

Dahlström, Carl and Mikael Holmgren. 2019. "The Political Dynamics of Bureaucratic Turnover." *British Journal of Political Science* 49(3):823–836.

Huber, John D. and Charles R. Shipan. 2008. Politics, Delegation, and Bureaucracy. In *The Oxford Handbook of Political Economy*, ed. Donald A. Wittmann and Barry R. Weingast. Oxford University Press pp. 256–272.

Kopecky, Peter, Jan-Hinrik Meyer Sahling, Francisco Panizza, Gerardo Scherlis, Christian Schuster and Maria Spirova. 2016. "Party patronage in contemporary democracies: Results from an expert survey in 22 countries from five regions." *European Journal of Political Research* 55(2):416–431.

O'Brien, Diana Z. and Yael Shomer. 2013. "A Cross–National Analysis of Party Switching." *Legislative Studies Quarterly* 38(1):111–141.

Peters, B. Guy and Jon Pierre. 2004. *The Politicization of the Civil Service in Comparative Perspective: A Quest for Control.* Abingdon: Routledge.

Toral, Guillermo. 2019. "The benefits of patronage: How the political appointment of bureaucrats can enhance their accountability and effectiveness?" *MIT, mimeo* .