Political Behavior (2019) 41:1071–1098
https://doi.org/10.1007/s11109-018-9483-3

ORIGINAL PAPER

Is the Relationship Between Political Responsibility and Electoral Accountability Causal, Adaptive and Policy-Specific?

Martin Vinæs Larsen[1]

Published online: 7 July 2018

Abstract

Will voters hold an incumbent more electorally accountable for the quality of a policy outcome if the incumbent's political responsibility for the underlying policy increases? To answer this question, this study exploits a reform of labor market regulation in Denmark that exogenously assigned more political responsibility for unemployment services to some municipal mayors. The study finds that in subsequent elections these mayors were held more electorally accountable for unemployment services, but not more accountable for other policy outcomes. This suggests that the relationship between political responsibility and electoral accountability is causal, adaptive and tied to specific policies. On balance, the electorate thus seems to be quite judicious when assigning electoral credit or blame, moderating the extent to which incumbents are held accountable for specific outcomes based on the extent to which these incumbents crafted and implemented the policies that shaped these outcomes.

Keywords Electoral accountability · Retrospective voting · Clarity of responsibility · Quasiexperiments

Voters tend to hold politicians accountable for how they perform in office. They do this by withholding electoral support if policy outcomes deteriorate (Healy and Malhotra 2013), by supporting governments that enact the types of policies they prefer (Erikson et al. 2002), and by calibrating their support for these policies based on the decisions governments make while in office (Soroka and Wlezien 2005). These

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11109-018-9483-3) contains supplementary material, which is available to authorized users.

✉ Martin Vinæs Larsen
mvl@ps.au.dk

[1] Department of Political Science, Aarhus University, Bartholins Allé 7, 8000 Aarhus C, Denmark

types of checks on elected officials are widespread, but seem to be more prevalent in some elections than they are in others.

One thing that seems to consistently predict the degree to which politicians are held electorally accountable is the extent to which a political system concentrates or disperses political responsibility; what is conventionally called "clarity of responsibility" (Powell and Whitten 1993; Whitten and Palmer 1999; Duch and Stevenson 2008; De Vries et al. 2011). For instance, studies have shown that voters are less likely to hold incumbents accountable for the quality of the economic situation if the incumbent operates in an open economy (Hellwig 2001; Fernández-Albertos 2006; Hellwig and Samuels 2007), if the incumbent is part of a coalition government (Nadeau et al. 2002; Hobolt et al. 2013), or if the incumbent operates in a federal system (León 2011; Cutler 2008; Anderson 2006). Electoral accountability can therefore be said to "strongly reflect the nature of policymaking in the society and the coherence and control the government can exert over that policy" (Powell and Whitten 1993, p. 398). However, it is unclear exactly what is implied in the idea that accountability "strongly reflects" the centralization of political responsibility in the hands of the government.

For one, it is not clear whether the correlation identified in the previous literature reflects a *causal* relationship between institutions that centralize responsibility and electoral accountability. While previous work has shown that the political performance of incumbents has a causal effect on the support for these incumbents (Stimson 2015; Healy et al. 2010), we know from the political economy literature that institutions, such as those dispersing or concentrating political responsibility, are fundamentally endogenous (Acemoglu 2005; Besley and Case 2000). Accordingly, it is possible that the observational co-occurrence between institutions that centralize political responsibility and electoral accountability is partly or completely driven by some extraneous, underlying factor.

Further, we do not know exactly how *adaptive* the relationship between electoral accountability and political responsibility is. That is, we do not know whether voters are quick to adjust to changes in political responsibility or whether the correspondence found in previous research is the result of a long-term process that slowly adjusts levels of accountability to levels of responsibility. Previous work has found that voters react fairly quickly to changes in policy outcomes (Wlezien 1995; Soroka and Wlezien 2005), which might suggest that voters will also be adaptive to changes in political institutions. However, since the previous literature on clarity of responsibility has primarily focused on institutional differences that were established a long time ago, we cannot know for sure.

Finally, we know little about the extent to which the relationship is *policy-specific*; that is, whether voters recognize and act on differences in the extent to which the same incumbent is responsible for individual policy outcomes (e.g., unemployment ctr. inflation). On the one hand, research on voters' thermostatic response to specific policy changes (Jennings 2009) and research on policy agendas (Jones and Baumgartner 2005) suggest that voters can discriminate between the outcomes of different policies. On the other hand, we cannot know whether voters act in a similar way when it comes to institutions that disperse political responsibility for specific policy outcomes. As such, previous research has primarily examined institutional

differences that implicate changes in incumbents' responsibility for a wide range of outcomes (e.g., federal contra unitary constitutions), making it impossible to discern whether voters are sensitive to policy-specific differences in political responsibility.

Causality, adaptiveness and policy-specificity are important features of the relationship between clarity of responsibility and electoral accountability because they all tell us something about the extent to which voters only hold incumbent politicians electorally accountable for the policy outcomes these politicians had a hand in shaping, and, thus, whether voters are able to use elections to select politicians who can competently manage the policies they are responsible for.

This article re-examines the relationship between political responsibility and electoral accountability in a context where it possible to cast some light on these different important features of the relationship between electoral accountability and the institutions that centralize responsibility for policy outcomes: a reform of labor market regulation in Denmark. The reform increased municipal mayors' responsibility for unemployment services, and *only* for unemployment services, by making the municipalities responsible for the administration of active labor market policies. In 14 municipalities, the reform was implemented 3 years before the 2009 municipal elections, and in the remaining 84 municipalities, it was implemented after these elections. The decision about which municipalities had to implement the reform before the 2009 election was taken primarily by the central government, and a closer examination of the selection process reveals that the central government plausibly chose municipalities independently of pre-reform levels of electoral accountability. In sum, this reform presents a rare instance in which assignment of political responsibility for a specific policy outcome changed abruptly and exogenously, making it possible to examine whether the institutions that centralize political responsibility affect electoral accountability in a way that is policy-specific, adaptive and causal.

Using the Danish Municipal Election Survey (Elklit and Kjær 2013), I show that voters in the municipalities where the labor market reform was first implemented, the treatment municipalities, held the mayor more electorally accountable for the quality of unemployment services in the election following the reform. Further analyses show that the voters in these treatment municipalities did *not* hold their mayor more electorally accountable for the quality of services unaffected by the reform. This immediate adjustment in electoral accountability for unemployment services, and only for these services, in response to an exogenous and recent change in political responsibility, suggests that the relationship between political responsibility and electoral accountability is causal, adaptive and policy-specific.

This article extends the literature on how incumbents' political responsibility shapes voters' assignment of electoral credit and blame for policy outcomes—a literature that has, broadly speaking, remained observational and paid little attention to changes in political responsibility for discrete policy outcomes (e.g., Powell and Whitten 1993; Whitten and Palmer 1999; Anderson 2000; Duch and Stevenson 2008; Hobolt et al. 2013; Harding 2015). Further, with a fairly consensual multiparty system (Houlberg and Pedersen 2015), which garners less attention from voters than national politics (Elklit and Kjær 2013), the Danish municipalities provide a hard case in a literature that has mostly focused on national politics, where sharp divides between opposition and government as well as higher levels of political

attention potentially amplify the relationship between clarity of responsibility and electoral accountability (Duch and Stevenson 2008; De Vries et al. 2011). By demonstrating that voters are able to react in a reasonable way to a change in political responsibility from one level of government to another, this study also challenges the scope of research showing that voters have a hard time attributing responsibility in multi-level systems (Cutler 2008; Sances 2017; Johns 2011). Instead, the voters in this study seem to be quite judicious when assigning credit and blame for the quality of policy outcomes.

Challenges in the Study of Clarity of Responsibility

When studying the relationship between political responsibility and electoral accountability, existing research has typically: (1) indexed different elections according to how much political responsibility economic and political institutions assign to the incumbent up for reelection; (2) measured how accountable the incumbent was held in the same elections by correlating electoral support for the incumbent with a subjective indicator (e.g., perceptions of the national economy) or an objective indicator (e.g., economic growth) of the quality of policy outcomes; and (3) linked the responsibility index with the measure of electoral accountability in a statistical model. Using this approach, a number of scholars have explored the relationship between institutions that centralize political responsibility and electoral accountability using different indices of incumbent responsibility (e.g., Powell and Whitten 1993; Whitten and Palmer 1999; Nadeau et al. 2002; Anderson 2006; Hellwig and Samuels 2007; Duch and Stevenson 2008; Hobolt et al. 2013; Carlin and Singh 2015), different policy outcomes (Tavits 2007; De Vries et al. 2010; Tilley and Hobolt 2011) and different types of elections (Ebeid and Rodden 2006; Berry and Howell 2007; Larsen 2016). Broadly speaking, these studies have found that in elections where incumbents have more responsibility for policy outcomes, they are also held more electorally accountable for the quality of these outcomes.[1]

This wealth of thorough and innovative studies has gotten us a long way when it comes to understanding how political responsibility shapes electoral accountability. However, if one wants to draw more detailed inferences about the relationship between clarity of responsibility and electoral accountability from the extant literature, one faces several challenges. In particular, based on previous studies, one would have a hard time evaluating whether political responsibility has a causal effect on electoral accountability, how quickly voters adapt to changes in political responsibility, and the extent to which voters are sensitive to differences in how politically

[1] Another set of studies have examined which psychological processes lead voters to attribute certain outcomes to incumbent politicians (e.g., Gomez and Wilson 2001; Tilley and Hobolt 2011). While this literature also examines responsibility in relation to retrospective voting, it does so in a very different way than the literature discussed here. As such, in this more psychological literature, responsibility is a subjective belief that voters hold, whereas in the literature described above, responsibility is an objective condition determined by the mix of political and economic institutions that characterize the nature of policy-making in a specific polity.

responsible incumbents are for different policy outcomes. That is, one would have a hard time evaluating whether this relationship is causal, adaptive and policy-specific.

Before moving on, one important caveat deserves special notice. As already mentioned in the introduction, a large number of studies have looked at the relationship between the political performance of incumbents and electoral support for these incumbents. This includes examining whether and to what extent voters' reactions to political performance are adaptive, causal and policy-specific (for prominent examples of studies that adress these features, see Jones and Baumgartner (2005), Healy et al. (2010), Stimson (2015)). However, that is not what this study is about. Rather this study is about how voters react when the institutional context underpinning political performance changes. In terms of a causal model, we are thus *not* interested in the direct of effect of political performance on voter behavior, but in characterizing a moderator, namely centralization of political responsibility, of this effect. In particular, this study is interested in knowing know whether this moderator has a causal effect on the effect of political performance on voter behavior, whether the moderator takes effect immediately and whether the moderator works at the level of individual policies.

Causality

Previous studies have almost exclusively analyzed the relationship between clarity of responsibility and electoral accountability by looking at the correlation between the presence of institutions that manipulate incumbent responsibility and the extent to which voters hold incumbents electorally accountable (Duch and Stevenson 2008). At the same time, however, most researchers agree that one can rarely estimate the causal effect of institutions using standard observational studies (Meyer 1995; Besley and Case 2000; Aghion et al. 2004), because institutions are typically endogenous to the outcomes of interests (Acemoglu 2005; Przeworski 2004). Accordingly, it is possible that the effect of institutions that disperse political responsibility is confounded.

This possibility looms large if one takes a close look at some of the specific institutions that have been used to get at the relationship between clarity of responsibility and electoral behavior. For instance, a number of studies have shown that a country's economic openness is negatively correlated with how electorally accountable its executive is held for the economic situation (Hellwig 2001; Fernández-Albertos 2006; Hellwig and Samuels 2007; Duch and Stevenson 2008). These studies argue that this correlation is driven by the fact that economic openness decreases political responsibility for economic outcomes. However, economic openness is also known to be correlated with the extent to which countries provide social protection to those who are unemployed (Cameron 1978; Rodrik 1996), and we know from studies of economic voting that economically vulnerable voters are more likely to punish and reward governing politicians for the state of the economy (Singer 2013; Fossati 2014; Pacek and Radcliff 1995). Accordingly, when researchers find that voters are less likely to hold their government electorally accountable for the economic situation in countries with an open economy, this might be because open economies have

extensive social protections for their citizens, leaving citizens in these countries less worried about short-term fluctuations in the economy.

Another example of a potentially endogenous institution can be found in studies demonstrating that single-party governments are more likely to be held electorally accountable for the economic situation than multi-party governments (Anderson 2000; Nadeau et al. 2002; Hobolt et al. 2013). This might be because it is harder to asses who is responsible for economic outcomes in a coalition government, but it might also be the result of another difference between coalition and single-party governments. However, a number of studies have documented that partisans generally refrain from holding their own party electorally accountable for economic outcomes (Bisgaard 2015; Kayser and Wlezien 2011; Rudolph 2006), attributing any poor performance to some other factor than the competence of their preferred party (Tilley and Hobolt 2011). Coalition governments are typically larger, electorally speaking, than single-party governments. Accordingly, there will probably also be more voters who feel attached to a government party in a coalition government. When comparing the level of electoral accountability for single and multi-party governments, one may therefore be picking up the effect of differences in the number of government partisans rather than differences in the levels of incumbent responsibility.

These examples are not exhaustive in the sense that they cover all institutions that have been used to index how politically responsible the incumbent is for policy outcomes. Even so, these examples hopefully illustrate how the existing literature is challenged when it comes to identifying the causal effect of political responsibility on electoral accountability.

Adaptive

Another interesting feature of the relationship between clarity of political responsibility and electoral accountability is how adaptive it is; that is, whether voters respond swiftly to short-term changes in political responsibility, continually adjusting how accountable incumbents are held for various outcomes, or whether this adjustment process works more slowly.

Previous studies have not paid much attention to the question of adaptiveness, mainly focusing on differences in incumbent responsibility that rarely change or change slowly and incrementally (e.g., Ebeid and Rodden 2006; Duch and Stevenson 2008). This makes it hard to know how voters respond to sudden shifts in political responsibility. Some studies do examine more dynamic aspects of political responsibility, focusing on institutions which allocate different degrees of responsibility for policy outcomes to incumbent politicians over time within the same political unit (e.g., Nadeau et al. 2002; Carlin and Singh 2015). By focusing on this type of time-sensitive variation in the assignment of political responsibility, such studies could potentially tell us something about how adaptive the relationship with electoral accountability is. Yet these studies have rarely leveraged the dynamic nature of these institutions when examining how they affect electoral accountability. Instead they pool, either completely or partially, the within and between unit variation, making it impossible to get at whether the time-sensitive (within unit) variation in political

responsibility correlates with the extent to which voters hold incumbents accountable. One exception is Anderson (2009), who looks at how voters in Belgium hold their central government accountable for economic conditions just before and in the decade following a reform that "federalized" the Belgian constitution. This study follows Anderson's approach in studying a reform that changed political responsibility, but I also try to improve on his design by examining changes in accountability a few years after the reform rather than comparing how voters react, on average, in the decade following the reform.

Adaptiveness is potentially quite important because if voters are not adaptive, they risk holding their incumbent accountable for the quality of an outcome that the incumbent is no longer responsible for—or they risk failing to hold the incumbent to account for the quality of an outcome the incumbent has recently become responsible for. Also, a lack of adaptiveness can give incumbents an incentive to neglect policy areas where they have recently become more politically responsible because incumbents know that they will not be held accountable for their performance in these areas.

Policy-Specific

The existing literature has primarily examined the relationship between clarity of responsibility and electoral accountability in terms of institutions that affect how responsible incumbents are for a large set of policy outcomes, such as constitutional design (Anderson 2006; Carlin and Singh 2015), which broadly shapes incumbents' ability to affect economic and social outcomes, or different parliamentary practices (Powell and Whitten 1993; Whitten and Palmer 1999; Nadeau et al. 2002), which shape incumbents' executive and legislative discretion across all policy areas. This focus on responsibility for a diverse and not clearly demarcated set of outcomes has made it difficult to assess how policy-specific voters are when they hold incumbents electorally accountable. In particular, we do not know whether voters link responsibility to accountability at the level of individual policy outcomes, weighing each outcome according to how responsible the government is for that specific outcome, or whether voters link political responsibility to electoral accountability at a more aggregate level, using different policy outcomes to form an overall evaluation of how their polity is doing, and then weigh this overall evaluation based on how responsible the incumbent is for policy outcomes in general.[2]

The previous literature cannot discriminate between a policy-specific and a more general relationship, because it looks at differences in incumbent responsibility for a diffuse set of policy outcomes (although for important exceptions, see Arceneaux

[2] It is not theoretically straightforward to predict which of these approaches voters will adopt. On the one hand, adopting a policy-specific strategy seems to be more rational if one simply wants to learn more about the incumbent's competence (for evidence of this, see the appendix of Achen and Bartels (2016)). On the other hand, voters are often interested in employing heuristics and mental shortcuts (Downs 1957; Kuklinski et al. 2000). One such mental shortcut might be to link responsibility and accountability at an aggregate rather than at a policy-specific level.

2006; Ruder et al. 2014). If one wanted to make inferences about policy-specificity, then one would need to examine a difference in political responsibility that only covered a discrete set of policy outcomes. In this case, it would be possible to examine policy-specificity by investigating whether voters *only* differed in how electorally accountable they held the incumbent for the policy outcomes for which there was an underlying difference in political responsibility, or whether electoral accountability for other outcomes was affected as well.

The Contribution of this Study

By focusing on a reform that changed political responsibility for a specific policy, the present study enables us to get at *policy-specificity*. That is, we can examine whether voters only hold local incumbents more electorally accountable for the policy outcome affected by the reform, or whether voters hold incumbents more electorally accountable for other policy outcomes as well. Further, because we examine the effect of the reform at the first election after its implementation, any effect that we do find will reflect a relationship that is reasonably *adaptive*. Finally, as discussed in more detail below, the implementation of the reform that changed political responsibility was arguably exogenous, making it possible to identify the *causal* effect of the reform on electoral accountability.

Analyzing whether the relationship between centralization of political responsibility and electoral accountability is causal, adaptive and policy-specific is important because it tells us something about how adept voters are at electing competent politicians. In particular, all these factors make it more likely that voters only hold incumbents electorally accountable for outcomes the incumbent had a hand in shaping, which should, in turn, make it easier for voters to identify whether incumbent politicians have their best interests in mind, and reward them with reelection if they do (cf. Anderson 2006; Duch and Stevenson 2008; Ashworth 2012; Achen and Bartels 2016).[3]

Research Design: Reform of Labor Market Regulation

In 2006 the administrative boundaries of Denmark were fundamentally redrawn, both in terms of geography, as 271 municipalities became 98, but also in terms of policy responsibilities, as the municipalities gained new responsibilities and lost others. Unfortunately, from a research standpoint, most of this extensive reform was implemented in all municipalities at the same time, making it hard to test how it affected the municipalities (although aspects of the reform have been leveraged in

[3] Too see this, note that if the relationship between centralization of responsibility and accountability is causal, then voters respond to changes in responsibility by holding incumbents more electorally accountable. If the relationship is policy-specific, then voters are more likely to shift their attention away from policy outcomes that incumbents have little responsibility for and towards outcomes that incumbents have more responsibility for. If the relationship is adaptive, then voters are more likely to act on the current distribution of political responsibility when holding incumbents accountable.

other contexts: see Lassen and Serritzlew (2011), Bhatti and Hansen (2011), Blom-Hansen et al. (2014)). One part of the reform, however, was not implemented at once, but in two steps: a reform of labor market regulation, which transferred the political responsibility for unemployment services from various agencies to municipal mayors.

The goal of this reform was to reduce the cost and increase the effectiveness of unemployment services (Eskelinen 2008). In short, unemployment services consist of advising unemployed workers, finding and financing retraining, assisting unemployed workers with special needs and helping employers look for employees.[4] Before the reform, unemployment services were provided by the national government, private unemployment insurance funds (if the unemployed citizen was a member) and by the municipality that the unemployed citizens resided in. The reform centralized responsibility for unemployment services in the hands of the municipalities. By doing this, the central government hoped to reduce the transaction costs involved in the old system, where several actors needed to cooperate, and they also wanted to give municipalities the opportunity to experiment more freely with what kind of unemployment services were most effective (Ministry of Employment 2010).

When the idea for the reform was initially floated, it was met with fierce resistance from the unions, who ran the unemployment insurance funds, as well as from the opposition in parliament, who were afraid that the reform would mean less generous services (Eskelinen 2008). Partly as a response to this, the reform was implemented in two steps, with full implementation being contingent on successful early implementation. In the end, 14 municipalities (out of 98) implemented the reform in the beginning of 2007 and the rest in 2010 (Order 1400 2006). The reform was narrow in scope, focusing only on these unemployment services. There was extensive debate about the reform, and it was the subject of a good deal of public debate (Eskelinen 2008), however, there were, as far as I have been able to determine, no concerted effort to inform citizens. See Sect. S1 of the supplementary materials for more details on scope of the reform.

The Labor Market Reform provides a unique opportunity to investigate how voters react when political responsibility is centralized. As such, we can use the municipal elections that took place in 2009 to compare the beliefs and behavior of the voters in the 14 municipalities where the mayor got more responsibility for unemployment services before the election—the treatment municipalities—with the beliefs and behavior of the voters in the 84 municipalities where the mayor did not get more responsibility until 2010—the control municipalities. If the relationship between political responsibility and electoral accountability is causal, adaptive and policy-specific, then voters in the treatment municipalities should hold their mayor

[4] Unemployment services constitute an important part of public service provision in Denmark, and Danish labor market policy has long been premised on the idea that the day-to-day interaction with the unemployed individual is important for reducing structural unemployment (Torfing 1999). This idea is mirrored in spending priorities. According to the OECD, expenditures towards unemployment services (i.e., active labor market policies) represented 1.82% of the Danish GDP in 2013 compared to just 0.23% in the United Kingdom (OECD 2014).

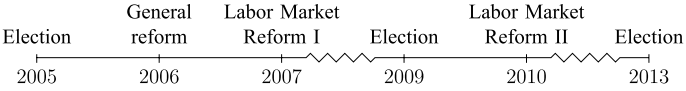

Fig. 1 Labor market reform timeline. Labor market reform I was in the 14 "treatment" municipalities, Labor market reform II was in the remaining "control" municipalities

more electorally accountable for unemployment services than voters in the control municipalities. See Fig. 1 for a timeline of the reform and its relation to the timing of municipal elections.

To draw such inferences, however, one needs to make two assumptions about the reform. One is about the nature of the reform; that the reform exclusively affected municipal mayors' political responsibility for unemployment services. The other is about the assignment of the municipalities to early implementation of the reform (i.e., assignment to treatment); that assignment was independent of existing and potential levels of electoral accountability.[5] Below, I explain why it is reasonable to make these assumptions about the reform, and then I present the data used to study the effect of the reform.

The Reform Only Affected Responsibility for Unemployment Services

The reform of labor market regulation made municipalities politically responsible for helping so-called "insured workers" get back to work if they lost their job (i.e., the three-fourths of all workers who were members of an unemployment insurance fund). Before the reform was implemented, the national government and the unemployment insurance funds were responsible for the insured workers, whereas the municipalities were responsible for uninsured workers. The reform removed the unemployment insurance funds and the national government from the equation and gave each municipality unilateral responsibility for all those who were out of a job in that municipality (Order 1400 2006; Eskelinen 2008).

It is important to note that the reform did not simply increase local politicians' "functional responsibility" for unemployment services (i.e., sense of obligation for unemployment services), but also their 'causal responsibility" (i.e., opportunity to affect the quality of unemployment services) (for details on these concepts, see Arceneaux (2006), p. 735). Put differently, after the reform, the municipalities had more power—in the form of policy discretion and resources—to shape unemployment services for the better or for the worse.

While the reform ostensibly had an effect on who was politically responsible for unemployment services, reforms tend to be messy and have a very diverse set of long and short term consequences. In light of this, one might suspect that the

[5] These assumptions roughly correspond to the exclusion and independence (or exogeneity) assumptions laid out by Dunning (2012) and Gerber and Green (2012). Along with the assumption of non-interference between units, they constitute the central assumptions needed to draw causal inferences. We do not discuss the non-interference assumption in detail, because political responsibility could not spillover to neighboring municipalities.

assignment to the implementation of the labor market reform had important side effects that could pose threats to the inferences I want to make below. However, if one examines the nature of the reform in more detail, such potential side effects are not forthcoming. Instead, the reform presents a very clean change in political responsibility for a specific policy: no responsibilities outside the area of unemployment services were conferred and no alternative regulation was implemented as part of the reform (Ministry of Employment 2010; Act 483 2009).[6]

The Change in Responsibility was Exogenous

If one wants to draw causal inferences based on the selection of some municipalities for the early implementation of the labor market reform, then this selection process should be independent of existing and potential levels of electoral accountability (Gerber and Green 2012). If it is not, one risks confounding the effect of the reform with the effect of being the type of municipality that is assigned to early implementation. We look for quantitative evidence of this type of selection below (cf. Table 1). However, before we do so, we want to note that several factors surrounding the assignment of municipalities to early implementation makes it likely that it was in fact independent of existing and potential levels of electoral accountability.

First, the selection process was confined to municipalities within a single country at a single point in time. This makes it possible to rule out a host of possible confounders, such as macro-social developments and country-specific factors such as political culture and history.

Second, the final decision about which municipalities were assigned to implement the reform early was made by the central government rather than the municipalities themselves. In particular, employees at the ministry as well as the minister prepared a list of municipalities that was then approved by the parties that voted for the reform in Parliament (Ministry of Employment 2006).

Third, and most importantly, it seems likely that the ministry's assignment of municipalities to early implementation was independent of the municipalities' existing or potential levels of electoral accountability for unemployment services. For one, it is not clear that the ministry would have known what the level of electoral accountability for unemployment services was in the individual municipalities. Even if the ministry knew the levels, it is not clear that the ministry would have had an incentive to assign municipalities to early implementation based on these levels. There could feasibly have been an incentive to pick municipalities that generally fared better when it came to handling unemployment services because these were more likely to make the reform look like a success (although I do not find any such imbalances between early and late implementers, cf. Table 1), but it is unclear why the ministry should be interested in implementing the reform in places where the level of electoral accountability for unemployment services was particularly high (or low). Finally, even if the ministry did know and, for some reason, favored types of

[6] See Sect. S1 of the supplementary materials for some additional evidence of the fact that the reform did not have any important side effects.

Table 1 Were treatment and control municipalities different?

Variable	Treatment	Control	SD	p value	n
Individual-level variables (2005)					
Informed	0.44	0.45	− 0.07	0.16	1996
Interested	0.63	0.64	− 0.06	0.31	1884
Unemployment performance	0.47	0.48	− 0.03	0.68	1454
Knowledge about municipal powers	0.70	0.71	− 0.05	0.37	2011
Elderly performance	0.64	0.67	− 0.09	0.15	1534
Housing performance	0.74	0.75	− 0.05	0.40	1944
Ideology	0.69	0.70	− 0.00	0.97	2011
Apathy	0.14	0.15	− 0.00	0.93	1988
Obligation	0.95	0.96	− 0.07	0.25	2000
Satisfaction with municipal democracy	0.52	0.52	− 0.02	0.71	1975
Pivotality	0.45	0.46	− 0.00	0.95	1875
Municipality-level variables (2006)					
Population density (log)	2.18	2.25	− 0.12	0.60	98
Citizens with more than high-school education (%)	23.10	21.54	0.19	0.51	98
Unemployment rate (%)	2.16	2.31	− 0.14	0.48	98
Citizens with non-Western origins (log)	2.39	2.32	0.27	0.27	98
Female municipal office-seekers (%)	28.13	30.39	− 0.46	0.13	98
Municipal tax rate (%)	24.55	24.77	− 0.23	0.31	98
Social transfers (log)	4.13	4.13	0.03	0.91	98
Services contracted out (0–100 scale)	21.90	22.78	− 0.22	0.35	98
Spending on active labor market policies (log)	3.21	3.20	0.06	0.77	98
Inhabitants (log)	4.82	4.58	0.69	0.00	98
Work in service-industry (%)	0.44	0.42	0.17	0.47	98
Work in manufacturing industry (%)	0.24	0.24	0.04	0.87	98
National government voters	0.42	0.38	0.31	0.17	98
National government mayors	0.64	0.45	0.38	0.20	98

Individual-level variables from the 2005 municipal election survey, see Sect. S2 of the supplementary materials for a detailed description. Municipal-level variables taken from Statistics Denmark. p-values from difference in means test. National government voters is the proportion of voters who voted for parties in government at the municipal election in 2005. Standardized difference computed as difference in means divided by standard deviation in the control group. Heavily skewed variables presented on a logarithmic scale

municipalities that had higher levels of electoral accountability, there were political forces at work that, arguably, muted any political favoritism.

When the reform was being negotiated, several actors were highly critical of giving the municipalities responsibility for unemployment services. As such, both the large unions and employer organizations as well as the minority government's usual ally in parliament, the Danish People's Party, were doubtful that the municipalities were up to the task (Kristensen 2008, p. 88). Accordingly, there was pressure on the Ministry of Employment not to "cherry pick" municipalities based on

past performance. As a person close to the selection process expressed it: "We were allowed to send up a test balloon, but it was extremely important that they [the municipalities] were balanced".[7] This sentiment is mirrored in a press statement published by the Ministry of Employment explaining how the 14 municipalities had been selected. In the statement, the Minister of Employment was quoted as saying that the goal had been to select "large as well as small municipalities, in cities as well as in rural areas". More generally, the Minister said "that the goal was to spread them out across the country" (Ministry of Employment 2006, author's translation). As such, specific types of municipalities were not targeted in the selection process. This is confirmed if we look at how the chairman of the organization Local Government Denmark, an organization representing the municipalities, reacted to the selection process. He said that the Ministry's decision insured that "a broad cross section of municipalities [are represented], both size-wise and geographically." (Ritzau (2006), author's translation). Additional evidence suggesting that the selection process was not politically motivated can be found if one looks at the reaction to the Ministry's decision among those who were very critical of the reform: the unions and the employer organizations. As far as I have been able to determine, none of these political organizations officially criticized the government for having selected a biased or problematic set of municipalities for early implementation.[8]

Taken together, these factors suggest that the selection of the 14 early-implementing municipalities was independent of existing and potential patterns of accountability. That is, based on the evidence presented here, there is reason to believe that the change in political responsibility for unemployment services was exogenous. This assertion is revisited below, where we show that important variables are balanced across treatment and control municipalities (cf. Table 1).

Data and Measuring Electoral Accountability

To analyze the electoral consequences of the reform, I use the Danish municipal election survey (Elklit and Kjær 2013). The 2009 election survey is of special interest, since this is where electorates in the treatment and control municipalities were governed by mayors with different levels of responsibility for unemployment services (cf. Fig. 1). Even so, the 2013 and 2005 surveys are used as well to test whether the electorates of the treatment and control municipalities differed before treatment (2005) and after all municipalities were treated (2013). Respondents in the municipal election surveys were recruited within six weeks of the municipal election using stratified random sampling in order to ensure that at least 30 respondents in each of the 98 Danish municipalities participated in the survey. The surveys are

[7] Interview with Jan Handeliowitz, former employee at the Ministry of Employment. Author's translation.

[8] This conclusion is based on an examination of all newspaper stories mentioning the reform in the month following the announcement of the assignment of municipalities to early-implementer status in the three major Danish broadsheets (Jyllands Posten, Politiken and Berlingske).

conducted partly via a web-survey and partly over the phone (for details about the surveys, see Elklit and Kjær (2013)).[9]

To measure the extent to which the mayor was held electorally accountable for unemployment services, I examine the correlation between voters evaluation of unemployment services and their propensity to support the municipal mayor (a typical measure of electoral accountability, cf. Stevenson and Duch (2013), Duch and Stevenson (2008), Carlin and Singh (2015)), interpreting a higher correlation as evidence that the mayor is being held more electorally accountable for the quality of unemployment services. To assess voters' evaluation of unemployment services, the following survey item is used: "How satisfied or unsatisfied are you in general with the municipality's efforts towards the unemployed?" Answers are recorded on a five-point Likert scale going from "very dissatisfied" to "very satisfied".[10] To measure support for the incumbent mayor, I look at whether respondents reported voting for the incumbent mayor's party at the municipal election. Respondents who did not vote and respondents who could not remember which party they voted for are omitted from the analysis.[11]

This measure of electoral accountability is not perfect, and will probably contain some measurement error. In particular, the measure might also capture, at least in part, the extent to which voters form beliefs about unemployment services based on who they vote for (so-called motivated reasoning, cf. Tilley and Hobolt (2011)). Accordingly, this measure might overestimate the *level* of electoral of accountability in each municipality. Even so, we will still be able to get an unbiased estimate of the *difference* between treatment and control as long as this measurement error is not correlated with treatment status (King et al. 1994, chap. 5).[12]

All survey items used in the analysis are described in Sect. S2 of the supplementary materials, and descriptive statistics on all variables can be found in Sect. S3 of the supplementary materials.

Analysis

The main goal of this analysis is to find out whether voters in the treatment municipalities held their mayor more electorally accountable for the quality of unemployment services than voters in the control municipalities. The analysis will also explore whether the mayors in the treatment municipalities were held more (or less) accountable for unemployment services in 2005, whether the national or regional

[9] The 2005 survey differs in this respect as it is not stratified according to municipality.

[10] The survey item on unemployment services was not included in the 2013 survey. Therefore, I cannot measure electoral accountability for unemployment services in the 2013 election.

[11] Support for the mayoral party is used to measure support for the mayor because voters do not elect mayors directly in Denmark. Rather, they elect members of a city council, and the city council then appoints a mayor right after the election (Houlberg and Pedersen 2015). Municipal elections in Denmark are held every 4 years in November. The electoral system is proportional representation and most municipalities have a multi-party system that mirrors the national party system.

[12] In "Alternative Explanations and Potential Mechanisms" this assumption is discussed further and tested empirically (see also Sect. S8 of the supplementary materials).

governments were held more accountable for unemployment services by voters in the treatment municipalities, and whether mayors in the treatment municipalities were held more electorally accountable for the quality of other services. In addition to this, I discuss the viability of some alternative explanations and the mechanism underlying the electoral effects of the reform. Before these analyses are presented, however, a balance test and a manipulation check is laid out in order to investigate whether assignment to treatment (ie., early implementation of the reform) was exogenous to electoral accountability, and whether being assigned to treatment had an impact on voters' beliefs about the distribution of political responsibility.

Balance Test and Manipulation Check

Table 1 compares treatment and control municipalities before they were treated on a number of individual-level and municipality-level variables. For the individual-level variables there are no statistically meaningful differences. Most importantly, treatment and control are balanced on several variables that should be correlated with voters' penchant for holding their mayor accountable, such as knowledge about municipal powers and interest in local politics (Vries and Giger 2014). This is consistent with the qualitative evidence laid out above, which suggests that implementation of the reform was assigned to municipalities independently of existing levels of electoral accountability. If particular types of electorates had a higher probability of being assigned to early implementation, it seems likely that one would be able to identify systematic differences across treatment and control municipalities, but no such differences are identified.

The municipal-level variables paint roughly the same picture. Across the different variables, only one shows a significant difference between the two groups—the treatment municipalities had a slightly larger number of inhabitants than the control municipalities. Even so, examining the standardized differences for the remaining municipal-level variables, there does seem to be some substantial, though statistically insignificant, differences across treatment and control. This is not that surprising. The number of observations at the municipal-level is relatively low, which means that the random variation between treatment and control could be relatively high. Nonetheless, these random imbalances might skew the results one way or another. When analyzing the differences between the treatment and control municipalities below, I take this issue into account by controlling for the municipal-level variables that have the largest standardized differences (i.e., proportion of national government voters/mayors, female office seekers, non-Western citizens and number of inhabitants).

Another relevant issue is whether voters in the treated municipalities actually updated their beliefs about the mayor's responsibility for unemployment services; that is, whether the reform actually registered with the voters. Unfortunately, there is no question in the municipal election survey that directly probes voters' beliefs about the extent of their mayor's responsibility for unemployment services. However, there are two questions asking respondents about how much political responsibility local politicians have for conditions in the municipality in

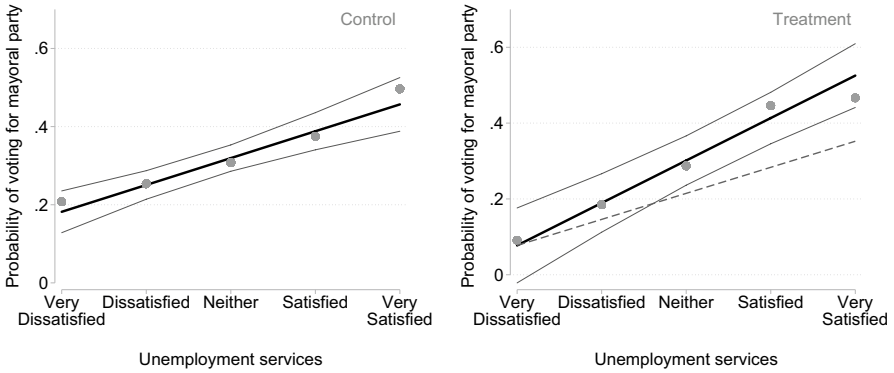

Fig. 2 Solid lines reflect the predicted probability of voting for the mayoral party across satisfaction with unemployment services in treatment and control municipalities with 95%. confidence intervals. The predictions are derived from a linear probability model with a treatment by unemployment services interaction. Standard errors are clustered at the municipal level. Interaction estimate is statistically significantly different from zero ($t(1521) = 2.08, p = 0.04$). Dots are conditional probabilities estimated from the sample ($n = 1522$). The dashed line in the "Treatment" plot mirrors the slope from the "Control" plot

general. The first of these questions ask voters whether the mayor and other local officials (rather than national politicians) has the primary responsibility for how the municipality developed in the last 4 years. The second question asks voters about the extent to which the mayor has had an effect on the well-being of the municipality. Analyzing voters' responses to these questions, I find that the voters in the treatment municipalities believed that their local politicians were more responsible for and had a greater influence on conditions in their municipality ($p < 0.05$; see Sect. S4 of the supplementary materials for details).

Electoral Accountability for Unemployment Services

Figure 2 plots the conditional probability of supporting the mayoral party in the 2009 election across voters' satisfaction with unemployment services in the treatment and in the control municipalities. The figure also plots a linear fit of the relationship between voters' satisfaction with unemployment services and support for the mayoral party. This graphical analysis allows us to compare the extent to which voters' evaluations of unemployment services shape incumbent support in the treatment versus the control municipalities.

The figure shows that support for the mayoral party was more closely related to voters' evaluation of unemployment services in the treatment municipalities. Accordingly, the increase in local political responsibility for unemployment services seems to be associated with an increase in the extent to which voters punished and rewarded local incumbents for the quality of these services.

To investigate further I estimate a regression model which sets the probability that the respondent voted for the mayoral party as a logistic function of the respondent's

Table 2 Logistic regression of probability of voting for the mayoral party

	(1)	(2)	(3)	(4)
Unemployment performance	0.82*	0.83*	0.94*	0.85*
	(0.23)	(0.23)	(0.36)	(0.35)
Treatment	− 0.61	− 0.59	− 0.52	− 0.63
	(0.39)	(0.40)	(0.43)	(0.41)
Treatment * unemployment performance	1.00*	0.98^{+}	1.25*	1.33*
	(0.51)	(0.52)	(0.62)	(0.58)
Administration controls municipality	− 0.36^{+}	− 0.46*	− 0.48*	− 0.43^{+}
	(0.21)	(0.21)	(0.24)	(0.24)
Elderly performance	0.76*	0.81*	1.03*	1.03*
	(0.23)	(0.25)	(0.33)	(0.30)
Housing performance	0.38	0.31	− 0.13	0.04
	(0.29)	(0.30)	(0.38)	(0.39)
Sociodemographic controls		✓	✓	✓
Political controls			✓	✓
Municipal level variables				✓
AME (Control)	0.17	0.17	0.10	0.09
	(0.05)	(0.05)	(0.04)	(0.04)
AME (Treatment)	0.36	0.35	0.26	0.25
	(0.08)	(0.08)	(0.06)	(0.05)
Difference (T–C)	0.19	0.18	0.15	0.15
p value of difference	0.03	0.04	0.04	0.01
Pseudo R^2	0.04	0.05	0.41	0.42
Log likelihood	− 921	− 894	− 561	− 552
Observations	1522	1500	1500	1500

Robust standard errors clustered on municipality in parentheses.

$^{+}p < 0.10$, * $p < 0.05$

evaluation of the municipality's performance in the area of unemployment services, an indicator variable determining whether the respondent lived in a treatment or a control municipality, as well as an interaction between the two. The model also includes a small number of control variables: voters' satisfaction with elderly care service and housing management as well as their beliefs about how powerful the municipal administration is.[13] These variables are meant to reduce the error term of the model and control for trends in performance and beliefs about how responsible local politicians are for economic and social outcomes. Since the treatment was assigned to municipalities, standard errors are clustered at the municipal level when estimating the model. The estimates from this model are presented in column one of Table 2.

[13] I also estimated a simpler logistic model, without any controls. The interaction estimate in this simple model is also statistically significant and of roughly the same size as the one presented in column one of Table 2.

The primary estimate of interest is the interaction between treatment and unemployment performance. This coefficient indicates whether voters' evaluation of the quality of unemployment services was more or less closely tied to the propensity to vote for the mayoral party in the treatment municipalities. The coefficient is statistically significant ($p < 0.05$) and positive. This means that satisfaction with the municipality's unemployment services mattered more in the municipalities where the mayor had more political responsibility for unemployment services. The remaining coefficients in the model have the expected sign, and, apart from housing performance, are all statistically significant.

A similar picture emerges if one derives the average marginal effects (AMEs) of satisfaction with unemployment services for the treatment and control municipalities (for a description of the statistical properties of AMEs, see Hanmer and Ozan Kalkan (2013)). The AMEs are reported at the bottom of column one in Table 2, and they reveal that in the control municipalities, the result of going from one end of the unemployment scale to the other is an average increase in the probability of voting for the mayoral party of 17 percentage points. In the treatment group, the result is an increase of 36 percentage points, a difference that is statistically significant ($p < 0.05$) and quite large: the AME for voters in the treatment municipalities is twice that for voters in the control municipalities.

These analyses indicate that there is a causal relationship between political responsibility and electoral accountability. In the municipalities where the mayor was exogenously assigned more responsibility for unemployment services, the mayor's party was also held more electorally accountable for the quality of these services. The analyses also imply that there is quite an adaptive relationship between centralization of political responsibility and electoral accountability. The reform which affected how responsible the municipal mayor was for unemployment services was implemented just 3 years prior to the election analyzed above. In spite of this, voters responded to the change, holding their mayor more electorally accountable.

Below, I show that these results are robust. In particular, I add more controls to the regression model estimated above and then I try out some alternative estimation techniques.

Additional Controls

First, a number of socio-demographic variables are included in the regression model (i.e., age, gender education, occupational status, and local media consumption). The controls are included in the model estimated in the second column of Table 2. Introducing these controls only shifts the estimates slightly, yet the difference in AMEs between the treatment and control municipalities remains statistically significant ($p < 0.05$). Second, a set of political variables are added to the model. To gauge partisanship, I include a control for whether the respondent would vote for the mayor's party if a *national* election was held tomorrow and a variable indicating whether the respondent shares the ideological orientation of the mayor. I also include a variable indicating whether the mayor was from the same party as one of the governing parties at the national level, something that might make the mayor susceptible to blame

for unemployment services (cf. Cutler 2008). These political controls are included in the third column of Table 2. Introducing the political controls reduces the *overall* AME of unemployment performance but does not substantially reduce the *difference* between the treatment and control municipalities, which remains statistically significant. Finally, a battery of municipal-level control variables are included. The municipal-level variables included are the ones for which the balance test revealed a substantial imbalance (cf. above). These controls are included in column four of Table 2. The interaction effect and difference in AMEs remain statistically significant in this specification as well.

Alternative Estimation Methods

Section S5 of the supplementary materials examines whether the results are sensitive to alternative ways of estimating the interaction effect and its sampling variability. Specifically, I use a multi-level logit model and a form of randomization inference. Using these alternative estimation methods, the difference in the AME of unemployment performance between treatment and control municipalities remains statistically significant ($p < .05$ for multi-level models, $p < 0.1$ for randomization inference).

Some Additional Tests

The balance check and discussion of selection into early-implementer status suggest that we can make causal inferences about the effect of the reform by comparing treatment and control municipalities. However, treatment and control municipalities might still, for one reason or another, be unbalanced on some unobservable variables that also affect accountability for unemployment services. To explore whether this is the case, one can look at whether there were differences in the extent to which voters held the mayoral party accountable before implementation of the reform.

To do so, we pool data from the 2005 and 2009 municipal election surveys in order to estimate a model that examines whether the difference in the effect of unemployment services between treatment and control municipalities is different across the two elections. Because of limited overlap in the questions asked in the 2009 and 2005 surveys, we can only use a small subset of the controls used in the analysis above (see Sect. S6 of the supplementary materials for details about which controls we use). It is also important to note that because of the large municipal reform in 2006, in which a large number of municipalities were amalgamated, the 2005 and 2009 data are not perfectly comparable (see Sect. S6 of the supplementary materials for details).

The key estimates from this difference-in-difference analysis, the AME of satisfaction with unemployment services across elections and treatment status, are plotted in Fig. 3. There are no apparent differences across treatment and control municipalities in 2005, and the difference identified in 2009 is statistically distinguishable

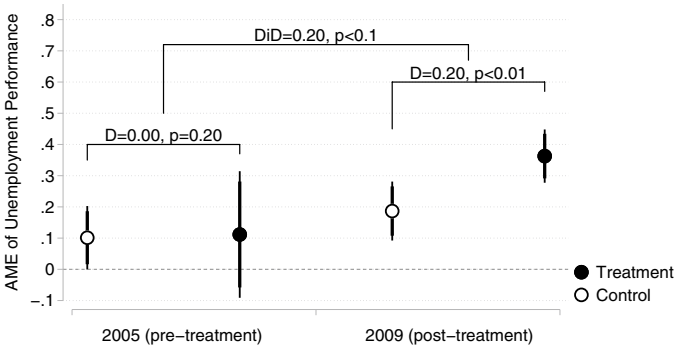

Fig. 3 Average marginal effects of unemployment performance on probability of voting for the mayoral party across election and treatment status. Derived from logistic regression model described in Sect. S6 of the supplementary materials; McFadden $R^2 = 0.032$, $n = 2582$. Wald tests used to compare the different AMEs. The vertical lines are 90% (thick) and 95% (thin) confidence intervals

from the difference in 2005 ($p < 0.1$). This analysis thus reaffirms the findings from the more simple cross-sectional analysis presented in Table 2.

Beyond these temporal dynamics, it is interesting to look at how voters hold governments at other levels accountable in treatment and control municipalities. In particular, we might expect voters to hold regional governments equally accountable for unemployment services across treatment and control municipalities, as there are no changes in their power over unemployment services, and the national government less accountable, as they have less power over unemployment services in the treatment municipalities. This analysis might speak both to the credibility of our causal inferences, examining whether electorates in treatment municipalities hold incumbent politicians more electorally accountable irrespective of how politically responsible these incumbents are, and as an additional empirical implication of an adaptive, causal and policy-specific relationship between clarity of responsibility and electoral accountability. To explore this expectation we re-estimate the model from column four of Table 2 using support for the regional government party and support for the national government parties as the dependent variables.[14] The key estimates from these two different models, the AME of satisfaction with unemployment services across treatment and control municipalities, are presented in Fig. 4.

We find no statistically significant differences across treatment and control municipalities, but in our sample the national government is held a little less accountable for unemployment services in the treatment municipalities. In general, however, national governments are not held very accountable for unemployment services. The AME are statistically indistinguishable from zero in both treatment and control municipalities. One reason for this might be that the survey item measuring satisfaction with unemployment services asks specifically about services *in the respondents municipality*, not in the country as a whole. The lack of a comparable measure of citizen satisfaction with unemployment services at the *national* level

[14] I adapt the models in one way, swapping the measure of support for the mayoral party at national elections for a measure of support for the regional/national government party/parties.

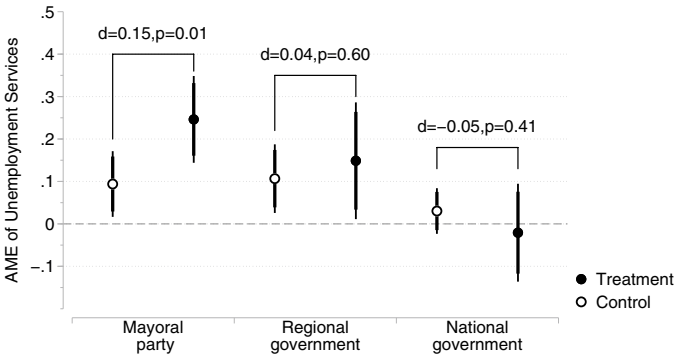

Fig. 4 Average marginal effects of satisfaction with unemployment services on probability of voting for either the mayoral party, the regional government party or the national government party. All average marginal effects derived from an augmented versions the model presented in the fourth column of Table 2. Wald tests used to compare the different AMEs. The vertical lines are 90% (thick) and 95% (thin) confidence intervals

might deflate the AME estimates, attenuating the observed difference between the control and treatment municipalities. One potential objection to this line of reasoning is that there is no question directed at the regional level either, but in spite of this we see relatively large AME's here. However, this might be explained by the fact that the regional elections are concurrent with the municipal elections (as opposed to the national elections), and therefore people might have their regional government more top of mind, or by the fact that the mayoral party and the regional party tends to overlap more than the national government party and the mayoral party.

The implication of this analysis is not very clear. While it is reassuring that we find no difference in the extent to which the regional government is held accountable for the quality of unemployment services, we would also expect the national government to be held less accountable in the treatment municipalities. However, the lack of any substantial difference in accountability at the national level might simply reflect that we do not have a good measure of citizen satisfaction with unemployment services at the national level. In sum, I think these results are consistent with the notion of a causal and adaptive relationship between centralization of political responsibility and electoral accountability, but they do not lend independent support to this notion.

Electoral Accountability for Other Outcomes

If voters link responsibility and accountability at a general level, rather than at the level of each specific policy, then changes in political responsibility for one policy outcome should lead voters to hold elected officials more electorally accountable for other policy outcomes as well. To investigate whether this was the case in this context, I look at how electorally accountable voters in the treatment and control municipalities held their mayor for the quality of two types of public services *not* affected by the labor market reform: housing and elderly care. These two types of services

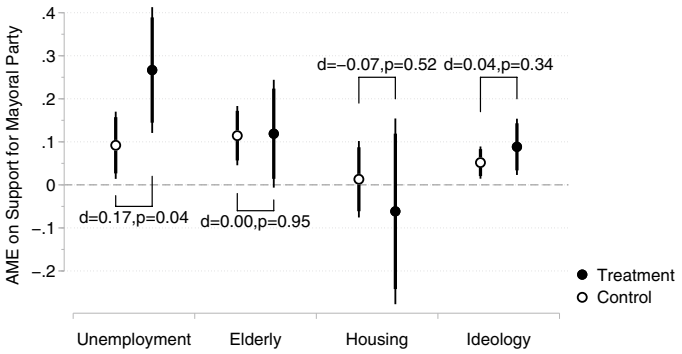

Fig. 5 Average marginal effects of different variables on probability of voting for the mayoral party in treatment and control municipalities. All average marginal effects derived from an augmented version the model presented in the fourth column of Table 2 which includes interactions between each variable and the treatment indicator; McFadden $R^2 = 0.41$, $n = 1500$. Wald tests used to compare the different AMEs. The vertical lines are 90% (thick) and 95% (thin) confidence intervals

are privileged because they represent a key policy that local governments typically deal with, housing, and a service that shares several features of the unemployment services examined above, elderly care.[15] In addition to this, I look at ideological congruence with the mayoral party to see if voters are more electorally responsive to their mayor's ideological orientation in the treatment municipalities.

To examine electoral accountability for these alternative policies, an extended version the model presented in the fourth column of Table 2 is estimated. This model adds interactions between the treatment indicator and voters' satisfaction with their municipality's housing management, their municipality's elderly care and voters' ideological congruence with their municipality's mayor. Figure 5 graphs the key estimates derived from this extended model—the AMEs of the alternative policy variables across voters living in the treatment and control municipalities. For comparison, the AMEs of unemployment performance are also plotted.

As can be seen from Fig. 5, there are no statistically discernible differences across treatment and control municipalities for the AMEs of elderly care, housing and ideology. This is in contrast to unemployment services, where there is a substantial and statistically significant difference.[16]

In sum, there is no sign that voters held their mayor more electorally accountable for other policy outcomes than the one for which additional political responsibility was conferred, suggesting that the relationship between centralization of political responsibility and accountability is policy-specific.

[15] In particular, elderly care only directly affects a certain target population (i.e., the elderly), similar to how unemployment services only affect the unemployed. Elderly care is also similar to unemployment services in that it is a public service consisting of direct contact with municipal employees.

[16] Section S7 of the supplementary materials analyzes the robustness of these results by running similar analyses for a number of other policy areas. Among the seven additional policies examined, there is not a single statistically significant difference between the treatment and control municipalities.

In addition to policy-specificity, these findings also speak to the question of whether the effect of the reform identified in Table 2 is causal. If there were differences across treatment and control municipalities in how accountable voters held their mayor for some other type of performance, then this might have been because the electorates in the treatment municipalities were more likely to attribute political responsibility for all types of outcomes to the mayoral party in the 2009 election—irrespective of any objective differences in political responsibility. In this sense, the analysis of the other policy areas can also be interpreted as a placebo test.

Alternative Explanations and Potential Mechanisms

When a policy area is subject to increased political subjugation it seems natural that it will become more politically contested. Accordingly, increased local political responsibility for unemployment services in a municipality could have led to increased local political attention to this issue in the campaign, priming the issue in the minds of the voter (Krosnick and Kinder 1990). Based on this, one might wonder: is the change in electoral accountability for unemployment services really based on the fact that this issue was "primed" in the treatment municipalities? This might be the case. More generally, priming may help explain why political responsibility is related to electoral accountability: responsibility leads to attention and attention leads to accountability. Some evidence suggests that priming can play such a role (Ruder et al. 2014; Hart 2016). If this is the case, it could pose a challenge to some of the inferences made above.

For one, increased attention to unemployment services might have led to increased polarization between mayoral partisans and non-partisans in beliefs about unemployment services. In particular, if unemployment services were framed by elite actors as a more important issue in the treatment municipalities, then mayoral partisans would also be more strongly motivated to engage in partisan rationalization when forming their beliefs about the quality of unemployment services (for evidence of such a mechanism, see Parker-Stephen (2013)). If this is the case, then the increased correlation between voters' assessment of unemployment services and support for the incumbent mayor might not just be a result of voters holding the mayor more electorally accountable, but also of voters relying more on their partisan preconceptions (i.e., reverse causation from voting to beliefs about unemployment services might be stronger in the treatment municipalities). To explore the viability of this alternative explanation, Sect. S8 of the supplementary materials examines whether beliefs about unemployment services are more correlated with *past* support for the mayoral party in the treatment municipalities. There is no evidence of this. As such, support for the mayor at the last election does not seem to have a greater bearing on voters' evaluation of unemployment services in the treatment municipalities.

A related concern is that the increase in electoral responsibility is not simply the result of a change in political responsibility but the result of a recent change in political responsibility. That is, changes in the distribution of political responsibility might have short lived priming effects that moderate voters' attention to

the policies for which responsibility is changed in the immediate aftermath of this change. The implication is that there was no permanent shift in how accountable the mayor was held for unemployment services and, accordingly, no lasting relationship between responsibility and electoral accountability. This alternative explanation is hard to test in the present context, as we cannot look at the long-term differences between the reformed (i.e., treated) and unreformed (i.e., control) municipalities. In the election following the one examined above, all municipalities had implemented the reform. Even so, at least we know that if the results obtained above are the consequence of priming, then this priming cannot have been very short-lived. After all, the election examined above occurred a few years after the reform was first implemented in the treatment municipalities, and while this is a short period of time compared to the rate at which other political institutions change (cf. the section on adaptiveness), it is not a short period of time compared to the news cycle of typical elections (Rosenberg and Feldman 2008).

In conclusion, it is not obvious that the side effects of potential increases in political attention can explain the local electorates' behavior in the wake of the labor market reform. It is important to note that priming, or some other mechanism related to increased political attention, might be what is driving electorates to hold local incumbents more electorally accountable in the treatment municipalities. However, since the existing literature has not really explored the psychological mechanisms underlying the relationship between political responsibility and electoral accountability, this article remains agnostic about the exact mechanisms involved. This study is simply interested in establishing that voters' response to the reform can be characterized as causal, adaptive and policy-specific—not in finding out why they responded in this way.

Conclusion

This article examined a reform of Danish labor market regulation in which some municipal mayors were assigned more political responsibility for a specific policy. The article found that in the election following this reform, voters held these mayors more electorally accountable for the outcomes of this policy. This is especially noteworthy because the increase in political responsibility was exogenous, because the increase took place just a few years prior to the election, and because the article also found that voters *did not* hold their mayor more electorally accountable for the outcomes of policies unaffected by the reform. As such, the findings suggest that when an incumbent's political responsibility for a specific policy changes, this has a practically immediate causal effect on how electorally accountable voters hold this incumbent for the outcomes of this specific policy.

How generalizable is this characterization of the relationship between centralization of political responsibility and electoral accountability? The increase in centralization examined in this study was abrupt and exogenous—which makes the case inherently special. Nonetheless, there are reasons to believe that the findings do generalize. Several features of the elections studied here are relatively common. For instance, this study focused on a multi-party system with proportional

representation. Most countries have multi-party systems with proportional representation. Some might argue that Denmark is a likely case for identifying a relationship between political responsibility and electoral accountability because of the relatively high levels of political knowledge and interest (Hansen and Pedersen 2014). However, this potential threat to generalizability is arguably ameliorated by focusing on local elections, which are generally followed less closely (Elklit and Kjær 2013). That is, a Danish voter in a local election might be more similar to an average voter in a national election when it comes to political engagement. Even so, more studies in other contexts are needed to pin down the exact scope of the article's findings.

Turning to implications, a more causal, adaptive and policy-specific relationship between centralization of political responsibility and electoral accountability should mean that voters tend to elect more competent politicians. As such, if voters are able to hold politicians more electorally accountable for the policy outcomes that the politicians are more responsible for, then it will also be more likely that voters will select politicians based on the quality of outcomes that accurately reflect these politicians' efforts and abilities (Duch and Stevenson 2008; Achen and Bartels 2016; Stimson 2015; Erikson et al. 2002). Even so, a causal, adaptive and policy-specific relationship between centralization of political responsibility and electoral accountability is no panacea for effective democratic control. For one, as Ashworth (2012) and others have argued, voters may end up holding politicians accountable for policy outcomes that voters think are relatively unimportant but that incumbent politicians are clearly responsible for (e.g., Ashworth et al. 2017). Nonetheless, the results suggest that to the extent that it is possible to sort competent politicians from incompetent ones, voters will try to do so by identifying how responsible politicians are for the state of specific policy outcomes and electorally punish or reward them accordingly.

Acknowledgements For comments on previous drafts, I would like to thank Marc Andre Bodet, Thad Dunning, Michael Lewis-Beck, Gabriel Lenz, Peter Thisted Dinesen, Peter Bjerre Mortensen, Cecilia Mo, Kasper Møller Hansen, Asmus Leth Olsen, Richard Nadeau, Michael Sances, Søren Serritzlew and Rune Stubager. For their work in collecting the Danish Municipal Election surveys, I would like to thank Ulrik Kjær, Christian-Elmelund Præstekær and Jørgen Elklit. Replication files for this paper are available in the Political Behavior Dataverse (https://dataverse.harvard.edu/dataverse/polbehavior).

References

Acemoglu, D. (2005). Constitutions, politics, and economics: A review essay on Persson and Tabellini's the economic effects of constitutions. *Journal of Economic Literature*, *43*(4), 1025–1048.

Achen, C. H., & Bartels, L. M. (2016). *Democracy for realists: Why elections do not produce responsive government*. Princeton: Princeton University Press.

Act 483. 2009. *Lov om ændring af lov om ansvaret for og styringen af den aktive beskæftigelsesindsats, lov om en aktiv beskæftigelsesindsats og forskellige andre love samt om ophævelse af lov om supplerende aktiveringstilbud til visse ledige medlemmer af en arbejdsløshedskasse*.

Aghion, P., Alesina, A., & Trebbi, F. (2004). Endogenous Political Institutions*. *The Quarterly Journal of Economics*, *119*(2), 565–611.

Anderson, C. D. (2006). Economic voting and multilevel governance: A comparative individual-level analysis. *American Journal of Political Science*, *50*(2), 449–463.

Anderson, C. D. (2009). Institutional change, economic conditions and confidence in government: Evidence from Belgium. *Acta Politica*, *44*(1), 28–49.
Anderson, C. J. (2000). Economic voting and political context: A comparative perspective. *Electoral Studies*, *19*(2), 151–170.
Arceneaux, K. (2006). The federal face of voting: Are elected officials held accountable for the functions relevant to their office? *Political Psychology*, *27*(5), 731–754.
Ashworth, S. (2012). Electoral accountability: Recent theoretical and empirical work. *Annual Review of Political Science*, *15*, 183–201.
Ashworth, S., Bueno de Mesquita, E., & Friedenberg, A. (2017). Accountability and information in elections. *American Economic Journal*, *9*(2), 95–138.
Berry, C. R., & Howell, W. G. (2007). Accountability and local elections: Rethinking retrospective voting. *The Journal of Politics*, *69*(3), 844–858.
Besley, T., & Case, A. (2000). Unnatural experiments? Estimating the incidence of endogenous policies. *The Economic Journal*, *110*(467), 672–694.
Bhatti, Y., & Hansen, K. M. (2011). Who marries' whom? The influence of societal connectedness, economic and political homogeneity, and population size on jurisdictional consolidations. *European Journal of Political Research*, *50*(2), 212–238.
Bisgaard, M. (2015). Bias will find a way: Economic perceptions, attributions of blame, and Partisan-motivated reasoning during crisis. *The Journal of Politics*, *77*(3), 849–860.
Blom-Hansen, J., Houlberg, K., & Serritzlew, S. (2014). Size, democracy, and the economic costs of running the political system. *American Journal of Political Science*, *58*(4), 790–803.
Cameron, D. R. (1978). The expansion of the public economy: A comparative analysis. *American Political Science Review*, *72*(04), 1243–1261.
Carlin, R. E., & Singh, S. P. (2015). Executive power and economic accountability. *The Journal of Politics*, *77*(4), 1031–1044.
Cutler, F. (2008). Whodunnit? Voters and responsibility in Canadian federalism. *Canadian Journal of Political Science*, *41*(03), 627–654.
de Vries, C. E., Edwards, E. E., & Tillman, E. R. (2010). Clarity of responsibility beyond the pocketbook: How political institutions condition EU issue voting. *Comparative Political Studies*, *44*(3), 339–363.
de Vries, C. E., Van der Brug, W., Van Egmond, M. H., & Van der Eijk, C. (2011). Individual and contextual variation in EU issue voting: The role of political information. *Electoral Studies*, *30*(1), 16–28.
Downs, A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy*, *65*(2), 135–150.
Duch, R. M., & Stevenson, R. T. (2008). *The economic vote: How political and economic institutions condition election results*., Political economy of institutions and decisions New York: Cambridge University Press.
Dunning, T. (2012). *Natural experiments in the social sciences: A design-based approach*. Cambridge: Cambridge University Press.
Ebeid, M., & Rodden, J. (2006). Economic geography and economic voting: Evidence from the US states. *British Journal of Political Science*, *36*(03), 527–547.
Elklit, J., & Kjær, U. (2013). *KV09: Analyser af kommunalvalget 2009*. Denmark: Syddansk Universitetsforlag.
Erikson, R. S., MacKuen, M. B., & Stimson, J. A. (2002). *The macro polity*. Cambridge: Cambridge University Press.
Eskelinen, L. (2008). *Jobcentrenes ledelse og organisering*. Copenhagen: AKF Forlag.
Fernández-Albertos, J. (2006). Does internationalisation blur responsibility? Economic voting and economic openness in 15 European countries. *West European Politics*, *29*(1), 28–46.
Fossati, D. (2014). Economic vulnerability and economic voting in 14 OECD countries. *European Journal of Political Research*, *53*(1), 116–135.
Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*. New York: WW Norton.
Gomez, B. T., & Wilson, J. M. (2001). Political sophistication and economic voting in the American electorate: A theory of heterogeneous attribution. *American Journal of Political Science*, *45*(4), 899–914.

Hanmer, M. J., & Kalkan, K. O. (2013). Behind the curve: Clarifying the best approach to calculating predicted probabilities and marginal effects from limited dependent variable models. *American Journal of Political Science*, *57*(1), 263–277.
Hansen, K. M., & Pedersen, R. T. (2014). Campaigns matter: How voters become knowledgeable and efficacious during election campaigns. *Political Communication*, *31*(2), 303–324.
Harding, R. (2015). Attribution and accountability: Voting for roads in Ghana. *World Politics*, *67*(4), 656–689.
Hart, A. (2016). *Economic voting*. Cambridge: Cambridge University Press.
Healy, A., & Malhotra, N. (2013). Retrospective voting reconsidered. *Annual Review of Political Science*, *16*, 285–306.
Healy, A., Malhotra, N., et al. (2010). Random events, economic losses, and retrospective voting: Implications for democratic competence. *Quarterly Journal of Political Science*, *5*(2), 193–208.
Hellwig, T., & Samuels, D. (2007). Voting in open economies the electoral consequences of globalization. *Comparative Political Studies*, *40*(3), 283–306.
Hellwig, T. T. (2001). Interdependence, government constraints, and economic voting. *The Journal of Politics*, *63*(04), 1141–1162.
Hobolt, S., Tilley, J., & Banducci, S. (2013). Clarity of responsibility: How government cohesion conditions performance voting. *European Journal of Political Research*, *52*(2), 164–187.
Houlberg, K., & Pedersen, L. H. (2015). Political consensus and fiscal outcomes. *Local Government Studies*, *41*(1), 78–99.
Jakobsen, M. L., & Mortensen, P. B. (2014). *Regelstaten: Væksten i danske love og bekendtgørelser 1989–2011*. Denmark: Djøf/Jurist-og Økonomforbundet.
Jennings, W. (2009). The public thermostat, political responsiveness and error-correction: Border control and asylum in Britain, 1994–2007. *British Journal of Political Science*, *39*(4), 847–870.
Johns, R. (2011). Credit where it's due? Valence politics, attributions of responsibility, and multi-level elections. *Political Behavior*, *33*(1), 53–77.
Jones, B. D., & Baumgartner, F. R. (2005). *The politics of attention: How government prioritizes problems*. Chicago: University of Chicago Press.
Kayser, M. A., & Wlezien, C. (2011). Performance pressure: Patterns of partisanship and the economic vote. *European Journal of Political Research*, *50*(3), 365–394.
King, G., Keohane, R. O., & Verba, S. (1994). *Designing social inquiry: Scientific inference in qualitative research*. Princeton: Princeton University Press.
Kristensen, K. M. (2008). *Jobcenterreformen: Som et lyn fra en klar himmel*. Aarhus C: Aarhus Universitet.
Krosnick, J. A., & Kinder, D. R. (1990). Altering the foundations of support for the president through priming. *American Political Science Review*, *84*(02), 497–512.
Kuklinski, J. H., Quirk, P. J., et al. (2000). Reconsidering the rational public: Cognition, heuristics, and mass opinion. In A. Lupia, M. D. McCubbins, & S. L. Popkin (Eds.), *Elements of reason: Cognition, choice, and the bounds of rationality* (pp. 153–182). Cambridge: Cambridge University Press.
Larsen, M. V. (2016). Economic conditions affect support for prime minister parties in Scandinavia. *Scandinavian Political Studies*, *39*(3), 226–241.
Lassen, D. D., & Serritzlew, S. (2011). Jurisdiction size and local democracy: Evidence on internal political efficacy from large-scale municipal reform. *American Political Science Review*, *105*(2), 238–258.
León, S. (2011). Who is responsible for what? Clarity of responsibilities in multilevel states: The case of Spain. *European Journal of Political Research*, *50*(1), 80–109.
Meyer, B. D. (1995). Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics*, *13*(2), 151–161.
Ministry of Employment. (2006). *Press release: 14 pilotcentre er udpeget*. Ministry of Employment.
Ministry of Employment. (2010). *Aftale om et enstrenget kommunalt beskæftigelsessystem*. Ministry of Employment.
Nadeau, R., Niemi, R. G., & Yoshinaka, A. (2002). A cross-national analysis of economic voting: Taking account of the political context across time and nations. *Electoral Studies*, *21*(3), 403–423.
OECD. (2014). *Public expenditure and participant stocks on LMP*. *http://stats.oecd.org/* .
Order 1400. (2006). *Bekendtgørelse om delegation af statens beskæftigelsesindsats til visse kommuner*.

Pacek, A. C., & Radcliff, B. (1995). Economic voting and the welfare state: A cross-national analysis. *The Journal of Politics*, *57*(1), 44–61.
Parker-Stephen, E. (2013). Clarity of responsibility and economic evaluations. *Electoral Studies*, *32*(3), 506–511.
Powell, G. B., & Whitten, G. D. (1993). A cross-national analysis of economic voting: Taking account of the political context. *American Journal of Political Science*, *37*, 391–414.
Przeworski, A. (2004). Institutions matter? *Government and Opposition*, *39*(4), 527–540.
Ritzau. (2006). *Kommuner hilser jobcentre velkommen*. Ritzaus bureau.
Rodrik, D. 1996. *Why do more open economies have bigger governments?* National Bureau of Economic Research.
Rosenberg, H., & Feldman, C. S. (2008). *No time to think: The menace of media speed and the 24-hour news cycle*. New York: A&C Black.
Ruder, A. I., et al. (2014). Institutional design and the attribution of presidential control: Insulating the president from blame. *Quarterly Journal of Political Science*, *9*(3), 301–335.
Rudolph, T. J. (2006). Triangulating political responsibility: The motivated formation of responsibility judgments. *Political Psychology*, *27*(1), 99–122.
Sances, M. W. (2017). Attribution errors in federalist systems: When voters punish the president for local tax increases. *The Journal of Politics*, *79*, 1286–1301.
Singer, M. M. (2013). What goes around comes around: Perceived vulnerable employment and economic voting in developing countries. *European Journal of Political Research*, *52*(2), 143–163.
Soroka, S. N., & Wlezien, C. (2005). Opinion-policy dynamics: Public preferences and public expenditure in the United Kingdom. *British Journal of Political Science*, *35*(4), 665–689.
Stevenson, R. T., & Duch, R. (2013). The meaning and use of subjective perceptions in studies of economic voting. *Electoral Studies*, *32*(2), 305–320.
Stimson, J. A. (2015). *Tides of consent: How public opinion shapes American politics*. Cambridge: Cambridge University Press.
Tavits, M. (2007). Clarity of responsibility and corruption. *American Journal of Political Science*, *51*(1), 218–229.
Tilley, J., & Hobolt, S. B. (2011). Is the government to blame? An experimental test of how partisanship shapes perceptions of performance and responsibility. *The Journal of Politics*, *73*(02), 316–330.
Torfing, J. (1999). Workfare with welfare: Recent reforms of the Danish welfare state. *Journal of European Social Policy*, *9*(1), 5–28.
Vries, C. E., & Giger, N. (2014). Holding governments accountable? Individual heterogeneity in performance voting. *European Journal of Political Research*, *53*(2), 345–362.
Whitten, G. D., & Palmer, H. D. (1999). Cross-national analyses of economic voting. *Electoral Studies*, *18*(1), 49–67.
Wlezien, C. (1995). The public as thermostat: Dynamics of preferences for spending. *American Journal of Political Science*, *39*(4), 981–1000.

Supplementary Materials (Online Appendix)

S1: The Broader Legislative Context of the Reform

Above I argued that the reform of labor market regulation changed how politically responsible municipalities were for unemployment services. A concern one might have with this argument, is that the national government somehow undid the effects of the reform by introducing detailed legislation instructing municipalities on how they should administer unemployment services, leaving the municipalities with no real administrative discretion. If this is the case, then implementing the reform would simply have meant trading a clear limit to the municipalities' political responsibility for an opaque limit. However, if one studies the reform legislation (Order 1400, 2006), there is no sign of any such detailed regulation instructing municipalities on how to administer unemployment services. Further, if one explores the amount of enacted national legislation related to labor market regulation around the implementation of the reform, one does not find any marked increase. On the contrary, an examination of the legislation coming from the Ministry of Employment between 2005 and 2011 reveals that, while additional statutes and laws were being instated, fewer were instated in this period than between 1998 and 2004 and between 1991 and 1997 (to examine this, I used data from Jakobsen and Mortensen, 2014). As such, I find no evidence suggesting that the national government tried to take back some or all of the political responsibility for unemployment services delegated to the municipalities as part of the labor market reform.

S2: Variable Descriptions

Table S.1 presents a short description of the different survey items used in the analysis.

Table S.1: Description of survey items from the municipal election surveys

Variable name	Question	Coding
Reelect mayor	'Who did you vote for in the municipal election?'	1 is for mayoral party voters, 0 is for the other party's voters.
Reelect regional government	'Who did you vote for in the regional election?'	1 is for regional government party voters, 0 is for the other party's voters.
Reelect national government	'Who would you vote for if a national election was held tomorrow?'	1 is for national government party voters, 0 is for the other party's voters.
Treatment	'What municipality do you live in?'	1 indicates 14 treatment municipalities, 0 the 84 control municipalities.
Unemployment performance	'How satisfied or unsatisfied are you in general with your municipality's efforts towards the unemployed?'	Five point scale going from 0 "Very unsatisfied" to 1 "Very satisfied".
Elderly performance	'How satisfied or unsatisfied are you in general with your municipality's efforts towards the elderly?'	Five point scale going from 0 "Very unsatisfied" to 1 "Very satisfied".
Housing performance	'How satisfied or unsatisfied are you in general with your municipality's efforts towards private and public housing?'	Five point scale going from 0 "Very unsatisfied" to 1 "Very satisfied".
Administration controls municipality	'In reality, the administration controls the municipality, not the politicians'	Five point scale going from 0 "Completely disagree" to 1 "Completely agree".

Partisanship	'Who would you vote for if a national election was held tomorrow?'	1 if respondent voted for mayoral party, 0 otherwise.
Mayoral party in government	Indicator variable of whether the mayor's party is in government	1 for mayors from the two governing parties in 2009, 0 for other mayors.
Ideology	Measures congruence between respondent's ideology (left or right-wing) and the ideology of the mayor. Mayoral ideology determined based on party (Conservative and Liberal party as right wing), respondent's ideology based on question about self-reported ideology.	Coded 1 if respondent shares ideology with mayor, coded 0 if respondent does not.
News consumption - local	'Thinking back, how important was local media as a source of knowledge about the municipal election campaign?'	Four point scale going from 1 "Not at all" to 5 "Very important".
Age	'How old are you?'	Measured in years.
Employment status	'Where are you currently employed?'	11 different categories including student, unemployed and retiree.
Knowledge about municipal powers	Five different questions about who has responsibility for various policy areas.	Proportion of correct answers.
Interest	'How interested would you say you are in politics?'	Four point scale going from 0 "Not at all" to 1 "Very".
Informed	'How informed would you say you are about municipal politics in your own municipality?'	Five point scale going from 0 "Not at all informed" to 1 "Very informed".

Influence	'The mayor has a great deal of influence on how the municipality develops'	Five point scale going from 0 "Completely disagree" to 1 "Completely agree".
Responsible	'Who do you think has the main responsibility for things going as they have in the past four years in your municipality?'	Respondents answering "Local politicians" or "The Mayor" coded 1. Respondents answering "National politicians" coded 0.
Apathy	'I cannot be bothered with the municipal election'	Five point scale going from 0 "Completely disagree" to 1 "Completely agree".
Obligated	'I feel obligated to vote at the municipal election'	Five point scale going from 0 "Completely disagree" to 1 "Completely agree".
Satisfied with democracy	'How satisfied are you with the local democracy?'	Four point scale going from 0 "Not at all satisfied" to 1 "Very satisfied'
Pivotality	'How likely is it that your vote will be pivotal?'	Five point scale going from 0 "Basically zero" to 1 "Very probable'

S3: Descriptive statistics

Tables S.2, S.3 and S.4 present descriptive statistics on the survey items used in the analysis of the 2005, 2009 and 2013 municipal election surveys.

Figure S.1 presents the distribution of the key unemployment performance variable across treatment and control.

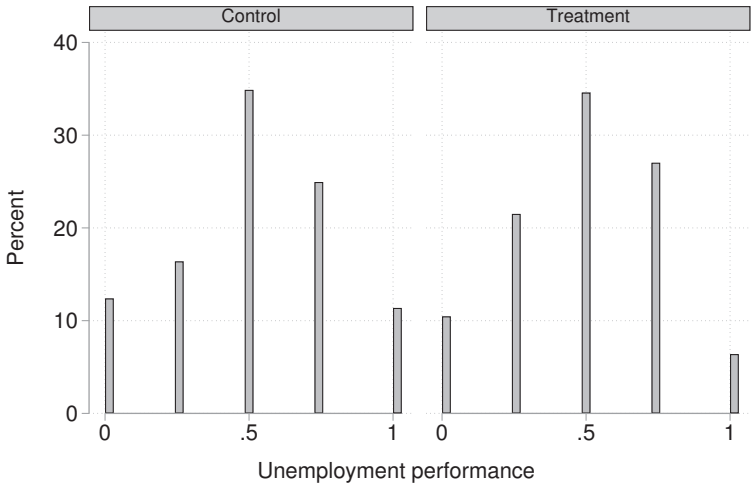

Figure S.1: Distribution of variable unemployment performance.

Table S.2: Descriptive statistics 2005

	Mean	SD	Min	Median	Max	n
Informed	0.56	0.23	0.00	0.50	1.00	1996
Interested	0.64	0.26	0.00	0.67	1.00	2009
Unemployment performance	0.59	0.31	0.00	0.75	1.00	1454
Knowledge about municipal powers	0.71	0.28	0.00	0.80	1.00	2011
Elderly performance	0.66	0.31	0.00	0.75	1.00	1534
Housing performance	0.75	0.26	0.00	0.75	1.00	1944
Ideology	0.70	0.46	0.00	1.00	1.00	2011
Apathy	0.15	0.27	0.00	0.00	1.00	1988
Obligation	0.96	0.16	0.00	1.00	1.00	2000
Satisfaction with municipal democracy	0.69	0.23	0.00	0.67	1.00	1975
Pivotality	0.45	0.28	0.00	0.50	1.00	1875

Table S.3: Descriptive statistics 2009

	Mean	SD	Min	Median	Max	n
Vote for mayoral party	0.31	0.46	0.00	0.00	1.00	2742
Vote for mayoral party at national elections	0.23	0.42	0.00	0.00	1.00	3199
Voted for mayoral party at regional election	0.25	0.43	0.00	0.00	1.00	3199
Voted for mayoral party at last municipal election	0.36	0.48	0.00	0.00	1.00	2642
Voted for mayoral party at last national election	0.26	0.44	0.00	0.00	1.00	3199
Influence	0.74	0.25	0.00	0.75	1.00	3175
Responsibility	0.67	0.47	0.00	1.00	1.00	2998
Unemployment performance	0.51	0.29	0.00	0.50	1.00	2296
treatment	0.15	0.36	0.00	0.00	1.00	3336
Housing performance	0.71	0.23	0.25	0.75	1.00	2920
Elderly performance	0.57	0.31	0.00	0.50	1.00	2771
Administration controls municipality	0.50	0.31	0.00	0.50	1.00	2895
Local media consumption	3.14	0.92	1.00	3.00	5.00	3336
National media consumption	2.33	1.06	1.00	2.00	5.00	3336
Age	54.53	13.79	18.00	55.50	91.00	3272
Shares ideology with mayoral party	0.44	0.50	0.00	0.00	1.00	3336
Mayor is from the same party as national government	0.49	0.50	0.00	0.00	1.00	3199
Elementary school	0.18	0.38	0.00	0.00	1.00	3336
High school	0.07	0.26	0.00	0.00	1.00	3336
Vocational high school	0.04	0.19	0.00	0.00	1.00	3336
Vocational school	0.25	0.44	0.00	0.00	1.00	3336
Shorter tertiary education	0.08	0.27	0.00	0.00	1.00	3336
Tertiary education	0.23	0.42	0.00	0.00	1.00	3336
Graduate degree	0.03	0.17	0.00	0.00	1.00	3336
Postgraduate degree	0.12	0.33	0.00	0.00	1.00	3336
Untrained worker	0.05	0.21	0.00	0.00	1.00	3336
Skilled worker	0.11	0.31	0.00	0.00	1.00	3336
Blue collar worker	0.13	0.34	0.00	0.00	1.00	3336
White collar worker	0.20	0.40	0.00	0.00	1.00	3336
Self employed	0.06	0.24	0.00	0.00	1.00	3336
Home maker	0.00	0.05	0.00	0.00	1.00	3336
Student	0.03	0.16	0.00	0.00	1.00	3336
Not looking for work	0.00	0.06	0.00	0.00	1.00	3336
Unemployed	0.04	0.19	0.00	0.00	1.00	3336
Retiree	0.34	0.47	0.00	0.00	1.00	3336
Will not say	0.01	0.10	0.00	0.00	1.00	3336
Other	0.03	0.18	0.00	0.00	1.00	3336

Table S.4: Descriptive statistics 2013

	Mean	SD	Min	Median	Max	n
Responsibility	0.78	0.41	0.00	1.00	1.00	3968
Influence	0.72	0.22	0.00	0.75	1.00	4254

S4: A Manipulation check

Mean responses for the two manipulation check questions are presented in the two left-most columns of Table S.5. It is important to note that while the differences between treatment and control are not very large, these questions are about conditions in the municipality in general, not just unemployment services. While I would expect that voters in the treatment municipalities believe the mayor is substantially more responsible for unemployment services, I would only expect that voters believe the mayor is slightly more responsible for the overall conditions in the municipality.

Table S.5: Manipulation check

	2009		2013	
	Responsible	Influence	Responsible	Influence
Control	0.66	0.74	0.78	0.72
	(0.01)	(0.01)	(0.01)	(0.00)
Treatment	0.71	0.76	0.77	0.72
	(0.02)	(0.02)	(0.01)	(0.01)
p-value	0.03	0.04	0.29	0.34
Observations	2998	3175	3968	4254

Standard errors in parentheses, one-sided p-value from difference in means test.

Was the 2009 difference due to pre-treatment differences in voters' beliefs? This seems unlikely given the extensive balance test presented in Table 1, but we cannot be sure because these questions were not posed in the 2005 survey. However, these questions were part of the 2013 municipal election survey and we can utilize the 2013 data to conduct a post-treatment balance test. Recall that, when the 2013 election came about, the reform was implemented in all municipalities. As such, if the differences in the 2009 survey were due to the asymmetry in political responsibility caused by the reform, these differences should have disappeared in 2013. The two right-most columns of table S.5 report means across the treatment and control municipalities from the 2013 survey. As expected, once all of the municipalities had implemented the reform, there was no longer any difference in the mean responses to the two manipulation check questions.

S5: Alternative estimation methods

In the analyses conducted below, I show that the key findings presented in Table 2 are robust to employing two alternative estimation methods. These methods relax some of the assumptions made in order to estimate the models in Table 2, and accordingly, they provide a more complete picture of the statistical evidence for the key conjecture of the analysis: that voters in the treatment municipalities held their mayoral party more electorally accountable for unemployment services than voters in the control municipalities.

The models estimated above did not take the hierarchical structure of the data – individual voters nested within municipalities – fully into account. In order to do this, I estimate a set of mixed effects multilevel logit models with the same configuration of variables used in Table 2. Estimates from these models are presented in Table S.6. The important estimates remain practically unchanged, although the standard error of the estimates increase slightly. Most importantly, the difference in AMEs remains statistically significant in three out of four models ($p \approx 0.05$). The logit interaction coefficients also remain statistically significant, although only at the ten percent level.

The tests used to asses the statistical significance of the interaction terms and differences in AMEs in Table 2 rely on a number of parametric assumptions. To get around these assumptions, I tried to derive the statistical significance using a form of randomization inference; a non-parametric method (cf. Gerber and Green, 2012). In particular, I used the following procedure:

1. Draw a random sample of 14 municipalities, and create a dummy which was equal to one if the respondent lived in one of these randomly drawn municipalities.

2. Estimate the models reported in column 1-4 of Table 2, but substituting the actual treatment variable for the dummy variable created in (1).

3. Store the estimated interaction effect between the simulated treatment dummy and unemployment performance obtained for each logit model estimated in (2).

4. Derive the the average marginal effect (AME) of unemployment performance in the simulated treatment and control municipalities for each of the models estimated in (2) and store the difference in AMEs.

Table S.6: Multi-level logistic regression of probability of voting for the mayoral party

	(1)	(2)	(3)	(4)
Unemployment performance	0.84*	0.84*	0.92*	0.88*
	(0.24)	(0.25)	(0.37)	(0.36)
Treatment	-0.60	-0.58	-0.49	-0.64
	(0.41)	(0.44)	(0.45)	(0.43)
Treatment * Unemployment performance	1.02+	1.02+	1.28+	1.36*
	(0.55)	(0.59)	(0.66)	(0.61)
Administration controls municipality	-0.31	-0.42+	-0.44+	-0.41+
	(0.21)	(0.21)	(0.24)	(0.24)
Housing performance	0.54+	0.53+	-0.01	0.07
	(0.30)	(0.32)	(0.40)	(0.40)
Elderly performance	0.83*	0.92*	1.06*	1.05*
	(0.23)	(0.25)	(0.31)	(0.30)
AME (Control)	0.17	0.17	0.10	0.09
	(0.05)	(0.05)	(0.04)	(0.04)
AME (Treatment)	0.36	0.35	0.23	0.24
	(0.08)	(0.09)	(0.06)	(0.05)
Difference (T-C)	0.18	0.18	0.14	0.15
p-value of difference	0.05	0.07	0.05	0.02
Sociodemographic controls		✓	✓	✓
Political controls			✓	✓
Municipal level variables				✓
Log likelihood	-911.88	-881.73	-553.65	-550.06
Observations	1522	1500	1500	1500

Robust standard errors clustered on municipality in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$

5. Repeat (1)-(4) 10,000 times resulting in 10,000 unique interaction coefficients and AME-differences for each model.

6. Calculate *p-values* for each model by looking at the proportion of simulated logit coefficients and AME estimates which are *larger* than the ones estimated for the actual treatment and control municipalities.

A random sample of the 10,000 simulations is plotted in Figure S.2 along with the calculated p-values. These p-values signify how likely it is to get an interaction or difference in AMEs of the size estimated in Table 2 or larger *if* there was no effect of being assigned to implement the labor market reform for any of the municipalities (a sharp null). The p-values do become slightly larger using this method, however, the p-values are still below 0.1 and thus reflect that the observed difference in the weight voters put on unemployment service between treatment and control municipalities is unlikely to have occurred by chance.

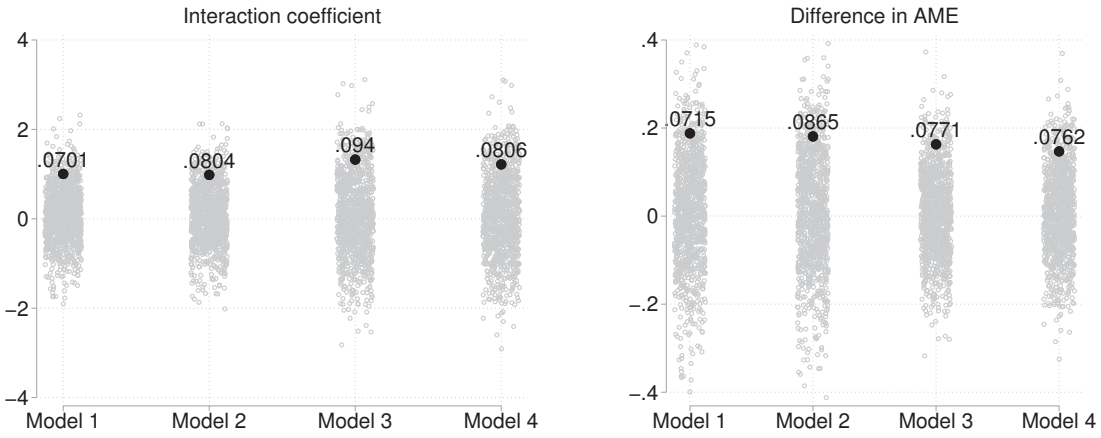

Figure S.2: A sample (n=1,000) of the simulated differences in AMEs and interaction effects from each of the four different logit models estimated in Table 2. These are computed using randomization inference (RI). The black dot signifies realized outcome, taken from Table 2, and the number attached to it is the RI p-value.

S6: Difference-in-difference

Conducting a difference-in-difference analysis is complicated by a few factors. Even though the key unemployment performance and vote intention questions were asked in both surveys, there is not a large overlap between the datasets when it comes to the control variables used in Table 2. As such, I cannot estimate a model with as large a number of controls, however, this problem is somewhat offset by the difference-in-difference approach's ability to control for any pre-treatment differences between treatment and control municipalities. A more serious challenge to including the 2005 data relates to the fact that some municipalities were in the process of being amalgamated due to the large reform which was implemented in 2006 (cf. Figure 1). As a result, almost half of the respondents voted in an amalgamated municipality, which was different from the one where their incumbent mayor had been elected, blurring patterns of accountability. I deal with this problem by by defining the dependent variable in '05 as voting for the party which had the mayoralty in the voter's existing (old) municipality. Even so, these amalgamations impede the strength of the analysis.

In Figure S.3, I show the AMEs of unemployment performance on support for the mayoral party in treatment and control municipalities in both 2009 and 2005. The AMEs are derived

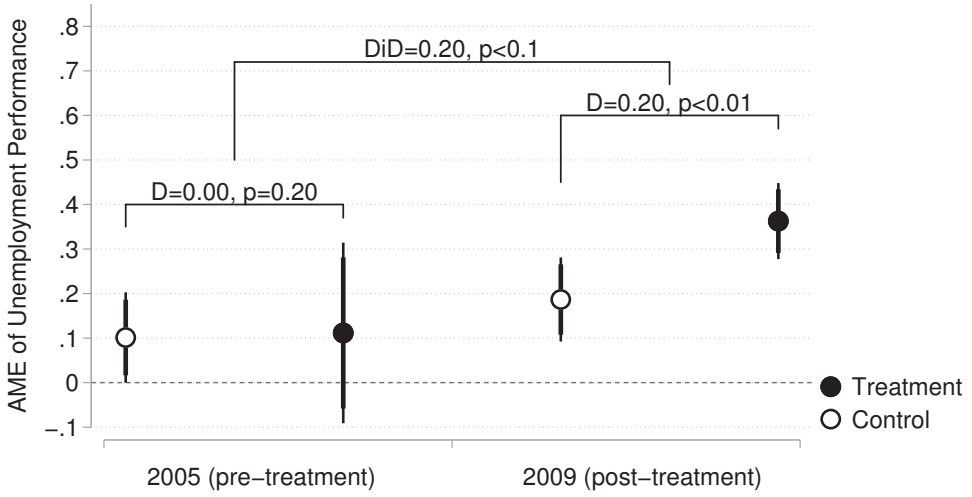

Figure S.3: Average Marginal Effects of unemployment performance on probability of voting for the mayoral party across treatment status and time period. Derived from logistic regression model described in the text; McFadden $R^2 = 0.032$, $n = 2,582$. Wald tests used to compare the different AMEs. The vertical lines are 90 pct. (thick) and 95 pct. (thin) confidence intervals.

from a logistic model estimated on a pooled dataset. This model sets voting for the mayoral party as a function of a three-way interaction between unemployment performance, treatment status and time period ('09 versus '05). The model also controls for housing and elderly care performance as well as for whether the mayor is of the same party as the national government. To take the different patterns of accountability across amalgamated and continuing municipalities into account (cf. above), I allow all performance variables to have different slopes depending on whether voters lived in a municipality which was amalgamated.

As can be seen from Figure S.3 there is no difference in the effect of voters evaluation of unemployment services on support for the mayoral party across treatment and control municipalities in 2005. Accordingly, before the reform of labor market regulation was implemented, there was no apparent difference in electoral accountability across treatment and control municipalities. In 2009, however, when the treatment municipalities had gotten more political responsibility for unemployment services, there is a statistically significant difference.The difference in difference estimate is only statistically significant at the ten percent level. The slight drop in statistical significance can be explained by the extra estimation error introduced by including the more noisy 2005 data.

S7: Analyzing additional policy areas

In this section, I examine differences in electoral accountability across the treatment and control municipalities for some additional policy areas. As such, I investigate whether voters in the treatment municipalities were more likely to electorally punish and reward the mayor for quality of services in nine different policy areas, which were not affected by the reform of labor market regulation (including the two examined in Figure 5). In particular, I use the logit model presented in column 4 of Table 2 as a template, swapping the unemployment performance variable for one of the alternative policy variables. I do this for all policy variables. For each of these nine new models, I then derive the AME of the policy variable on voters' propensity to vote for the mayoral party in both the treatment and in the control municipalities. Finally, I test the AME in the treatment municipalities against the AME in the control municipalities using a Wald test. The results of these analyses are reported in Table S.7.

Table S.7: Differences across treatment and control for other policy outcomes

Policy Area	Treatment	Control	Standard Error	p-value
Unemployment per	0.25	0.09	0.07	0.01
Housing	0.05	0.04	0.09	0.85
Daycare	0.08	0.08	0.10	0.94
Recreation	0.01	0.08	0.08	0.44
Schools	0.08	0.03	0.08	0.56
Library	0.04	0.09	0.07	0.47
Culture	0.19	0.09	0.09	0.28
Business	0.18	0.10	0.10	0.43
General services	0.21	0.08	0.09	0.16
Elderly Services	0.14	0.11	0.09	0.79
Health Services	0.17	0.08	0.09	0.29
Total	0.13	0.08	0.09	0.47

The models from which the average marginal effects are derived include the full set of controls.

As is revealed by looking at the right-most column of Table S.7, the AME of voters' assessment of the quality of the services provided in these nine different policy areas do not significantly differ across treatment and control municipalities.

S8: No Evidence of Increases in Partisan Motivated Reasoning

An alternative explanation for our findings is that there are voters in the municipalities who (dis)like the mayoral party, and when they find out that their mayor has become more responsible for unemployment services, they increase (or decrease) their estimate of service quality in this area accordingly. If this is the case, voters' satisfaction with unemployment services should be more strongly correlated with *past* support for the mayor in the municipalities where the mayor got more responsibility for unemployment services.

In order to examine this possibility, we re-estimate the four logistic regression models from Table 2 using self-reported support for the (current) mayoral party at the previous election as the dependent variable. (The models are thus only estimated using respondents who said that they could remember which party they voted for at the last election.) Table S.8 presents the results from these analyses. In these models, which predict past voting, the interaction effect between treatment status and unemployment performance is negligible and statistically insignificant. So is the difference in AMEs across treatment and control municipalities. From this, we can conclude that the increased correlation between satisfaction with unemployment services and support for the mayor is not be driven by voters who already supported the mayor at the last election becoming more satisfied with unemployment services, or by voters who did not support the mayor becoming less satisfied.

More broadly, these analyses show that there is no sign of increases in partisan motivated reasoning when it comes to how satisfied voters are with unemployment services in the treatment municipalities. This corroborates the initial conclusion that voters hold their mayor more electorally accountable for the quality of unemployment services.

Table S.8: Logistic regression of voting for the mayoral party at the *last* election

	(1)	(2)	(3)	(4)
Unemployment performance	0.53*	0.55*	0.45	0.43
	(0.21)	(0.23)	(0.32)	(0.32)
Treatment	-0.18	-0.20	-0.17	-0.14
	(0.28)	(0.28)	(0.34)	(0.34)
Treatment * Unemployment performance	0.02	0.05	0.00	-0.03
	(0.51)	(0.52)	(0.76)	(0.76)
Administration controls municipality	-0.18	-0.28	-0.09	-0.11
	(0.19)	(0.20)	(0.26)	(0.26)
Elderly performance	0.10	0.22	-0.03	-0.04
	(0.20)	(0.21)	(0.25)	(0.25)
Housing performance	0.86*	0.74*	0.84*	0.87*
	(0.28)	(0.28)	(0.35)	(0.36)
Sociodemographic controls		✓	✓	✓
Political controls			✓	✓
Municipal level variables				✓
AME (Control)	0.12	0.12	0.06	0.06
	(0.05)	(0.05)	(0.04)	(0.04)
AME (Treatment)	0.12	0.13	0.06	0.05
	(0.11)	(0.10)	(0.10)	(0.10)
Difference (T-C)	0.00	0.01	-0.00	-0.01
p-value of difference	1.00	0.96	0.98	0.95
Pseudo R^2	0.02	0.04	0.34	0.34
Log likelihood	-965	-937	-644	-642
Observations	1476	1461	1461	1461

Robust standard errors clustered on municipality in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$