

IZA DP No. 9497

**The Deterrent Effect of Voting Against Minarets:  
Identity Utility and Foreigners' Location Choice**

Michaela Slotwinski  
Alois Stutzer

November 2015

# The Deterrent Effect of Voting Against Minarets: Identity Utility and Foreigners' Location Choice

**Michaela Slotwinski**

*University of Basel*

**Alois Stutzer**

*University of Basel  
and IZA*

Discussion Paper No. 9497  
November 2015

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### **The Deterrent Effect of Voting Against Minarets: Identity Utility and Foreigners' Location Choice\***

This paper uses the vote on the Swiss minaret initiative as a natural experiment to identify the causal effect of negative attitudes towards immigrants on foreigners' location choices and thus indirectly on their utility. Based on a regression discontinuity design with unknown discontinuity points and administrative data on the population of foreigners, we find that the probability of their moving to a municipality that unexpectedly expressed strong reservations decreases initially by about 60 percent. The effect levels off over a period of about 5 months. Consistent with a reduction in the identity utility for immigrants in general, the reaction is not confined to Muslims, whereby high-skilled foreigners seem to be most sensitive to the newly revealed reservations.

JEL Classification: D83, J61, R23, Z13

Keywords: attitudes, foreigners, identity utility, location choice, RDD

Corresponding author:

Alois Stutzer  
Faculty of Business and Economics  
University of Basel  
Peter Merian-Weg 6  
4002 Basel  
Switzerland  
E-mail: [alois.stutzer@unibas.ch](mailto:alois.stutzer@unibas.ch)

---

\* Data on the foreign resident population were kindly provided by the Federal Statistical Office in cooperation with the Federal Office for Migration. We are further grateful to Jack Porter and Ping Yu for sharing the basic code for their method, which we then coded in Mata. Michaela Slotwinski acknowledges financial support from WWZ Forum and the nccr - on the move funded by the Swiss National Science Foundation.

# 1 Introduction

The question of whether people care how they are perceived by others is important for understanding individual decision-making as many choices involve aspects of social interaction. This obviously holds for bonding decisions, but it also applies to people's choice of organization where they wish to work or spend their leisure time. The issue also arises when people decide about the community they want to live in. To what extent are the resident population's attitudes relevant for such decisions? By analyzing how foreigners react to unexpectedly revealed reservations towards their group, we illuminate the role of identity utility in location choices.

Attitudes towards foreigners influence the interaction between immigrants and the native population. This matters for the successful realization of gains from trade between these two parties as well as for the migrants' integration (see, e.g., the review on cultural integration by Algan et al., 2012). Specifically, a host country's culture of welcome might affect where mobile foreigners are willing to locate, and is crucial for regions that need to attract experts from abroad. There is a large body of literature which discusses the sources and expression of negative attitudes towards foreigners as reflected, for example, in right-wing extremism (e.g., Hainmueller and Hopkins, 2014). However, foreigners' behavioral reactions to these attitudes and the consequences that they have for their welfare are less often addressed. While economic models and analyses of foreigners' location choices offer an approach to learn about foreigners' valuations of the overall attractiveness of an environment, this field of research typically neglects the political attitudes of the resident population. Over and above, the attempt to analyze the relationship between the presence of foreigners and residents' attitudes towards them poses a severe methodological challenge: an inherent simultaneity bias (Dustmann and Preston, 2001). On the one hand, the presence of foreigners may affect natives' preferences and attitudes. Getting to know foreigners may reduce prejudices, as postulated by the so-called contact hypothesis, or the presence of foreigners may aggravate negative attitudes, as natives fear pressure on the labor market and the welfare system, expect a rise in the crime rate or the alienation of their own culture. On the other hand, attitudes of natives towards foreigners may affect the presence of immigrants. Foreigners probably move less to areas where residents have reservations about them, where they fear discrimination or even physical abuse.

In this paper, we analyze the role of attitudes towards foreigners in their choice of residence based on a unique setting which allows us to address the identification problem of simultaneity. We integrate identity utility, as introduced by Akerlof and Kranton (2000), in a model of location choice to capture the idea that perceived negative attitudes towards foreigners affect their utility. The model predictions are empirically tested exploiting a quasi-experimental setting for Switzerland. In a national ballot, citizens voted on whether the further construction of minarets in Switzerland should be prohibited (the so-called

minaret initiative). Against the recommendation of the Federal Assembly and the predictions of leading opinion pollsters, the amendment to the Swiss constitution was accepted with a clear majority. The ballot outcome manifested reservations and anti-immigrant attitudes in regions of Switzerland which had previously been hidden. There were municipalities where voters unexpectedly deviated strongly from their past voting behavior on migration-related issues. We exploit this fact and study whether the inflow of foreigners to these particular municipalities declined in the aftermath of the vote, which would be consistent with lower expected utility and evidence that individuals react to perceived attitudes towards their social category.

The empirical analysis draws on administrative data for the universe of foreigners living in Switzerland which allow us to study their moving behavior before and after the national vote on the minaret initiative. Based on a simple simulation study, we expect the effect, if it exists, to materialize as a sharp decline in the probability that a foreigner would choose to move to a municipality that had unexpectedly revealed its reservations towards foreigners some time after the vote. Since the point in time at which the effect can be observed is not deterministically predictable, the change in moving patterns is analyzed based on a regression discontinuity design with unknown discontinuity points (Porter and Yu, 2015). We find such discontinuous jumps in the moving pattern of foreigners in several cantons. We observe that foreigners are deterred from locating to municipalities that unexpectedly revealed negative attitudes towards them. The estimated effect is sizable. The probability that a foreigner chooses one of these municipalities drops, on average, by about 8 percentage points relative to a pre-intervention level of thirteen percent, and thus by about 60 percent. Several validation checks suggest that the assumptions underlying the RDD are valid; neither the number of foreigners moving nor the municipalities that unexpectedly reveal their reservations about immigrants are systematically different before and after the threshold dates. The effect holds to a similar extent for groups of foreigners other than Muslims. This suggests that foreigners identify with the specific minorities who are restricted and that they perceived the support for the initiative as being a general expression of negative attitudes towards them. Still, the effect differs across specific groups of foreigners and over time. While it is substantial for first generation foreigners, there is no discernible effect for second generation foreigners. The latter are often well assimilated and seem to be less affected by the revealed attitudes. This latter finding emerges from testing a potential alternative explanation that landlords will dare to discriminate more against foreigners once they learn about the reservations of their fellow citizens. This supply-driven reaction, however, would affect foreigners in general. We further find that foreigners in high-skill occupations react more than foreigners in low-skill occupations. The effect fades after about 5 months. This might be due to the lower salience of the issue after the elapse of some time and/or due to an adjustment in the housing market.

In the latter case, lower rents compensate foreigners for the disutility of living in an environment with negative attitudes towards them.

Our analysis contributes to several strands of literature. First, there is a growing literature on the economics of identity which builds on the seminal work of Akerlof and Kranton (2000). In research on the integration of foreigners, an emphasis is on whether immigrants see themselves more or less as members of the host society. Respective studies explore how this aspect of foreigners' identity is related to economic outcomes (see, e.g., Dustmann, 1996; Casey and Dustmann, 2010; Georgiadis and Manning, 2013). In a lab experiment, Mechtel et al. (2014) study identity utility based on people's group choices. In our analysis, we derive evidence on identity utility from observed decision-making regarding the choice of residency.

Second, our study also contributes to the growing literature on the interdependent relationship between the presence of immigrants and attitudes towards them.

Regarding the effects of an inflow of foreigners on the attitudes of the native population, several studies take the potential reverse causality into account and try to solve the identification problem by instrumental variables (IV) strategies. In pioneering work, Dustmann and Preston (2001) analyze the effect of the local ethnic composition on natives' attitudes towards foreigners in the UK. They propose an IV approach by using the ethnic composition of broader areas as an instrument for the composition in local areas. In their study, the simultaneity bias matters and correcting for it reveals that a higher share in the (ethnic) minority population leads to more hostile attitudes in the majority population. The same strategy is, for example, applied in Dill (2013) to explore the effect of the presence of immigrants on right-wing extremism in Germany. In a study for Switzerland, Kuhn and Brunner (2014) make use of this approach to measure the effect of the foreign population in Swiss municipalities on voting behavior in national referendums on migration issues. They find that a larger fraction - in particular of culturally distant - foreigners is related to lower support for laws which propose more lenient rulings on migration practices. Another strand of studies uses an IV strategy initially suggested by Card (2001) to analyze the effect of immigration on the labor market. This strategy is based on the idea that immigrants tend to move to regions where immigrants of the same ethnicity are already present, and thus use the historical migration pattern as an instrument for current immigration (see, e.g., Barone et al., 2014; Méndez Martínez and Cutillas, 2014; Halla et al., 2014). Gerdes and Wadensjö (2008) study how an increase in the population's share of immigrants affects election outcomes using exogenous variation in the share of immigrants due to the government's placement practice for asylum seekers in Denmark. They find that the support for anti-immigrant parties increases and that the most pro-immigrant party loses votes.

These studies, having taken the potential endogeneity into account find, overall, that attitudes towards foreigners become more negative with a higher fraction of foreigners in the local population. They also indicate that endogeneity seems to play a role, thus suggesting that attitudes could be an influencing factors in immigrants' location choices.

To the best of our knowledge, there are five studies that are concerned with the question of whether foreigners' location choices react to attitudes towards them. An analysis by Tolnay and Beck (1992) documents the relationship between racial violence in U.S. counties and the outmigration of the threatened group between 1910 and 1930. In related work, Henry (2009) finds that African Americans are less likely to move to places where they face a higher risk of becoming a victim of a hate crime. While these two studies focus on actual violence, we concentrate on the effect of negative attitudes which are expressed in an otherwise tranquil environment. Waisman and Larsen (2008) use a quasi-random variation in the placement of asylum seekers in Sweden to investigate the welfare consequences of negative attitudes towards foreigners. They use attitude measures based on survey data and find some evidence that immigrants' location choice is influenced by natives' attitudes and that negative attitudes are associated with lower wages. They attribute their findings to labor market discrimination. Damm (2009) investigates regional factors of migrants' location choices in Denmark. She also uses a quasi-random placement of asylum seekers and finds that the hazard to leave the assigned municipality rises sharply with the share of votes for right-wing parties. She does not attribute this finding to negative attitudes per se, but to the welfare-seeking behavior of immigrants that is hampered by right-wing parties which seek to restrict welfare for immigrants. The most recent study is by Gorinas and Pytliková (2013) who investigate whether attitudes towards immigrants have an impact on migration flows on an international level. They perform a cross-national study and find that the surveyed dismissive attitudes of natives are negatively correlated with migrant inflows. Thereby migrants from OECD countries are the most adversely affected group in their study.

While these five studies indicate that negative attitudes potentially affect minorities' location choices, negative attitudes are either only inferred and/or their potential endogeneity is not accounted for.

In our work, we pursue a novel strategy and measure attitudes towards immigrants directly by ballot data from national referendums. These voiced attitudes are policy-relevant and are less likely to be biased by socially desirable answers than survey responses. Moreover, attitudes expressed in the context of a ballot can be clearly observed and interpreted by the public.

The remainder of the paper is structured as follows. Section 2 provides background information on the institutional setting as well as on the informational conditions defining the municipalities that unexpectedly reveal a shift towards an anti-immigration position.

Section 3 formalizes our idea how perceived attitudes influence an individuals' utility and how this affects his or her location choice. Based on these considerations, Section 4 outlines the empirical strategy. Our data source is described in Section 5, and the results are presented in Section 6. Section 7 concludes.

## 2 Background and Informational Setting

At the center of our study, there is information about individual voting behavior in a national referendum that is used to infer residents' revealed attitudes towards foreigners and the change of these publicly visible attitudes over time. In order to understand this specific informational setting, we first provide some background information about the specific proposition and direct democratic decision-making in Switzerland. In a second step, we explain the circumstances in voting behavior that are considered an unexpected revelation of reservations towards foreigners. Finally, we discuss how the ballot outcome entered foreigners' information base.

### 2.1 The vote on the minaret initiative

On November 29, 2009, a majority of Swiss voters approved an initiative that banned the further construction of minarets in Switzerland.<sup>1</sup> Swiss citizens are accustomed to expressing their opinions at the poll. Direct democratic decision-making is very common in Switzerland, even at the national level. Citizens openly discuss the proposals and can rely on being provided with ample coverage of the issues in the media. Newspapers, for example, publish the statements and voting recommendations of political parties and opinion leaders. Frey (1994) analyses the role of public discussion in the process of direct democratic decision-making, with an emphasis on the Swiss experience.

In the aftermath of 9/11 with its increased fear of terrorism, a discussion arose about whether practicing Muslims in Switzerland would threaten the democratic order. In this anxious atmosphere, two Swiss centre-right conservative parties, i.e. the Swiss People's Party and the Federal Democratic Union, started preparations for the minaret initiative against the further construction of minarets. The required number of signatures was collected by July 8, 2008, and the initiative was officially submitted on July 29, 2008.

---

<sup>1</sup>The Swiss direct democratic system, among other things, allows the electorate to propose amendments to the federal constitution by launching a so-called popular initiative. Initiators have to collect 100,000 valid signatures amounting to about two percent of the population of Swiss citizens. Once this hurdle is overcome, voters decide in a national referendum on whether the constitution should be changed. This is the case if the initiative is approved by the majority of voters overall as well as in a majority of cantons. If this holds, the federal government is obliged to implement the will of the majority. For an overview about the Swiss political system, see, for example, Linder (2010).



At this point in time, there were only four minarets in the whole of Switzerland. For their construction a permit was necessary, requiring that the construction plans comply with the cantonal and communal rules. Due to these existing restrictions, it was not to be expected that many minarets would be built in Switzerland. It was also clear that the proposed amendment would not change anything in the religious practice of Swiss Muslims. Mosques were already present and their right to exist was not affected by the initiative.

During the campaign and beyond, the initiative was widely discussed in the Swiss media and attracted global attention. While supporters argued that minarets have a political dimension, as they represent an Islamic religious and political symbol of power and that their presence might thus endanger Swiss democracy, opponents equated minarets with steeples and questioned whether their prohibition would be compatible with human rights and freedom of religion. The degree of public attention that this issue attracted can partly be ascribed to the campaign advertisement. The initiative's campaign poster showed a woman in a black niqab and the Swiss flag with looming minarets representing rockets (see Figure B.4 in the Appendix). The Federal Commission against Racism judged that this visual tactic would jeopardize public peace and therefore banned it from being displayed in some towns and cantons.<sup>2</sup> This dispute aggravated the tensions between supporters and opponents, as the faction backing the ban believed it to be necessary, whereas those against the ban saw it as censorship and a restriction of the freedom of speech. Overall, the campaign raised fundamental discussions about religious freedom, cultural diversity and tolerance in Switzerland and was very salient in the media. The discussion was not limited to Muslims in Switzerland, but also involved the issue of migration and the treatment of minorities in general.

Most parties and the Swiss Federal Assembly recommended that the initiative be rejected. However, contrary to general expectations as well as the forecasts of leading opinion researchers that the initiative would not pass, it was accepted with a clear majority of 57.5 percent of the votes. Specific minority rights were thus restricted in a constitutional amendment. The decision came into effect at the national level and no minarets have been built since then. In the exercise of their religion, however, nothing has changed for Muslims in Switzerland.

## 2.2 Attitudes towards foreigners across municipalities

The acceptance of the initiative did not only come as a surprise, but the voting pattern across municipalities did not match the rather stable political spectrum of views on migration manifested in past initiatives. There were municipalities that unexpectedly revealed

---

<sup>2</sup>The campaign poster was, for example, forbidden in Basel, Freiburg, Lausanne, Morges, Neuenburg, Nyon, and Yverdon.

anti-immigrant attitudes, which were not known to external observers until then. We thus see the event as an exogenous shock in *perceived* attitudes towards foreigners.

The vote on the minaret initiative provides a particularly attractive setting to capture information about citizens' attitudes. It can be conceptualized as a low-cost decision involving virtually no instrumental considerations about economic consequences, but rather allowing fully expressive voting behavior. Previous votes like the initiative "for the regulation of immigration" held on September 24, 2000, provoked serious concerns about negative economic repercussions in case of an approval. The voting behavior on the minaret initiative thus revealed rather undiluted signals about attitudes towards foreigners and allows us to measure potential changes in the publicly visible positioning of municipalities, i.e., changes relative to the best approximation as derived from previous voting behavior. Importantly, the conditions to get representative information about attitudes across space are the same for the people in our study as for us the researchers. We consider the changes in perceived attitudes as exogenous. The outcome is not anticipated and reveals new information. Information, which foreigners can incorporate in their perception about the political orientation of people across municipalities.

In order to use the exogenous shock to measure the effect of perceived attitudes on foreigners' location choices, we identify the municipalities that revealed previously unknown preferences based on a comparison with relevant votes in the past. In particular, we rely on four votes that occurred within a time frame close to the minaret initiative and which are sufficiently contextually related to it, i.e., in that they also involve the expression of attitudes towards foreigners. We code the voting results such that a higher vote share reflects more restrictive attitudes towards migration and foreigners. These votes include (i) the initiative "For the regulation of immigration" ("Initiative für eine Regelung der Zuwanderung") held on September 24, 2000 (ii) the initiative "Against the abuse of the asylum law" ("Initiative gegen Asylrechtsmissbrauch") held on November 24, 2002 (iii) the referendum on "The federal decree regarding the regular naturalization and the easier naturalization of young, and second-generation foreigners" ("Referendum zum Bundesbeschluss über die ordentliche Einbürgerung sowie über die erleichterte Einbürgerung junger Ausländerinnen und Ausländer der zweiten Generation") held on September 26, 2004 and (iv) the referendum on "The federal decree about the acquisition of citizenship rights by third-generation foreigners" ("Referendum zum Bundesbeschluss über die ordentliche Einbürgerung sowie über erleichterte Einbürgerung junger Ausländerinnen und Ausländer der dritten Generation") also held on September 26, 2004.

Specifically, a municipality is characterized as having unexpectedly revealed negative attitudes towards foreigners if the following criteria are fulfilled:

- The average support for restrictive migration policies in the past lies below the mean of the canton they belong to.

- The swing to a more restrictive position, as expressed in the vote on the minaret initiative, is larger than the mean change of all the municipalities within the canton.
- The support for the minaret initiative exceeds 50% of the municipality's active electorate.

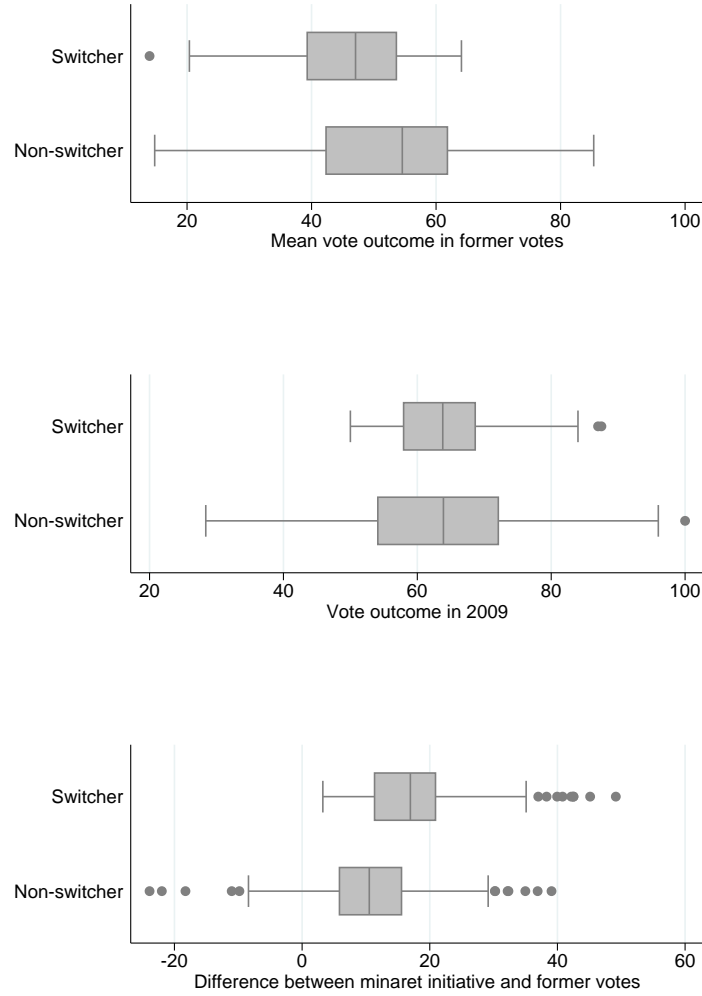
We label these municipalities with the generic term "switcher municipalities".

The first criterion is motivated by the idea that the respective municipalities were perceived as being relatively tolerant up to the vote on the minaret initiative. The second criterion ensures that only municipalities which experienced a large shift are considered to be switchers, given that nearly all the municipalities shifted to the right. The third criterion guarantees that the municipality voted in favor of the minaret initiative thus providing a signal that the majority agrees with the proposition. The latter point is important for the media coverage (see below). Based on these three criteria, we classify approximately 24 percent of all municipalities in Switzerland as switchers.<sup>3</sup> The following graphs in Figure 1 show how the voting results are distributed when comparing switcher and non-switcher municipalities. Descriptive statistics on the criteria are reported in Table B.2. First, citizens in switcher municipalities voted less restrictive on migration issues in former votes (used to calculate the ex ante perceived level of critical attitudes towards foreigners). Second, however, in the minaret initiative municipalities from the two groups cast similarly restrictive votes, on average. Third, when focusing on the change in the support of a restrictive position towards migration, a larger difference is observed for switcher municipalities than for non-switcher municipalities.

In a supplementary analysis, we, moreover, check whether the switcher municipalities merely became more similar to the municipalities that expressed stronger reservations in former votes. As a generally large fraction of the voting population approved a restrictive position in the minaret initiative, it could well be that the relative positioning of the municipalities in terms of expressed reservations has not changed at all. Accordingly, we rank municipalities within their cantons, once with regard to their average yes vote share in the four migration-related propositions that took place before the minaret initiative, and once on the basis of the vote outcomes in the minaret initiative. Figure 2 shows the average rank changes for switcher and non-switcher municipalities in all Swiss cantons. A positive change indicates that the municipality increased its rank by expressing a relatively more restrictive position towards immigration. A clear picture emerges that switcher municipalities gained in rank and non-switcher municipalities lost in rank. The relative strength of expressed reservations towards foreigners has thus changed between switcher and non-switcher municipalities.

---

<sup>3</sup>Please note that other reasonable definitions classify mostly the same municipalities as switchers and thus which particular criteria are used does not qualitatively change our results.

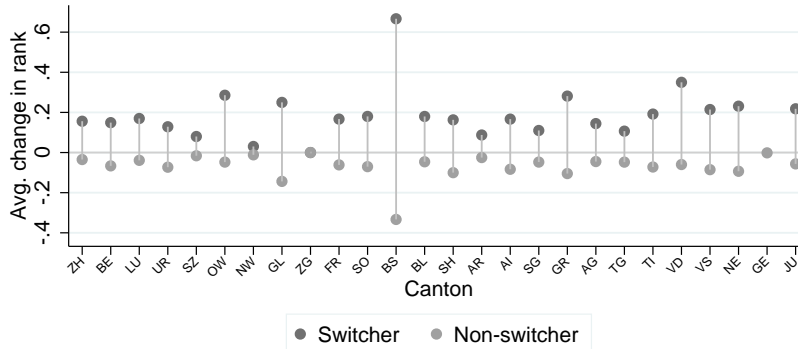


**Figure 1:** Voting outcomes in switcher and non-switcher municipalities

### 2.3 Information dissemination

In the aftermath of the unexpected approval of the minaret initiative, there was substantial post-election news coverage. On all the Swiss media channels, the unreckoned outcome was discussed and the municipal voting patterns made prominent news. Most Swiss newspapers presented maps indicating which municipalities of a canton supported or rejected the initiative. Sometimes, even the vote district results were discussed and displayed.<sup>4</sup> For Swiss citizens and foreigners alike, it was easy to learn about attitudes across municipalities and to spread the word about the municipalities that were expressing more pronounced reservations towards foreigners.

<sup>4</sup>Figures B.5 and B.6 in Appendix B provide examples of maps displaying the outcome where approval is marked green and rejection is marked red.



**Figure 2:** Rank change of switcher and non-switcher municipalities. A positive rank change indicates that the involved municipalities voted more restrictively on migration in the minaret initiative than in previous votes (compared to the other municipalities in a cantonal ranking). The rank change is measured as a percentage of the number of municipalities within the canton.

### 3 Theory

To derive our main hypothesis and to point out important conditions for identification, we formalize our ideas in a simple location choice model that takes identity utility into account. We explicitly restrict our model to the location choice and take the decision to move as given. Thus, we model the choices of individuals once they have decided that they will move and face a continuum of municipalities and their characteristics in their choice set. We argue that their decision to move is much less affected by the new information about attitudes towards foreigners than their decision about where to move. First, the decision to move is much more costly than a change in the decision about where to move. Second, individuals residing in municipalities for which negative attitudes towards foreigners are disclosed might have already been aware of their neighbours' attitudes and would know that nothing had really changed. We test this supposition in a supplementary empirical analysis (see Section 6.5) and do not find systematic reactions to the vote outcome in the decisions to move.

We start this section by setting out the equilibrium conditions for foreigners' choices between municipalities when they want to move; here, we evaluate choices in terms of utility losses from ostracism and the feeling of being unwelcome. We then derive our main hypothesis about how changes in perceived attitudes towards foreigners affect their moving behavior. In order to understand the preconditions for a test of the hypothesis, we simulate how new information about attitudes is incorporated and reflected in the observed aggregate moving behavior.

### 3.1 Location choice including identity utility

To model location choices of foreigners, we assume that the economy is in equilibrium and that individuals take municipality characteristics as given when choosing their location. Foreigners are a minority, and their decisions influence neither the wage level, as it is driven by the majority, nor the equilibrium for housing prices, in the short run.

The utility  $U_{i,j,t}$  of an individual  $i$  ( $i = 1, \dots, n$ ) living in municipality  $j$  ( $j = 1, \dots, J$ ) at time  $t$  depends on several municipality characteristics  $C_{j,t}$  and individual characteristics  $Id_{i,t}$ . Municipality characteristics that influence individual utility are taxes  $T$ , housing prices  $p$ , the provision of public goods  $g$ , but also natural beauty or social ties. Individual characteristics involve an individual's income, the preference for housing, and other individual characteristics such as family status. An individual's utility from residing in  $j$  in  $t$  is thus defined as  $U_{i,j,t}(C_{j,t}; \dots; Id_{i,t})$ , whereby we abstain from defining a specific functional form. We use standard assumptions concerning marginal utilities and marginal rates of substitution, assuring that households can be compensated for higher values of variables they dislike by lower values of others of the same type. For example, a tax rise can be compensated by a drop in housing prices.

In order to capture that individuals' utility might be affected by the perceived attitudes of others, we model an individual's utility as dependent on his or her identity. People have a sense of self and care about how they perceive themselves as human beings and how they are perceived by others. In our specific application, we aim at identifying the effect of a negative change in perceived attitudes towards an individual's social group or category, i.e., being a foreigner.

In social psychology, it is well documented that perceiving even slightly repellent behavior of others can produce psychological and physical distress. Investigations on ostracism show that individuals strongly react to perceived exclusion, as it threatens the fundamental needs of belonging and increases negative affect. For example, regions of the brain that are important for pain detection are activated. Moreover, individuals are highly sensitive in detecting situations of exclusion (see, e.g., Williams, 2009; Williams and Nida, 2011).

We follow the seminal work on the economics of identity by Akerlof and Kranton (2000) to define central aspects of a person's identity utility. In particular, identity utility  $I_i$  depends on  $i$ 's action  $a_i$ , on the actions of others  $a_{-i}$ , and on the assigned social category  $c_i$ , i.e.,  $I_i = I_i(a_i, a_{-i}, c_i)$ .<sup>5</sup> The actions of others which, in our case, focus on the revealed attitudes of natives can differ across municipalities  $j$  and in time  $t$ . We define the utility function, its derivatives, and marginal rates of substitution as follows:

$$U_{i,j,t}(C_{j,t}; \dots; Id_{i,t}; I_{i,j,t})$$

---

<sup>5</sup>For our application, we can abstract from the extent to which an individual's attributes match the ideals of his or her social category and actions match prescriptions on how to behave.

$$\frac{\partial U}{\partial T} < 0; \frac{\partial U}{\partial p} < 0; \frac{\partial U}{\partial g} > 0; \frac{\partial U}{\partial I} > 0$$

$$MRS_{I,T} := \frac{dI}{dT} > 0; MRS_{I,p} := \frac{dI}{dp} > 0; MRS_{I,g} := \frac{dI}{dg} < 0.$$

$I$  can take on positive values or negative ones to capture the possible disutility from living in a municipality. Individuals can be compensated for lower identity utility by advantageous values of other relevant factors. We restrict the social category  $c_i$  to a one-dimensional continuous measure between  $c_i = 0$  (*native*) and  $c_i = 1$  (*foreigner*). The assignment to a particular category depends on individuals' subjective perception.

We do not make any specific assumption about how the various factors influence  $I$  except for the behavior of others  $a_{-i,j,t}$ . We assume that individuals update their information about their set of choices at the level of municipalities when they move. Moreover, the behavior of the people in a municipality towards an individual's social category can be approximated as  $a_{-i,j,t} = a_{-i,j,t}(c_i * at_{j,t})$ , i.e., as a function of the social category  $c$  and  $at_{j,t}$ , which describes the municipalities' attitudes towards the social category of individual  $i$ . The value of  $at_{j,t}$  ranges between 0 and 1 where 0 stands for rather positive or neutral attitudes and 1 for rather negative or hostile attitudes towards people with higher values of  $c_i$ .

$$\frac{\partial I}{\partial at} < 0 \text{ and thus } \rightarrow \frac{\partial U}{\partial at} < 0.$$

In sum, the more negative the attitudes towards someone's social group are, the less identity utility he or she will draw from living in a particular municipality.<sup>6</sup> We assume that individuals regularly update their perceptions of  $at_{j,t}$  based on the past voting record of the municipalities' taking into account votes related to their social category. Thus,  $at_{j,t}$  is a function of past vote outcomes  $\bar{v}_{j,t}$ . Overall, the framework includes the possibility that individuals' utility is affected by the revealed preferences of the voting population in their place of residence, and accordingly that this influences the utility of locating in a particular municipality.

### 3.2 Changes in perceived attitudes and moving decisions

Revealed attitudes towards foreigners in a municipality may change over time and thus also individuals' optimal location choices. Individuals choose their location such that their utility is maximized within their choice set. Thus, individual  $i$  will choose to locate in municipality  $j$  if, and only if,

$$U_{i,j,t} > U_{i,k,t} \text{ for all } k \neq j,$$

---

<sup>6</sup>Related evidence based on panel data in Germany shows that higher vote shares of right-wing parties are associated with lower reported life satisfaction for immigrants (Knabe et al., 2013).

the utility to live there is higher than in all other municipalities in the choice set.

We will later compare the location choice of individuals around a point in time. We assume that the location choice of individuals is in equilibrium before  $t$  and that new information about attitudes across municipalities ( $at_{j,t}$ ) is revealed between  $t$  and  $t + 1$ , thus leading individuals to update their utility expectations. In our specific setting, the newly revealed information will be the voting outcome on the minaret initiative in Switzerland. This new information can be seen as an exogenous change in  $at$ , as it was absolutely unexpected. Thus, if an individual who would have chosen  $j$  in  $t$ , chooses a different municipality  $k$  in  $t + 1$ , the following should hold.

$$U_{i,j,t} > U_{i,k,t} \text{ for all } k \neq j, \quad (1)$$

as the individual has chosen municipality  $j$  in  $t$ , but

$$U_{i,j,t+1} < U_{i,k,t+1} \text{ for at least one } k \neq j, \quad (2)$$

as the individual chooses another municipality in  $t + 1$ .

Thus, municipality  $j$  has become relatively less attractive:

$$\underbrace{\frac{U_{i,t+1,k}}{U_{i,t+1,j}}}_{>1 \text{ from (2)}} > \underbrace{\frac{U_{i,t,k}}{U_{i,t,j}}}_{<1 \text{ from (1)}}$$

either because the utility to live in  $j$  has fallen from  $t$  to  $t + 1$ , or the utility to live in  $k$  has risen; or both.

There are two crucial conditions that need to hold in order to ascribe a potential change in the location choice to the change in perceived local attitudes towards the social group  $c$ .

- I. Municipality characteristics have not changed systematically between  $t$  and  $t + 1$  ( $C_{j,t} = C_{j,t+1} = C_j$ ).
- II. Moving individuals are, on average, the same in  $t$  and  $t+1$ . Thus, individual-specific characteristics have not changed systematically between  $t$  and  $t+1$  ( $Id_{i,t} = Id_{i,t+1} = Id_i$ ).

Given that these conditions hold, the only component in the utility function that could have changed the relative attractiveness of municipality  $j$  is the identity utility from perceived attitudes ( $at_{j,t}$ ) towards an individual's social category  $c_i$ .<sup>7</sup> Thus, a systematic change in the location choice can be attributed to a change in perceived attitudes.

---

<sup>7</sup>Please note that the two conditions are close to the identifying assumptions in a RDD context. Condition I is close to the *conditional independence assumption*, and condition II is close to the *local randomization assumption*.



Whether somebody is affected by a change in attitudes towards foreigners depends on the extent to which he or she identifies with this social group, i.e., the value of  $c_i$ . Residents with a foreign passport might identify strongest with the social group of foreigners due to their ancestry, ethnicity or prior experiences. As an individual to some extent chooses his or her identity, in our test we explicitly allow that an individual's self-ascribed social category might also react to the new information. This idea is grounded in the *rejection-identification model* in social psychology (Tajfel, 1978; Branscombe et al., 1999). This theory asserts that when a minority group member perceives prejudice against his or her group, that member will develop a stronger identification with the group in order to reduce the psychological cost of perceived prejudice. The model further states that this mechanism might not be limited to clearly-defined group identities. Group identities may develop as a result of perceived prejudice, i.e., minority groups tend to identify with other minority groups under threat. It is well possible that various groups of foreign residents react to votes on topics that do not relate directly to them, but to other groups such as religious minorities, because the issues concern minorities, where being an 'outsider' becomes salient. International students, for example, tend to identify themselves as an independent group in response to perceived prejudice (Schmitt et al., 2003). And individuals who observe someone being ostracized, whose perspective they are told to adopt, identify with the victim and experience personal distress akin to being the target of the observed ostracism (Wesselmann et al., 2009). Moreover, in recent experimental and survey studies, Mechtel et al. (2014) provide laboratory evidence that social identity affects individuals' decisions and Rudert et al. (2014) conduct a survey which reports the responses of foreigners in Switzerland after a popular vote on migration in 2014 in order to investigate whether individuals feel ostracized. The latter study finds that foreigners' perceptions of their environment reflect the vote's outcome and that immigrants' need for belonging is less satisfied where citizens voted more restrictively.

Building on these considerations from economics and social psychology, we hypothesize that a negative change in perceived attitudes widens the pool of people who identify themselves as foreigners, and also causes them psychological distress. In our context, the minaret initiative, we would expect not only Muslims but also other groups of foreign residents to experience ostracism, as their identity as foreigners in Switzerland becomes salient. Specifically, we would also expect EU citizens to be negatively affected by the initiative, if the identity as a foreigner is the moderating factor.

### 3.3 Identification of identity utility based on location decisions

The choice set of municipalities remains the same before and after the signal about residents preferences is revealed. Moreover, we assume that the other municipality characteristics do not systematically change around the voting date. We will later on validate

this assumption. Accordingly, it should be possible to identify the effect of the signal on the location choice by comparing the location decisions of foreigners who choose their residency before and after the vote has taken place. As we do not find evidence that the decision to move is influenced by the vote (see Section 6.5), it can be assumed that individuals who make their location choice before and after the new information has been revealed are, on average, the same. Individuals decide to move at some point in time. Whether they find a new home on a specific day can be considered as random with respect to a specific voting date. Consequently, it is more or less random whether an individual looking for a new apartment is successful in his or her search the day before or after the vote has taken place. While an individual who finds a new apartment the day before the vote cannot incorporate the information on the outcome that will be revealed the next day in his or her choice, an individual who finds an apartment the day after can utilize the information on the outcome. Hence, individuals who decide to move and looking for an apartment narrowly before the voting day can be used as a control group for those who make this decision shortly after the vote has taken place (Thistlethwaite and Campbell, 1960). If revealed preferences of the voting population about their attitudes towards foreigners play no role for the relative attractiveness of municipalities, foreigners' location decisions should, on average, be the same before and after the vote. However, if the mover is affected by perceived attitudes towards his or her group, we should observe a clear change in the mover's location decision.

In order to apply this theoretical identification strategy, we would like to compare the location choices of people who decided to move before the new information became public with those who decided afterwards. In particular, we would like to exploit the presence of switcher municipalities (described in Section 2), where we assume that new information was unexpectedly revealed. We thus aim to compare the probability that a foreigner chooses to move to one of the identified switcher municipalities before receiving the new information and after receiving it. If it were possible to determine a priori on which date the effect on location choices materializes, the test could be approached using a regression discontinuity design right away, i.e., the local randomization around the specific date could be used. However, it is not evident when this date might be. Shortly after the vote on the minaret initiative, many observed movers must have made their location choice before the vote took place and who were therefore unable to incorporate the new information. The longer the time is that elapses after the polling day, the larger will be the number of movers who have made their location choice based on the updated information set. The two groups might therefore overlap for some time until only movers who could incorporate the new information are observed. The corresponding time lag relative to the polling day might be affected by the length of notice periods and the situation in the housing market. In the next subsection, we provide a framework to analyze how the new information is

incorporated into individuals' observed location choices. In Section 4, we translate the insights into our empirical strategy.

### 3.4 Incorporation of new information in observed location choices

The location choice equilibrium for foreigners materializes in a stable aggregate probability of choosing a particular municipality type. This holds over time and refers in our context to the switcher and non-switcher municipalities. For the individual location decision, we, however, only observe the moving date, but not the decision date. Accordingly, "observed" individual location choices might be based on different information sets. There is a timespan after the vote, during which we observe a mixture of those people who made their location choice before the information became public and those who decided subsequently. The effect of the change in perceived attitudes, given there is an effect, materializes to the extent that the search generation that could incorporate the new information outweighs the prior search generations. This transition between the search cohorts could be smooth or sharp. To gain an idea of how the effect could materialize, we perform a simulation analysis of individual location choices.

Appendix A presents the details of the analysis. We simulate a simple location search with overlapping search generations. Two institutional scenarios are considered, and a negative treatment effect on the probability of moving to a switcher municipality is assumed. First, only search frictions exist, and people move once they have found an apartment. Second, a notice period is taken into account. It is clearly revealed that in the second setting, the new information is incorporated later on in the aggregate outcome than in the first setting, i.e., in the inflow of foreigners to the switcher municipalities. Moreover, the adjustment pattern also reveals a much steeper drop in the probability of moving to a switcher municipality in the scenario with a notice period than in the scenario without one. With search frictions alone, in most cases, a rather gradual adjustment to the new equilibrium occurs.

Based on the assumption that it is random when an individual chooses to move and finds an apartment, a regression discontinuity framework is suitable to analyze the incorporation of the new information. We apply the method of a regression discontinuity design with unknown discontinuity points (as introduced in the following Section 4) to measure the drop. The simulation clearly indicates that in an environment with legal restrictions that delay moving - such as a notice period - the method can find a discontinuity. This is not always the case in the scenario with search frictions alone. Given the rental agreements in Switzerland that stipulate mandatory as well as discretionary notice periods, the effect of a change in attitudes towards foreigners should be identifiable in observed moving dates.

## 4 Empirical strategy

Based on the theoretical considerations, we derived predictions on how individual location choices in response to new information translate into changes in aggregate migration flows between municipalities. In an environment with institutional restrictions, any displacement effect, given one exists, is expected to materialize as a jump (or discontinuity) in the probability that a foreigner will choose to locate in a switcher municipality some time after the vote has taken place. The regression discontinuity design (RDD) is an appropriate method to measure treatment effects at such points. However, in most RDD applications the discontinuity point is known, since it is, for instance, determined by the institution being analyzed. In our application though, this is not the case. The discontinuity point is the result of the interaction between the state of the housing market and institutional conditions such as the notice period. As a consequence, it is not deterministically predictable a priori when the discontinuity is expected to occur.

Porter and Yu (2015) have recently proposed an approach that is particularly well suited for the problem at hand. They exploit the fact that testing for the presence of a treatment effect in such a context is similar to nonparametric testing for a structural change. They develop a nonparametric procedure to perform regression discontinuity analyses when discontinuity points are unknown, but predicted by theory. They describe a procedure that starts with a specification test. This tests the null hypothesis of no treatment effect, or that of no discontinuity in the relationship between a dependent variable  $y$  and a running variable  $x$ . Subsequently, the discontinuity point or threshold is estimated, given that the specification test rejects the null hypothesis. Finally, a standard RDD analysis is performed.

### 4.1 Regression discontinuity designs with unknown discontinuity points

In classical studies with regression discontinuity designs, the forcing variable, also referred to as the running variable, is used to determine the treatment status. When the value of this variable exceeds a certain threshold, observations are treated, but not so below this value. In the classical case, this threshold is deterministically known and defined by some policy. If individuals have no perfect control over the value of the forcing variable, the random variation narrowly around the threshold can be exploited to estimate local average treatment effects at this particular point (see, e.g., Imbens and Lemieux, 2008, and Lee and Lemieux, 2010 for reviews of the RDD).

However, the threshold value might not be known *ex ante* even though theory predicts that there should be one. Card et al. (2008) are - to our knowledge - the first who try to estimate an unknown threshold value in order to estimate effects at this point. In their

application, they are interested in the presence of a minority population in a municipality and the share after which the majority starts to leave the municipality; i.e., a tipping point based on the Schelling model about dynamics in segregation (Schelling, 1971).

Here, we follow the approach proposed by Porter and Yu (2015) and apply the RDD with unknown discontinuity points. In our application, we concentrate on the sharp design. We further only discuss the case with one discontinuity point, even though their approach allows for multiple discontinuity points.<sup>8</sup>

The relationship of interest between the dependent variable  $y$  and the running variable  $x$  when estimating a RDD can be expressed by the following reduced-form expression:

$$y = m(x) + \epsilon = m_c(x) + \tau_c d_c + \epsilon.$$

Where  $m(x)$  is some function of  $x$ ,  $d_c$  is a treatment indicator ( $d = \mathbf{1}(x > c)$ ),  $E[\epsilon|x, d_c] = 0$  and  $\tau_c$  is the potential treatment effect at  $c$ .  $m_c(x) = m(x) - \tau_c d_c$  and  $c \in \Pi = [\underline{c}, \bar{c}]$ .

The starting point of the procedure is to test whether there is a discontinuity in the relationship. Therefore the proposed approach by Porter and Yu (2015) first tests two hypotheses in order to disentangle whether there is a selection effect, a treatment effect, both, or neither of the two:

$$\begin{aligned} H_0^{(1)} &: \text{no effects} \\ H_1^{(1)} &: \text{selection only} \end{aligned} \tag{3}$$

$$\begin{aligned} H_0^{(2)} &: \text{no effects \& selection only} \\ H_1^{(2)} &: \text{treatment effect only \& both selection and treatment effect} \end{aligned} \tag{4}$$

They further propose conducting these tests sequentially and starting with the test of (4). If  $H_0^{(2)}$  cannot be rejected, a test of (3) should be performed. They develop a test statistic for which the only difference in testing (3) or (4) is the choice of a different smoothing parameter. When testing (3),  $m(x)$  is estimated nonparametrically, thereby overersmoothing the estimate so that the resulting bias generates power. When testing (4), however,  $m(x)$  is undersmoothed, and the bias in the selection case will disappear asymptotically, while a jump in  $m(x)$  would still generate power. So the bandwidth in testing (3) is larger than the bandwidth in testing (4). As we are interested exclusively in cases where there is a treatment effect, we will only test (4). If  $H_0^{(2)}$  cannot be rejected, we abstain from further analysis. According to the approach by Porter and Yu (2015), it can be assumed that there is a discontinuity in  $m(x)$  if  $H_0^{(2)}$  is rejected by the specification test. It can be proceeded with the estimation of the threshold value  $c$ . Once  $c$  is known, the size of the discontinuity  $\tau_c$  can be estimated. Before discussing these latter aspects, we describe how the specification test is performed.

---

<sup>8</sup>A discussion of the fuzzy case and an in-depth discussion of the theoretical derivations and the method are presented in the original paper (Porter and Yu, 2015).

Porter and Yu (2015) derive a consistent specification test for the hypothesis. In order to find critical values, the empirical distribution of the test statistic under the null hypothesis is simulated.

Consider the relationship between  $y$  and  $x$  as

$$y = \bar{m}(x) + e. \quad (5)$$

Under  $H_0^{(1)}$   $E[e|x] = E[\epsilon|x] = 0$  for  $x \in \Pi$  and under  $H_1^{(1)}$   $E[e|x] \neq 0$ . By observing that  $E[eE[e|x]\mathbf{1}_x^\Pi] = E[E[e|x]^2\mathbf{1}_x^\Pi] = E[(m(x) - \bar{m}(x))^2\mathbf{1}_x^\Pi] \geq 0$  and that the equality only holds if  $H_0^{(1)}$  is true, they construct a consistent test for (3) based on  $E[eE[e|x]]\mathbf{1}_x^\Pi$ . They state that  $E[e|x]$  can be estimated by a kernel estimator, and choose to estimate a density-weighted version of  $E[eE[e|x]\mathbf{1}_x^\Pi]$ , and hence  $E[eE[e|x]f(x)\mathbf{1}_x^\Pi]$  in order to overcome the random denominator problem in kernel estimation.  $e_i$  is estimated by the nonparametric residual from (5) ( $\hat{e}_i = y_i - \hat{y}_i$ ) and  $E[e_i|x_i]f(x_i)$  by a kernel estimator to obtain a feasible test statistic.  $\hat{y}_i$  is a kernel estimator of  $\bar{m}(x_i)$  defined as

$$\hat{y}_i = \frac{1}{n-1} \sum_{j \neq i} y_j L_{b,ij} / \hat{f}_i$$

and  $\hat{f}_i$  is estimated by

$$\hat{f}_i = \frac{1}{n-1} \sum_{j \neq i} L_{b,ij},$$

where  $L_{b,ij} = \frac{1}{b} l(\frac{x_i - x_j}{b})$  is a second-order kernel<sup>9</sup> and  $b$  is the bandwidth. Note that  $(x_i, y_i)$  is excluded in estimating  $\bar{m}(x_i)$ .  $E[e_i|x_i]f_i$  is estimated by  $\frac{1}{n-1} \sum_{j \neq i} \hat{e}_j L_{h,ij} \mathbf{1}_j^\Pi$ . Finally, the test statistic is based on

$$I_n = \frac{nh^{1/2}}{n(n-1)} \sum_i \sum_{j \neq i} \mathbf{1}_i^\Pi \mathbf{1}_j^\Pi L_{h,ij} \hat{e}_i \hat{e}_j.$$

Under  $H_0^{(1)}$ ,  $\hat{e}_i$  is a good estimator for  $\epsilon_i$ , but under  $H_1^{(1)}$  it includes a bias in the neighborhood of  $c$ .<sup>10</sup> This generates power.

In order to use  $I_n$  in testing (4), it has to be made sure that  $E[eE[e|x]f(x)\mathbf{1}_x^\Pi]$  is approximately zero under  $H_0^{(2)}$ . Porter and Yu (2015) derive that, when estimating the test statistic, the bandwidths have to be smaller when testing  $H_0^{(2)}$  than those used for  $H_0^{(1)}$  in order to eliminate the bias in approximating  $m(x)$  by  $\bar{m}(x)$ .

They provide the asymptotic distribution of  $I_n$ . The asymptotic variance of the test statistic can be consistently estimated by

$$v_n^2 = \frac{2h}{n(n-1)} \sum_i \sum_{j \neq i} \mathbf{1}_i^\Pi \mathbf{1}_j^\Pi L_{h,ij}^2 \hat{e}_i^2 \hat{e}_j^2$$

<sup>9</sup>Please note that we use a uniform or rectangular kernel for the specification testing and the estimation of  $\hat{c}$  in our application. We do not want to give too much weight to single days because of the possible periodical patterns in individuals' moving behavior.

<sup>10</sup>Whereby for these estimates the following is recommended: Take  $h < b$ ;  $b \sim n^{-c}$ ;  $h \sim n^{-(2c+\delta)}$  where  $\delta$  is an arbitrarily small positive number and  $\frac{1}{4} \leq c < \frac{1}{2}$ .

for the testing of both hypotheses. The resulting studentized test statistic is  $T_n = I_n/v_n$ . As the convergence rate of  $T_n$  to a standard normal is very slow, they propose approximating the finite-sample distribution by using the wild bootstrap of Wu (1986) and Liu (1988). The bootstrap samples are generated by imposing the null hypothesis; i.e., such that they will mimic the null distribution of  $T_n$ . The algorithm is specified as follows:

Step 1: Generate the two-point wild bootstrap residual for  $i = 1, \dots, n$ .  $\epsilon_i^* = \hat{\epsilon}_i(1 - \sqrt{5})/2$  with probability  $(1 + \sqrt{5})/(2\sqrt{5})$ , and  $\epsilon_i^* = (1 + \sqrt{5})/2$  with probability  $(\sqrt{5} - 1)/(2\sqrt{5})$ . Then  $E^*[\epsilon_i^*] = 0$  and  $E^*[\epsilon_i^{*2}] = \hat{\epsilon}_i^2$ .<sup>11</sup>

Step 2: Generate the bootstrap resample  $\{y_i^*, x_i\}_{i=1}^n$  by  $y_i^* = \hat{y}_i + \epsilon_i^*$ . Only data with  $x_i \in [\underline{c} - b, \bar{c} + b]$  are needed to construct  $I_n^*$ . Then calculate the bootstrap residuals  $\hat{\epsilon}_i^* = y_i^* - \hat{y}_i$ .<sup>12</sup>

Step 3: Use the resulting samples to calculate the test statistic  $T_n^* = I_n^*/v_n^*$  using the same bandwidths as in the estimation of  $T_n$ . Where

$$I_n^* = \frac{nh^{1/2}}{n(n-1)} \sum_i \sum_{j \neq i} \mathbb{1}_i^\Pi \mathbb{1}_j^\Pi L_{h,ij} \hat{\epsilon}_i^* \hat{\epsilon}_j^*$$

and

$$v_n^{*2} = \frac{2h}{n(n-1)} \sum_i \sum_{j \neq i} \mathbb{1}_i^\Pi \mathbb{1}_j^\Pi L_{h,ij}^2 \hat{\epsilon}_i^{*2} \hat{\epsilon}_j^{*2}.$$

Step 4: Repeat the three former steps B<sup>13</sup> times and use the resulting empirical distribution to approximate the null distribution of  $T_n$ .  $H_0$  can consequently be rejected if  $T_n > T_{n(\alpha B)}^*$ .  $T_{n(\alpha B)}^*$  is the upper  $\alpha$ -percentile of the empirical distribution of  $T_n^*$ .

As discussed before, this test can equally be used to test  $H_0^{(2)}$ . If  $H_0^{(2)}$ , which implies that there is no discontinuity in the relationship, can be rejected by the specification test, the threshold value  $c$  and the discontinuity size  $\tau_c$  can be estimated. Porter and Yu (2015) determine the discontinuity point by

$$\hat{c} = \arg \max_{c \in \Pi} \hat{\tau}^2(c),$$

and thus by checking whether  $c = x_i$  ( $x_i \in \Pi$ ) maximizes  $\hat{\tau}^2(c)$ .

The estimate of  $\hat{\tau}(c)$ , i.e., the nonparametric estimate of the discontinuity in the relationship between  $y$  and  $x$  at  $c$ , is thereby estimated as in the classical RDD literature. In

<sup>11</sup>Note that  $\epsilon_i^*$  is, in effect, simulated based on the centered  $\hat{\epsilon}_i$ ,  $\bar{\epsilon}_i = \hat{\epsilon}_i - \bar{\epsilon}$ . For more details, see Porter and Yu (2015).

<sup>12</sup> $\hat{y}_i^*$  is defined as  $\hat{y}_i$ , except that  $y_i$  in (5) is replaced by  $y_i^*$ .

<sup>13</sup>In our application, B is set to 1,500.

the sharp design, this amounts to the estimation of the difference in the limits of  $m$  when approaching the threshold from the left and from the right.

$$\hat{\tau}(c) = \hat{m}_+(c) - \hat{m}_-(c).$$

This is normally done using a local polynomial estimator as, for example, described in Porter (2003), Hahn et al. (2001), and Imbens and Lemieux (2008). Importantly, Porter and Yu (2015) show that  $\hat{\tau}_c$  is a natural by-product of the estimation of the discontinuity point, and that  $c$  can be treated as if known when estimating treatment effects in RDDs. The asymptotic distribution of  $\hat{\tau}(\hat{c})$  is the same as if  $c$  were known. After the estimation of  $c$ , one can perform a standard RDD analysis.

## 4.2 Estimation in the regression discontinuity design

We briefly describe how we estimate the local average treatment effect in our RDD<sup>14</sup> setting with time as the assignment variable, given that the null hypothesis of no treatment effect could be rejected.<sup>15</sup>

A peculiarity of this approach is that it is necessary to control for periodical patterns before running the analysis. We control for these patterns, which are very probable in moving patterns, using a battery of indicator variables. As often done in RDD studies, we first run a linear regression of the dependent variable on covariates. Subsequently, the residual of this regression is used as the dependent variable in our analysis (see, e.g., Lee and Lemieux, 2010).<sup>16</sup> For notational convenience, we will nonetheless refer to our dependent variable as  $y_{it}$  or the probability that an individual moves to a switcher municipality.

Given that the necessary assumptions for a sharp RDD hold and that the estimated threshold value  $\hat{c}$  is estimated based on the approach proposed by Porter and Yu (2015), the treatment effect  $\tau$  can be estimated by the following limits<sup>17</sup>

---

<sup>14</sup>A practical guide for the estimation of a RDD is provided by Imbens and Lemieux (2008).

<sup>15</sup>There are few studies that use time as the assignment variable so far. Davis (2008) is, to our knowledge, the first. He uses the introduction date of a program aimed at reducing pollution in Mexico City as the threshold. Another example is the study by Paola et al. (2012). They use a RDD approach over time to investigate the effect of the introduction of a system with penalty points for driving offenses in Italy. They identify the causal effect of the new system on road safety.

<sup>16</sup>We first run the specification

$$y_t = \beta_0 + \beta_1 x_t + u_t,$$

where  $y_t$  is our dependent variable and thus an indicator of whether an individual moved to a switcher municipality. The variable  $x_t$  includes indicator variables for the month, the day of the month, and the weekday of the movement. We perform this procedure separately for each canton and use data between 2006 and 2008 for the estimation. We then predict the residual  $u_t$  for the whole time span and use it as our dependent variable.

<sup>17</sup>The identification of a treatment effect in a RDD setting is only possible at  $x = c$ , i.e., a LATE is identified at  $x = c$  (Hahn et al., 2001).



$$\widehat{E(\tau_i|x_i = c)} = \lim_{x \rightarrow c^+} \widehat{E(y_i|x_i = x)} - \lim_{x \rightarrow c^-} \widehat{E(y_i|x_i = x)}. \quad (6)$$

Where  $y$  stands for the dependent variable,  $x$  represents the assignment variable (in our application the moving date),  $\lim_{x \rightarrow c^+}$  stands for the limit approaching the threshold from the right, and  $\lim_{x \rightarrow c^-}$ , the limit coming from the left.

For the point estimates in our analysis, we follow Hahn et al. (2001) and Porter (2003) and estimate the limits in (6) by a LLR using a uniform kernel, as in the specification testing, and variance estimates as proposed in Porter (2003).<sup>18</sup> To obtain, for example, an estimator for  $\lim_{x \rightarrow c^-} E(Y_i|x_i = x)$ , i.e., to the left of the threshold, a regression with a constant term with coefficient  $\alpha_{yl}$  ( $l$  for left) and a linear term in  $x_c$  with coefficient  $\beta_{yl}$ , to correct for the influence of  $x_c$  on the outcome, is specified ( $y_i = \alpha_{yl} + \beta_{yl}x_{ci} + \epsilon$ ). Thereby,  $x_c$  corresponds to the forcing variable  $x$  centered at the threshold  $c$  ( $x_c = x - c$ ). This estimation includes all data points within the range  $x - h \leq x < c$  (or  $x_c - h \leq x_c < 0$ ). The estimate  $\hat{\alpha}_{yl}$  then emerges from the solution to the following minimization problem

$$\min_{\alpha_{yl}; \beta_{yl}} \sum_{i=1}^N (y_i - \alpha_{yl} - \beta_{yl}x_{ci})^2 K_h(x_{ci} - 0) I(x_{ci} < 0)$$

(Hahn et al., 2001). The analogous procedure is applied for the estimation of the other limit in equation (6).  $h$  represents the suitable bandwidth  $h > 0$  and  $K_h(x_i - c)$ , the kernel weighting function, defined as  $K_h(x_i - c) = \frac{1}{h} K\left(\frac{x_i - c}{h}\right)$ . Since  $x$  is centered at  $c$ , the estimate of the constant corresponds to the value of the regression function at the threshold. The treatment effect estimator in the sharp RDD is given by  $\hat{\tau} = \hat{\tau}_c = \hat{\alpha}_{Yr} - \hat{\alpha}_{Yl}$ .

The asymptotic distribution of the treatment effect estimator can be expressed as

$$\sqrt{Nh}(\hat{\tau}_Y - \tau_Y) \xrightarrow{d} N(0, V_{\tau_Y}) \text{ as } Nh \rightarrow \infty.$$

The specification for the estimation on both sides of the determined threshold can be formulated as

$$y_i = \alpha + \beta(\text{moving\_date} - \hat{c}) + \epsilon.$$

For graphical representation, we plot a local linear regression of  $y$  over  $x$  using a rectangular kernel to gain an impression of the basic relationship. This is done either separately from both sides of the threshold or over the whole range of  $x$ , depending on the analysis at hand.

---

<sup>18</sup>Since the LLR framework is well understood, our analysis, like many others, abstains from applying a more complicated model (such as the GAM) and uses a linear probability model, which directly models  $P(Y = 1) = X'\beta$ . This model can be directly estimated using LS, or LLR in the present case. The drawbacks of this approach are clear. Using this linear model, it is possible to obtain predicted probabilities below zero or above one, which is at odds with the definition of a probability. However, in many applications this approach seems to be a useful simplification (see, e.g., Cameron and Trivedi, 2005, p. 466ff.).

### 4.3 Assumptions of the regression discontinuity design

The crucial assumptions for the validity of a sharp RDD are the *local randomization* and the *continuity of conditional expectations*, which provides that, without the treatment, both groups are alike at the threshold. The latter assumption makes sure that in a world without the treatment or in a world where everyone is treated, there would be no discontinuity in the conditional expectations at the threshold, i.e., that the effect found can be attributed to the received treatment. Local randomization describes the situation that it is more or less random whether an individual decides to move before or after the vote, i.e., that individuals are not able to perfectly manipulate their value of the assignment variable (Lee and Lemieux, 2010; Hahn et al., 2001). Note that these assumptions in some respect also evolve from our theoretical reasoning in Section 3. In our setting, they translate to the assumptions that individuals deciding to move before or after the vote are on average the same and that the characteristics of switcher municipalities, except for their voting behavior, have not changed systematically. Further, the basic assumption also implies that the moving probability of individuals does not change at the threshold. Otherwise, this would correspond to a manipulation of the running variable. We present a series of validation checks for the specific assumptions after the results in Section 6.

## 5 Data

We primarily use comprehensive register data on the whole foreign population in Switzerland for the years between 2006 and 2011, which was kindly provided by the Swiss Federal Statistical Office (BFS) in cooperation with the Federal Office for Migration (BFM). We thus rely on high quality data when it comes to moving dates. We restrict our estimation sample to be the universe of all foreigners in Switzerland holding either the residence permit B, for temporary residency, or C, for permanent residency. We thus exclude foreigners with other types of permits such as refugees, who are not allowed to choose their location freely. We exclusively use data on individuals moving within Switzerland, and thus exclude those just immigrating or leaving the country. In total, this sample includes roughly 7,854,025 observations and 559,951 incidences of foreigners moving.

Further, we use data on municipality characteristics and voting outcomes. This data is freely available on the website of the Swiss Federal Statistical Office. Data on the municipality tax rates has been kindly provided by Parchet (2014).

## 6 Results

The results of our empirical analysis are presented in four steps. First, we perform the analysis for one canton, i.e., the canton of Thurgau, to show how we apply the empirical strategy. In a second step, we pool the observations of those cantons for which a discontinuity is detected. Third, we test and reject an alternative supply-side explanation of the general phenomenon; i.e. landlords' discriminatory behavior. In a fourth step, we report the results of the validation of the identifying assumptions.

The empirical analysis uses data for 22 Swiss cantons. Separate analyses are performed for each canton, as the threshold date might vary depending on the situation on the housing market and institutional variation.<sup>19</sup> We abstain from testing for a treatment effect in four cantons, namely the cantons of Uri, Obwalden, Appenzell-Innerrhoden and Geneva. The former three cantons are rather small and, although we draw on register data, there are too few observations to perform the analysis. Furthermore, we exclude the canton of Geneva, since no municipality in this canton accepted the minaret initiative, and thus there is no switcher municipality.

### 6.1 Exemplary analysis for a single canton

In order to provide an intuition for the application of the methodology used, we first exemplify our econometric proceeding for the canton of Thurgau (TG). The dependent variable of interest is an indicator of whether a moving individual chooses to locate in a switcher municipality.<sup>20</sup>

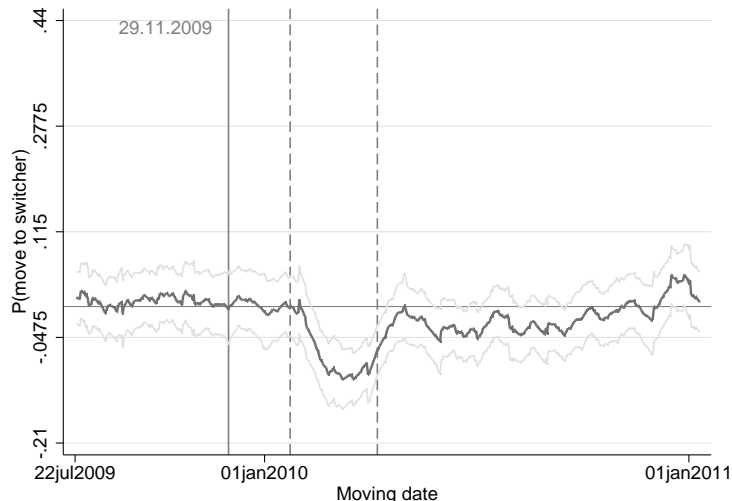
We first visually inspect how the probability that a foreigner chooses to move to one of the switcher municipalities evolves over time, whereby we undersmooth the estimate using the same bandwidth as in the specification testing. As can be seen in Figure 3, the probability is rather stable until the vote, which is marked by the solid vertical line. Some time after the vote has taken place, the probability seems to drop temporarily. For the detection of a possible threshold, we define a time span within which we expect the jump to occur, if there is one. This time span is marked by the two dashed vertical lines.

We perform the specification test as described in Section 4; i.e., testing the null hypothesis of no treatment effect, within the defined time span. We find that the null hypothesis is rejected for this canton. The test statistic is 2.62 and exceeds the critical value of

---

<sup>19</sup>Thus, we test for a discontinuity in the following cantons: Zürich (ZH), Bern (BE), Luzern (LU), Schwyz (SZ), Nidwalden (NW), Glarus (GL), Zug (ZG), Fribourg (FR), Solothurn (SO), Basel-Stadt (BS), Basel-Landschaft (BL), Schaffhausen (SH), Appenzell-Ausserrhoden (AR), St. Gallen (SG), Graubünden (GR), Aargau (AG), Thurgau (TG), Ticino (TI), Vaud (VD), Valais (VS), Neuchâtel (NE), and Jura (JU).

<sup>20</sup>As addressed before, we use the residual of a regression on time indicators to control for periodical patterns, thus in "normal" times it varies around zero.



**Figure 3:** Probability of moving to a switcher municipality in the canton of Thurgau. Local linear smooth of the probability that a foreigner moves to a switcher municipality, using a bandwidth of 30 days.

0.79 at the 5% significance level. These results are also listed in Table B.3 in Appendix B. The estimated threshold date for the canton of Thurgau is March 2, 2010 and, thus about 3 months after the vote has taken place. We proceed with a classical RDD analysis and plot a local linear smooth separately from both sides of the estimated threshold. In Figure 4, a negative jump in the probability that a foreigner chooses to locate in a switcher municipality is observed at the estimated threshold date. Moreover, it seems that the effect fades and that after some time the probability that a foreigner moves to a switcher municipality returns to its former level.

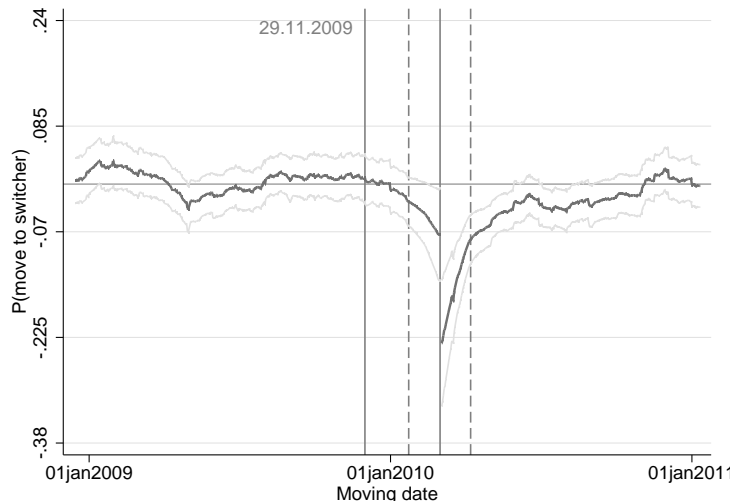
When we estimate the treatment effect at the detected threshold, we find a statistically significant negative discontinuity at the particular date. The probability of moving to a switcher municipality drops by 16 percentage points at the threshold. The details of the estimation result are listed Table B.4 in Appendix B. This is first evidence demonstrating that foreigners are deterred from locating to switcher municipalities where voters unexpectedly expressed stronger reservations towards foreigners.

## 6.2 Pooled analysis

### Overall effect

We pool the data of 12 cantons for which we detect a discontinuity.<sup>21</sup> They account for 60% of all moving individuals in our data. The corresponding results of the specification

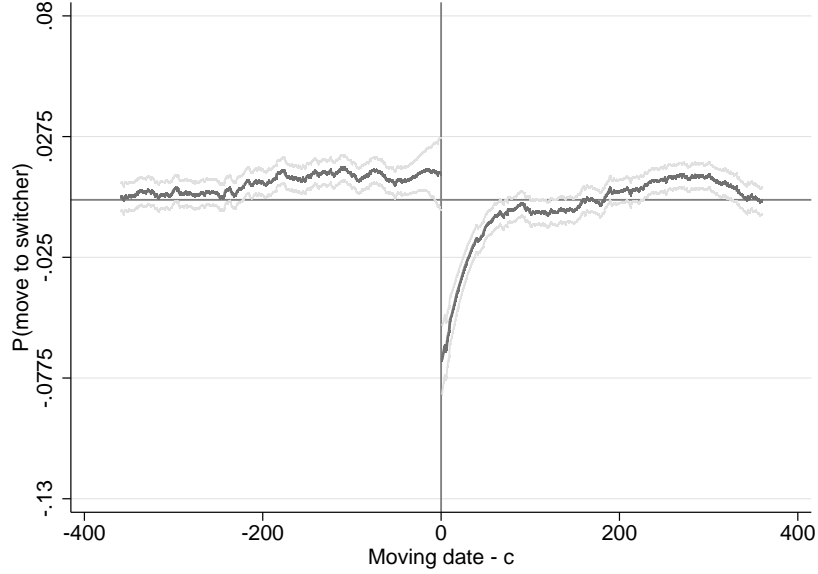
<sup>21</sup>These cantons are: ZH, BE, SZ, NW, GL, SO, AR, GR, TG, TI, VD, and VS.



**Figure 4:** Probability of moving to a switcher municipality in the canton of Thurgau - RDD graph around the threshold date March 2, 2010. Local linear smooth of the probability that a foreigner moves to a switcher municipality separately from both sides of the threshold, and using a bandwidth of 60 days.

test are presented in Table B.3 in Appendix B.<sup>22</sup> The null hypothesis of no effect is not rejected for the remaining ten cantons. This does not necessarily exclude the possibility that foreigners in these cantons also reacted. However, the situation on the housing market or the incorporation of the new information might not have followed a pattern which would have allowed the applied method to capture the effect. For our analysis of the pooled observations, we center the moving dates at the detected thresholds, such that the resulting threshold for the pooled results is zero. The threshold dates are estimated to lie between about 1 month and 5.5 months after the vote. Combining data for all moving foreigners in cantons with a discontinuity and repeating the RDD plot of LLR smooths from both sides of the threshold, we obtain the pattern presented in Figure 5. A pronounced negative jump at the threshold is observed. Moreover, the probability of moving to a switcher municipality seems to return to its former level after some time.

<sup>22</sup>Regarding the necessary bandwidth choice in the applied nonparametric procedure, we abstain from using a technique that either depends on the variance of the assignment variable or cross-validation. First, our running variable is a date and the variation of such a variable makes no sense. Cross-validation would be theoretically appealing. However, it is far too computationally intensive, given the large number of observations and the simulation-based approach. We rather rely on eye-balling and group cantons according to size when choosing the bandwidth that is as small as possible but which makes the graphs reasonably smooth. We start with a bandwidth size of 45 days for the largest cantons ZH, BE, AG, and VD. This bandwidth size ensures that each estimation includes one beginning and one end of the month, when relocations are frequent. The next group contains the cantons LU, FR, SO, BL, SG, TG, TI, and VS for which a bandwidth of 60 days is applied. We use 75 days for the cantons SZ, ZG, BS, SH,GR, and NE. For the smallest ones, i.e., NW, GL, AR, and JU, a bandwidth of 90 days is chosen. These bandwidths are used for the RDD graphs, the RDD estimates and the threshold search. As the bandwidth for the testing has to be considerably smaller than for these estimates, we choose to use half of it in the testing procedure.



**Figure 5:** Probability of moving to municipalities revealing stronger reservations against foreigners - RDD graph for pooled data. Local linear smooth of the probability that a foreigner moves to a switcher municipality separately from both sides of the threshold, and using a bandwidth of 45 days.

Estimating the treatment effect at the threshold as described in Section 4, the estimated drop in the probability that a foreigner moves to a switcher municipality is about 4 to 8 percentage points, depending on the applied bandwidth. The corresponding estimation results are presented in Table 1. They are rather robust to the bandwidth choice. The estimated effect is sizable. Before the threshold dates, the average probability that foreigners move to a switcher municipality was about 13 percent. The jump at the threshold of about 4 to 8 percentage points, thus amounts to a drop by about 30 to 60 percent.

**Table 1:** Probability that foreigners move to municipalities revealing stronger reservations towards foreigners - RDD estimates using pooled data

Dependent variable: Moving to a switcher municipality						
	(1)	(2)	(3)	(4)	(5)	(6)
$\tau$	-0.0824*** (0.0112)	-0.0819*** (0.00997)	-0.0601*** (0.00804)	-0.0389*** (0.00698)	-0.0358*** (0.00628)	-0.0385*** (0.00566)
$P(\text{switch})_{\text{before}}$	0.13	0.13	0.13	0.13	0.13	0.13
h	45	60	90	120	150	180
N left	5,971	7,415	11,636	16,904	21,278	25,968
N right	5,876	8,156	12,729	17,020	21,232	26,000

*Notes:* Local linear sharp RDD estimates for different bandwidths and orders of polynomials. Standard errors in parentheses.  $P(\text{switch})_{\text{before}}$  refers to the average probability of moving to a switcher municipality within a range of 60 days before the threshold date for the respective sample. Significance levels: \*  $.05 < p < .1$ , \*\*  $.01 < p < .05$ , \*\*\*  $p < .01$ .

## Temporal development of the effect

As is evident from the graph in Figure 5, the effect fades after some time. To gain an impression of how long the effect persists, we run a linear regression model including indicators for time periods after the break on the pooled data for a time span of half a year around the threshold dates. We estimate a model of the form

$$y_i = \beta_0 + \sum_{k=1}^6 \gamma_k \mathbf{1}(x_{c,i} \geq (k-1) \cdot 30 \quad \& \quad x_{c,i} < k \cdot 30) + \epsilon_i,$$

where  $y$  is our dependent variable, i.e., the residual of a regression of time fixed effects on an indicator set to one if an individual moves to a switcher municipality,  $x_c$  is the moving date centered at the estimated threshold for the corresponding canton, and  $k \cdot 30$  stands for a number series rising in intervals of 30 days:  $\{0, 30, \dots, 180\}$ . The  $\gamma$  coefficients consequently indicate whether the probability of moving to a switcher municipality is different from the level before the break within the specified time span. The estimation results are presented in Table 2. It turns out that the effect becomes weaker over time, as suggested by the graphical evidence. There is no systematic difference in the level anymore about 150 days, or 5 months, after the break. This leveling off could have occurred, because the issue is no longer salient, i.e., expressed attitudes are no longer being discussed in the media. Alternatively, or in some combination, the housing market could have adjusted to a new equilibrium in which foreigners are compensated by, for example, lower rents for their disutility of residing in a municipality in which attitudes towards them are more negative.

**Table 2:** Temporal development of the probability that foreigners move to switcher municipalities - estimates using pooled data

Dependent variable: Moving to a switcher municipality							
Coef.	$\beta_0$	$\gamma_{30}$	$\gamma_{60}$	$\gamma_{90}$	$\gamma_{120}$	$\gamma_{150}$	$\gamma_{180}$
	0.010*** (0.002)	-0.049*** (0.005)	-0.015*** (0.005)	-0.012** (0.005)	-0.025*** (0.005)	-0.011** (0.005)	-0.005 (0.005)
No. of obs.	51,817						
$R^2$	0.0019						
F	16.18						

*Notes:*  $\beta_0$  captures the constant term.  $\gamma$ . are coefficients for deviations from the level before the break for intervals of 30 days. Standard errors in parentheses. Significance levels: \*  $.05 < p < .1$ , \*\*  $.01 < p < .05$ , \*\*\*  $p < .01$ .

## Effects for different groups of foreigners

The vote on the minaret initiative prohibited Muslim communities from constructing further minarets. The change in the substantive law was thus narrow and affected solely one particular group within the foreign population of Switzerland. However, the public discourse was much broader in that it addressed general concerns about migration, as

well as the role of (religious) tolerance and cultural identity in Switzerland. Based on the content of the initiative, we would expect that foreigners, in particular those with a Muslim background, experience lower identity utility if they move to a municipality where the initiative had unexpectedly received wide support. However, due to the general discourse, the outcome of the vote might well have been perceived as a revelation of attitudes towards foreigners in general. According to our understanding of the event, we share such an interpretation. Moreover, identity theories predict that other (foreign) minority groups identify with the beleaguered minority group. Accordingly, similar effects for different groups of foreigners are expected.

In order to learn about any heterogeneity across groups of foreigners, we re-estimate the effect for various subsamples of foreigners. As the results in Table 3 show, we find similar effects for foreigners holding either temporary or a permanent residence permit, for European and non-European foreigners, as well as for foreigners from Muslim countries.<sup>23</sup> This is evidence that the change in perceived attitudes affected the attractiveness of particular municipalities for the group of foreigners across nationalities and cultural backgrounds.

**Table 3:** Reactions in moving behavior for different groups of foreigners - RDD estimates using pooled data for subsamples

Dependent variable: Moving to a switcher municipality						
Sample	(1) Baseline	(2) Temp. residency	(3) Perm. residency	(4) Non-European	(5) European	(6) Muslim
$\tau$	-0.0819*** (0.00997)	-0.0841*** (0.0122)	-0.0761*** (0.0168)	-0.0934*** (0.0231)	-0.0795*** (0.0110)	-0.0577*** (0.0216)
$P(\text{switch})_{\text{before}}$	0.13	0.11	0.15	0.09	0.14	0.1
h	60	60	60	60	60	90
N left	7,415	4,441	2,974	1,072	6,340	1,199
N right	8,156	4,930	3,226	1,170	6,984	1,363

*Notes:* Local linear sharp RDD estimates for different subsamples. Standard errors in parentheses.  $P(\text{switch})_{\text{before}}$  refers to the average probability of moving to a switcher municipality within a range of 60 days before the threshold date for the respective sample. Significance levels: \*  $.05 < p < .1$ , \*\*  $.01 < p < .05$ , \*\*\*  $p < .01$ .

### Effects for foreigners with different occupational skill levels

So far, the evidence suggests that, in general, residents of Switzerland with a foreign nationality expected to experience a loss in identity utility if they were to move to a municipality that unexpectedly revealed increased reservations towards foreigners. Independent of nationality and cultural background, effects might, however, differ depending on individual skill levels. Two arguments are of particular interest with regard to for-

<sup>23</sup>We have no information about an individuals' religious affiliation. Instead, we infer that individuals originating from countries where Islam is the main religion are more likely to be Muslim. The classification on the main religion of countries is based on the *Cross-National Socio-Economic and Religion Data, 2011* and was downloaded from the Association of Religion Data Archives, [www.TheARDA.com](http://www.TheARDA.com).



eigners with low- and high-skill levels, respectively. Regarding the first group, studies investigating the determinants of attitudes towards foreigners frequently argue that concerns about the welfare system play an important role (see, e.g., Hainmueller and Hiscox, 2010; Dustmann and Preston, 2007). Thus, immigrants with a low level of education, who are more likely to rely on the welfare system, might be the main drivers as well as the main targets of negative attitudes towards foreigners. It is, however, open whether they also suffer the highest loss in identity utility when residing in an environment with strong reservations against foreigners. Instead, foreigners with a high education might be particularly sensitive towards such reservations. They might also be more interested in politics and more likely to follow the discussion in the media than their less well educated compatriots.

In our data, we have only limited information about the formal education of the individuals in our sample to test for differential effects across skill levels. Instead, we rely on information about individuals' occupations. As the immigration authorities do not collect this information systematically, it is only available for about 15% of individuals in our data. We infer people's (occupational) skill level by matching occupations to standards of required formal education. It is important to note that immigrants work in occupations that do not match their skill level more often than nationals. Their formal educational degrees might not be accepted, this might be especially the case for, e.g., refugees from non-western countries. Miscategorizations are therefore likely, in particular for people in occupations with formally low skill requirements. To mitigate this miscategorization problem as much as possible we use only individuals originating from countries from which formal education is most likely accepted in Switzerland. These are the countries of the European-Union, countries on the American continent, as the United States, Canada, and Brasil. We further include Australia, and New Zealand. Concentrating on these individuals, we expect that occupation is a good approximation of the skill level.

We use the ISCO-08 classification of the International Labour Office (2012) to classify the skill level needed to perform specific occupations. The four point classification considers not only required formal education, but also the task complexity of the respective occupation. We differentiate between high, upper-medium, medium and low skilled individuals.<sup>24</sup>

The number of observations for the estimates in this particular sample get scarce. Nonetheless we can interpret them qualitatively. We find that the high-skilled group is most sensitive to the revelation of new information about citizens' attitudes towards foreigners. The

---

<sup>24</sup>We code somebody as high-skilled if his or her occupation corresponds to an ISCO skill level of 4, i.e., the second stage of tertiary education or first stage of tertiary education (medium duration). Upper-medium skilled, i.e., ISCO skill level of 3 and thus the first stage of tertiary education (short or medium duration). Medium skilled people refer to an ISCO skill level of 2, i.e., lower to post-secondary education. Foreigners considered low-skilled are in an occupation of ISCO skill level 1 which requires primary education only.

**Table 4:** Reactions in moving behavior for foreigners with different occupational skill levels - RDD estimates using pooled data

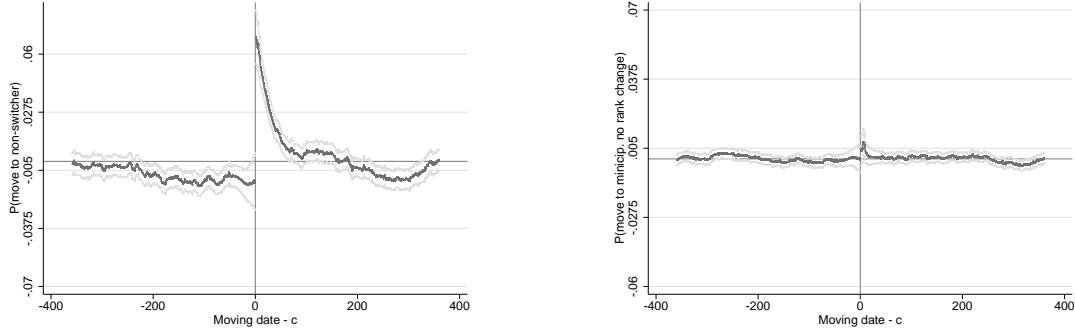
Dependent variable: Moving to a switcher municipality												
Sample	(1) High skill	(2) High skill	(3) High skill	(4) Up-med. skill	(5) Up-med. skill	(6) Up-med. skill	(7) Med. skill	(8) Med. skill	(9) Med. skill	(10) Low skill	(11) Low skill	(12) Low skill
$\tau$	-0.203** (0.0803)	-0.151** (0.0658)	-0.155*** (0.0595)	-0.104 (0.0716)	-0.0643 (0.0940)	-0.0324 (0.0955)	-0.109 (0.0931)	-0.00349 (0.0789)	0.0353 (0.0662)	0.0600 (0.116)	0.00904 (0.0968)	0.0233 (0.0744)
$P(\text{switch})_{\text{before}}$	0.09	0.09	0.09	0.12	0.12	0.12	0.10	0.10	0.10	0.08	0.08	0.08
h	60	90	120	60	90	120	60	90	120	60	90	120
N left	77	112	156	33	55	69	70	124	181	50	83	125
N right	65	108	140	31	45	62	82	124	164	51	85	120

*Notes:* Local linear sharp RDD estimates for different subsamples. Standard errors in parentheses.  $P(\text{switch})_{\text{before}}$  refers to the average probability of moving to a switcher municipality within a range of 60 days before the threshold date for the respective sample. Significance levels: \*  $.05 < p < .1$ , \*\*  $.01 < p < .05$ , \*\*\*  $p < .01$ .

estimate is significantly different from zero despite the few observations. For the other groups, the discontinuity estimates seem to fall the lower the skill level is. However, the estimates for these other groups are not significantly different from zero at conventional levels. Whatever the cause, the results suggest that the group of immigrants that triggers most of the resentments is not the group that suffers the most from them. It is, instead, rather the high-skilled immigrants who seem to react the most.

### 6.3 Validation check for the relevance of the change in the relative political positioning of the switcher municipalities

The definition of the switcher municipalities results in a dichotomous categorization of municipalities. Accordingly, we see per construction an equivalent positive jump in the probability of moving to a non-switcher municipality at the threshold. The corresponding pattern is presented in Figure 6a. RDD estimation results are reported in Table 5 in column 1, 2, and 3. Moreover the classification itself rests on a central argument emphasizing that the voters in a switcher municipality not only expressed a more restrictive position towards migration than in the past, but that the swing in expressed reservations is also more pronounced than for the average municipality of a canton (second condition of the definition for switcher municipalities outlined in Subsection 2.2). In order to validate the importance of this rank condition, we re-estimate the basic specifications with the pooled data for an alternative classification of municipalities. We construct the dependent variable such that it captures the probability of moving to a municipality that has not changed its rank, i.e., its position regarding its support of critical attitudes towards foreigners compared to the other municipalities in the canton. Specifically, we focus on municipalities that are not switcher municipalities and have not experienced a



(a) Probability of moving to a non-switcher municipality

(b) Probability of moving to a municipality with no rank change

**Figure 6:** Probability of moving to alternative groups of municipalities. Local linear smooth of the probability that a foreigner moves to a specific group of municipalities separately from both sides of the threshold, and using a bandwidth of 45 days.

rank change. If our findings are driven by a change in the positioning, we would not expect to see a pronounced jump, either positive or negative, in the probability of moving to these municipalities at the estimated thresholds. The resulting pattern is presented graphically in Figure 6b. The results for the RDD treatment effect estimates are reported in Table 5 in columns 4, 5, and 6. We do not find a systematic reaction for this category of municipalities.

In an additional test, we analyze whether the positive jump within the group of non-switcher municipalities is consistent with our reasoning. We define two further alternative classes of municipalities, one including non-switcher municipalities with a positive rank change, and thus a relative right shift, after the vote on the minaret initiative and one including non-switcher municipalities with a negative rank change, and thus a relative left shift. Re-estimating the jump in the probability to move to one of these municipality types at the threshold, we find that the positive jump in the non-switcher municipalities seems to be purely driven by municipalities which reduced the rank. The estimation results can be found in Table 5, columns 7 to 12. This finding is consistent with our reasoning that foreigners are deterred from moving to municipalities revealing reservations, and rather choose to move to municipalities that got relatively more attractive in this dimension.

We interpret this as further support that the observed negative jump in the probability of choosing a switcher municipality is driven by a mover’s perception of an unexpected negative shift in the social attitudes of these municipalities towards foreigners.

## 6.4 Test of an alternative supply side explanation

So far, the empirical regularity is interpreted within the theoretical framework based on demand-side reactions to perceived negative attitudes. However, the observed patterns in

**Table 5:** Change in the probability that foreigners move to different samples of non-switcher municipalities - RDD estimates using pooled data

Dependent variable: Moving to a particular type of non-switcher municipality						
Dep. Var.	(1) all	(2) all	(3) all	(4) $\Delta rank = 0$	(5) $\Delta rank = 0$	(6) $\Delta rank = 0$
$\tau$	0.0824*** (0.0112)	0.0819*** (0.00997)	0.0601*** (0.00804)	0.00431 (0.00504)	-4.22e-05 (0.00443)	-0.00159 (0.00354)
$P(minor)_{before}$	0.87	0.87	0.87	0.02	0.02	0.02
h	45	60	90	45	60	90
N left	5,971	7,415	11,636	5,971	7,415	11,636
N right	5,876	8,156	12,729	5,876	8,156	12,729
Dep. Var.	(7) $\Delta rank > 0$	(8) $\Delta rank > 0$	(9) $\Delta rank > 0$	(10) $\Delta rank < 0$	(11) $\Delta rank < 0$	(12) $\Delta rank < 0$
$\tau$	-0.00190 (0.0123)	0.0191* (0.0107)	0.0128 (0.00842)	0.0800*** (0.0158)	0.0628*** (0.0138)	0.0488*** (0.0109)
$P(minor)_{before}$	0.11	0.11	0.11	0.74	0.74	0.74
h	45	60	90	45	60	90
N left	5,971	7,415	11,636	5,971	7,415	11,636
N right	5,876	8,156	12,729	5,876	8,156	12,729

*Notes:* Local polynomial sharp RDD estimates. Standard errors in parentheses.  $P(minor)_{before}$  refers to the average probability of moving to a municipality of the respective type within a range of 60 days before the threshold date. Significance levels: \*  $.05 < p < .1$ , \*\*  $.01 < p < .05$ , \*\*\*  $p < .01$ .

aggregate moving behavior across municipalities would also be consistent with a supply-side reaction. Native landlords in switcher municipalities might feel free to discriminate against foreigners, once they are informed about their co-residents' attitudes towards immigrants. We perform a supplementary test to address the alternative supply-side explanations.

If the effect is supply-driven and native landlords become less willing to rent their property to foreigners, they are expected to discriminate against foreigners in general. When considering second-generation immigrants who hold the same foreign names, the same potentially foreign appearance, and the same residence permit types as their parents, one would thus expect landlords to discriminate against this group in the same way that they discriminate against the first generation immigrants. However, if the phenomenon is, instead, demand-driven, a smaller reaction, is expected for the second-generation immigrants, as they are in general better assimilated. They grew up in the Swiss institutional environment, for example, being familiar with referendums on migration issues. We therefore repeat our estimates for the group of second-generation foreigners, i.e., individuals who were born in Switzerland but have not been naturalized. Table 6 reports the estimates of the treatment effect. Figure 7 presents the evidence graphically. We find a much smaller reaction for second generation foreigners, whereby no statistically significant

treatment effect is estimated. We interpret this finding as evidence that the main effect is primarily demand-driven.

**Table 6:** Reaction in moving behavior for second-generation immigrants - RDD estimates using pooled data

Dependent variable: Moving to a switcher municipality					
	(1)	(2)	(3)	(4)	(5)
$\tau$	-0.0174 (0.0359)	-0.00497 (0.0280)	0.00188 (0.0240)	0.00144 (0.0214)	-0.00911 (0.0195)
$P(\text{switch})_{\text{before}}$	0.16	0.16	0.16	0.16	0.16
h	60	90	120	150	180
N left	882	1,342	1,925	2,387	2,921
N right	927	1,436	1,919	2,377	2,845

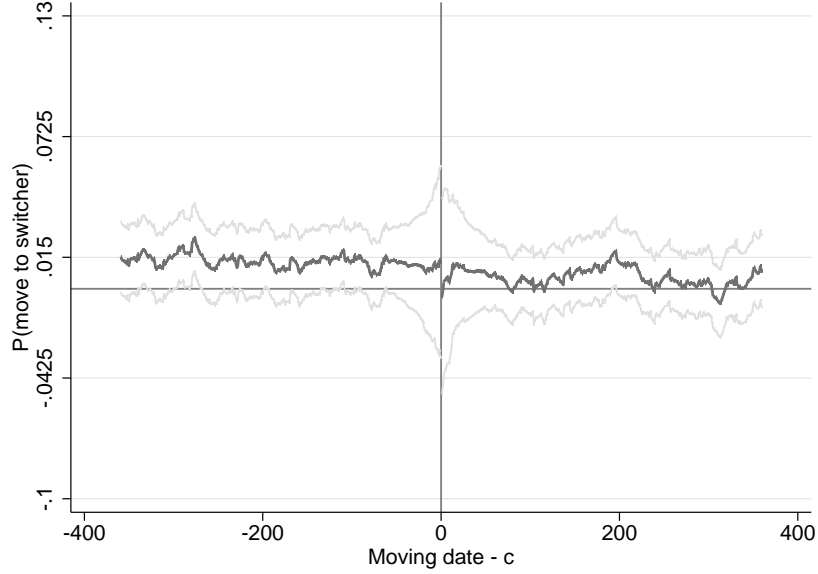
*Notes:* Local linear sharp RDD estimates.  $P(\text{switch})_{\text{before}}$  refers to the average probability of moving to a switcher municipality within a range of 60 days before the threshold date for the respective sample. Significance levels: \*  $.05 < p < .1$ , \*\*  $.01 < p < .05$ , \*\*\*  $p < .01$ .

## 6.5 Validation of the identification assumptions

There are some crucial assumptions that need to hold for a valid RDD setting.

First, local randomization, i.e., individuals deciding to move just before and just after the vote are, on average, the same. There is no formal test for this assumption and we do not observe the date of the moving decision. However, most RDD studies check whether individual characteristics are balanced around the threshold to validate it. We follow this strategy and estimate a sharp RDD, taking some individual characteristics, which are available in our data, as dependent variables, and using the estimated threshold date as threshold value. We report the results of this exercise in Table 7. We find only one significant estimate for Age and a bandwidth of 45 days. As it is not robust to the bandwidth choice and seems to be rather small we are not alarmed that our results might be driven by this difference. We do not find systematic differences in the other individual characteristics around the threshold.

Second, the characteristics of the switcher municipalities have not changed systematically after the vote in 2009 with two exceptions. Table 8 reports the results of the corresponding validation estimates. Regarding the two systematic changes: on the one hand, there is a decrease in the resident population, and on the other hand, there is a slight increase in the rate of vacant apartments. Both findings are rather complementary evidence in support of our main hypothesis, as they are consistent with the observed individual reactions of foreigners. It can be expected that there will be, at least temporarily, more vacant apartments if a sizable group of the population shuns switcher municipalities.



**Figure 7:** Probability of moving to a switcher municipality for second-generation immigrants. Local linear smooth of the probability that a second-generation immigrant moves to a switcher municipality separately from both sides of the threshold, and using a bandwidth of 90 days.

**Table 7:** Individual characteristics of movers around the threshold dates

Dep. Var.	(1) Age	(2) Age	(3) P(married)	(4) P(married)	(5) Duration CH	(6) Duration CH	(7) P(par. CH)	(8) P(par. CH)
$\tau$	-1.238*** (0.443)	-0.549 (0.382)	-0.0173 (0.0190)	-0.0178 (0.0165)	0.000222 (0.381)	0.464 (0.328)	0.000141 (0.00226)	-3.18e-05 (0.00205)
h	45	60	45	60	45	60	45	60
N left	5,971	7,415	5,971	7,415	5,971	7,415	5,971	7,415
N right	5,876	8,156	5,876	8,156	5,876	8,156	5,876	8,156

*Notes:* Local linear sharp RDD estimates for different dependent variables. Standard errors in parentheses. Dependent variables: *Age* of the individual, *P(married)* an indicator of whether a person is married, *Duration CH* the duration of stay in Switzerland, and *P(par. CH)* an indicator of whether a parent of an individual is Swiss. Significance levels: \*  $.05 < p < .1$ , \*\*  $.01 < p < .05$ , \*\*\*  $p < .01$ .

Third, there is no change in the probability of moving due to the event. It is important that this condition holds, for example, because a large increase in the prevalence of moving within the group of non-switcher municipalities would be reflected in a relative decline in the probability of moving to a switcher municipality. In order to investigate how many foreigners move per day around the threshold, we generate counts of movers within each canton and moving date. This approach requires that we control for periodical patterns before running the analysis. There are dates at which it is much more likely to observe individuals moving due to the municipalities' mandatory moving dates and contractual arrangements on tenancy changeover. We again follow, for example, Davis (2008) in controlling for these periodical patterns by using a battery of indicator variables. We first

**Table 8:** Municipality characteristics before and after the vote on the minaret initiative

Dep. Var.	<i>Tax rate</i>	<i>Pop</i>	<i>Pop_welfare</i>	<i>Pop_poor</i>	<i>Pop_rich</i>	<i>Vacant</i>	<i>N_comp</i>	<i>Natur</i>
<i>Time</i>	-0.286*** (0.058)	32.551*** (7.971)	0.014 (0.031)	0.000 (0.000)	0.010*** (0.001)	-0.056 (0.081)	-0.056 (0.144)	0.075 (0.061)
$\mathbb{1}(\text{year} \geq 2009)$	0.058 (0.140)	18.495*** (4.774)	-0.070 (0.079)	-0.001* (0.000)	-0.003** (0.001)	-0.196 (0.116)	-0.290 (0.356)	-0.237 (0.191)
<i>Switcher municipality</i>	0.044 (0.083)	-1424.448*** (326.222)	-0.279** (0.104)	-0.001 (0.000)	0.026*** (0.007)	-0.269 (0.293)	-4.356*** (1.425)	-0.018 (0.017)
<i>Switcher</i> * $\mathbb{1}(\text{year} \geq 2009)$	-0.035 (0.048)	-38.086** (15.128)	0.008 (0.078)	0.000 (0.000)	-0.001 (0.001)	0.499* (0.243)	0.311 (0.255)	0.035 (0.031)
No. of obs.	9,404	9,404	6,810	10,028	10,028	9,403	8,376	9,355
No. of clusters	24	24	25	25	25	24	24	24
$R^2$	0.78	0.07	0.13	0.14	0.37	0.07	0.05	0.00

*Notes:* OLS estimations for annual data between 2007 and 2010, including cantonal fixed effects. Standard errors are adjusted for clustering at the cantonal level. Significance levels: \* .05 < p < .1, \*\* .01 < p < .05, \*\*\* p < .01. Dependent variables: *Tax rate* is the municipality tax rate, *Pop* is the municipality's population, *Pop\_welfare* is the share of the population that depends on welfare, *Pop\_poor* is the proportion of the population classified as poor, *Pop\_rich* is the fraction of the population that is classified as rich, *Vacant* is the number of vacant apartments per 1,000 inhabitants, *N\_comp* measures how many companies have been incorporated, and *Natur* measures the naturalization rate of foreigners in the municipality. Differences in sample sizes occur, because low number of cases in a municipality are not reported in the official statistics for reasons of privacy.

run a linear regression of the dependent variable on covariates, and, subsequently, the residuals of this regression are used as the dependent variable in our analysis (see, e.g., Lee and Lemieux, 2010).<sup>25</sup> For convenience, we still refer to this variable as the number of moving individuals. We first estimate a sharp RDD on the number of foreigners moving in general. The corresponding RDD graph is presented in Figure B.8a. The results for two RDD estimates are reported in columns 1 and 2 of Table 9. We find no significant change in the number of movers. Second, we repeat the exercise using counts of individuals leaving switcher municipalities. The graphical evidence is presented in Figure B.8b, and the corresponding estimates are reported in columns 3 and 4 of Table 9. As before, we do not find a systematic change in the probability that a foreigner moves from this group of municipalities. The latter finding could also be interpreted as evidence that the interaction between foreigners and citizens in switcher municipalities has not radically changed after the vote. If it had, we would expect to see foreigners leaving these municipalities. Third, we check whether the aggregate data still allows us to detect the pattern we found in the individual data. We repeat the estimation using the residual counts of individuals moving to a switcher municipality. As is evident from Figure B.9 and the estimates in columns 5 and 6 of Table 9, a drop in the probability of moving to a switcher municipality is still observed. Foreigners shun switcher municipalities after the vote.

<sup>25</sup>We first run the specification

$$y_t = \beta_0 + \beta_1 x_t + u_t,$$

where the dependent variable is  $y_t$ , the number of moving individuals at a certain date, and  $x_t$  includes indicator variables for months, the day of the month and weekdays. We estimate over the period between 2006 and 2008, thus before the event, in order to control for moving behavior in normal times. We then predict the residual  $u_t$  for the entire time period and use it as our dependent variable.

**Table 9:** Number of moving individuals around the threshold dates

Dep. Var.	(1) $N_{move}$	(2) $N_{move}$	(3) $N_{from\_switch}$	(4) $N_{from\_switch}$	(5) $N_{to\_switch}$	(6) $N_{to\_switch}$
$\tau$	-0.0476 (4.393)	-0.609 (4.310)	-0.250 (0.423)	-0.103 (0.380)	-1.348*** (0.461)	-0.989** (0.389)
h	60	90	60	90	60	90
N left	501	753	501	753	501	753
N right	492	738	492	738	492	738

*Notes:* Local linear sharp RDD estimates. Standard errors in parentheses. Dependent variables:  $N_{move}$  corresponds to the number of foreigners moving.  $N_{from\_switch}$  refers to the foreigners leaving switcher municipalities, and  $N_{to\_switch}$  refers to foreigners relocating to switcher municipalities. Significance levels: \*  $.05 < p < .1$ , \*\*  $.01 < p < .05$ , \*\*\*  $p < .01$ .

Finally, as in many RDD studies, we run a placebo test by repeating the test procedure for the same time periods in the previous year (-365 days). We cannot reject the null hypothesis of no treatment effect in either of the cantons. Results for this placebo test are reported in Table B.5.

## 7 Concluding remarks

This paper uses the vote on the Swiss minaret initiative in 2009 as a natural experiment to identify the effect of reservations towards immigrants on their location choice and thus indirectly on the utility they derive from residing in a particular municipality. Using the quasi-random variation in the exact date of the decision to move of individuals and applying an RDD with unknown discontinuity points, we are able to detect the point in time after which the effect materializes and to estimate the reaction to the treatment at this point in time. We find discontinuous drops in the probability that a foreigner moves to a municipality which has unexpectedly been revealed to have reservations about immigrants in 12 out of 22 cantons. The estimated reaction is sizable. The probability of an immigrant relocating to a switcher municipality drops, on average, by about 60 percent. This initial effect levels off over a period of 5 months before it vanishes.

The approval of the amendment involved no real, immediate consequences regarding the Muslims' ability to practice their religion. However, the symbolic restriction of prohibiting the construction of new minarets was heavily discussed and interpreted as a signal of limited openness towards foreigners. Accordingly, we consider an expected loss of identity utility due to perceived exclusion to act as a mediator in the observed reaction to the vote outcome. This interpretation is supported by the finding that the reaction is not confined to Muslims or non-European immigrants. Non-Muslims are also found to react strongly to the vote outcome. This finding is in line with research on ostracism and the rejection-identification model in social psychology which asserts that minority groups tend



to identify with other minority groups when the latter are under threat and that even a vague feeling of being ostracized can produce psychological distress.

Moreover, we find that second-generation immigrants who are often well assimilated react much less strongly, if at all, to the revealed attitudes in the minaret vote. This suggests that the observed moving pattern after the vote is not driven by landlords who learn about negative attitudes in their surrounding and feel free to discriminate against foreigners. In contrast, foreigners in high-skill occupations react much more strongly than foreigners in medium- and low-skill occupations. Well-educated immigrants therefore seem to be particularly sensitive to reservations towards foreigners. This observation highlights a tension when countries try to attract highly-skilled professionals from abroad.

Overall, a location choice model that takes identity utility into account is well suited to predict and reconcile the observed breaks in foreigners' relocating patterns between municipalities after the vote on the minaret initiative. The analysis reveals that the political attitudes of prospective co-residents are potentially an important determinant of location choices for some groups in the population. This holds, in particular, for foreigners who in many regions of the world experience reservations or even outright rejection. Thus, we do not think that the findings reported in this analysis are confined to immigrants in Switzerland. Switzerland is rather exceptionally well suited to the investigation of this phenomenon. Switzerland's direct democratic institutions allow researchers to observe attitudes as revealed in popular votes with a high spatial resolution.

## References

- Akerlof, G. and R. Kranton (2000). Economics and Identity. *Quarterly Journal of Economics* 115(3), 715–753.
- Algan, Y., A. Bisin, A. Manning, and T. Verdier (2012). *Cultural Integration of Immigrants in Europe*. Oxford: Oxford University Press.
- Barone, G., A. D’Ignazio, G. d. Blasio, and P. Naticchioni (2014). Mr. Rossi, Mr. Hu and Politics: The Role of Immigration in Shaping Natives’ Political Preferences. *IZA Discussion Paper No. 8228*.
- Branscombe, N., M. T. Schmitt, and R. Harvey (1999). Perceiving Pervasive Discrimination Among African Americans: Implications from Group Identification and Well-Being. *Journal of Personality and Social Psychology* 77(1), 135–149.
- Cameron, A. and P. Trivedi (2005). *Microeconometrics: Methods and Applications*. Cambridge and UK: Cambridge University Press.
- Card, D. (2001). Immigrant Inflows, Native Outflows, and the Local Market Impacts of Higher Immigration. *Journal of Labor Economics* 19(1), 22–64.
- Card, D., A. Mas, and J. Rothstein (2008). Tipping and the Dynamics of Segregation. *Quarterly Journal of Economics* 123(1), 177–218.
- Casey, T. and C. Dustmann (2010). Immigrants’ Identity, Economic Outcomes and the Transmission of Identity Across Generations. *Economic Journal* 120(542), F31–F51.
- Damm, A. P. (2009). Determinants of Recent Immigrants’ Location Choices: Quasi-Experimental Evidence. *Journal of Population Economics* 22(1), 145–174.
- Davis, L. W. (2008). The Effect of Driving Restrictions on Air Quality in Mexico City. *Journal of Political Economy* 116(1), 38–81.
- Dill, V. (2013). Ethnic Concentration and Extreme Right-Wing Voting Behavior in West Germany. *SOEPpapers on Multidisciplinary Panel Data Research* 565.
- Dustmann, C. (1996). The Social Assimilation of Immigrants. *Journal of Population Economics* 9(1), 37–54.
- Dustmann, C. and I. Preston (2001). Attitudes to Ethnic Minorities, Ethnic Context and Location Decisions. *Economic Journal* 111(470), 353–373.
- Dustmann, C. and I. Preston (2007). Racial and Economic Factors in Attitudes to Immigration. *The B.E. Journal of Economic Analysis & Policy* 7(1).

- Frey, B. S. (1994). Direct Democracy: Politico-Economic Lessons from Swiss Experience. *The American Economic Review* 84(2), 338–342.
- Georgiadis, A. and A. Manning (2013). One Nation Under a Groove? Understanding National Identity. *Journal of Economic Behavior & Organization* 93, 166–185.
- Gerdes, C. and E. Wadensjö (2008). The Impact of Immigration on Election Outcomes in Danish Municipalities. *IZA Discussion Paper No. 3586*.
- Gorinas, C. and M. Pytliková (2013). *Essays on Marginalization and Integration of Immigrants and Young Criminals – A Labour Economics Perspective: Do Attitudes Toward Immigrants Influence International Migration?* Ph. D. thesis, University, Aarhus.
- Hahn, J., P. Todd, and W. van der Klaauw (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica* 69(1), 201–209.
- Hainmueller, J. and M. J. Hiscox (2010). Attitudes Toward Highly Skilled and Low-skilled Immigration: Evidence from a Survey Experiment. *American Political Science Review* 104(1).
- Hainmueller, J. and D. J. Hopkins (2014). Public Attitudes Toward Immigration. *Annual Review of Political Sciences* 17, 225–249.
- Halla, M., A. Wagner, and J. Zweimüller (2014). Immigration and Voting for the Extreme Right. *Mimeo, Department of Public Finance, University of Innsbruck*.
- Henry, R. (2009). Does Racism Affect a Migrant’s Choice of Destination? *IZA Discussion Paper No. 4349*.
- Imbens, G. and T. Lemieux (2008). Regression Discontinuity Designs: A Guide to Practice. *Journal of Econometrics* 142(2), 615–635.
- International Labour Office (2012). ISCO-08. *International Standard Classification of Occupations: Structure, Group Definitions and Correspondence Tables 1*.
- Knabe, A., S. Rätzl, and S. L. Thomsen (2013). Right-Wing Extremism and the Well-Being of Immigrants. *Kyklos* 66(4), 567–590.
- Kuhn, A. and B. Brunner (2014). Immigration, Cultural Distance and Natives’ Attitudes Towards Immigrants: Evidence from Swiss Voting Results. *IZA Discussion Paper No. 8409*.
- Lee, D. and T. Lemieux (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature* 48(2), 281–355.

- Linder, W. (2010). *Swiss Democracy: Possible Solutions to Conflict in Multicultural Societies*. Basingstoke: Palgrave Macmillan.
- Liu, R. (1988). Bootstrap Procedures under Some Non-I.I.D. Models. *The Annals of Statistics* 16(4), 1696–1708.
- Mechtel, M., F. Hett, and M. Kröll (2014). Endogenous Social Identity and Group Choice. *Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014 No. E16-V2*.
- Méndez Martínez, I. and I. M. Cutillas (2014). Has Immigration Affected Spanish Presidential Elections Results? *Journal of Population Economics* 27(1), 135–171.
- Paola, M. d., V. Scoppa, and M. Falcone (2012). The Deterrent Effects of the Penalty Points System for Driving Offences: A Regression Discontinuity Approach. *Empirical Economics* 45(2), 965–985.
- Parchet, R. (2014). Are Local Tax Rates Strategic Complements or Strategic Substitutes? *IdEP Economic Papers, University of Lugano* (2014/07).
- Porter, J. (2003). Estimation in the Regression Discontinuity Model. *Mimeo, Department of Economics, Harvard University*.
- Porter, J. and P. Yu (2015). Regression Discontinuity Design with Unknown Discontinuity Points: Testing and Estimation. *Journal of Econometrics* forthcoming.
- Rudert, S., Janke S., and Greifeneder R. (2014). Under Threat by Popular Vote: Naturalistic Social Exclusion Due to the Swiss Vote Against Mass Immigration. *Mimeo, University of Basel*.
- Schelling, T. C. (1971). Dynamic Models of Segregation. *Journal of Mathematical Sociology* 1(2), 143–186.
- Schmitt, M. T., R. Spears, and N. R. Branscombe (2003). Constructing a Minority Group Identity out of Shared Rejection: The Case of International Students. *European Journal of Social Psychology* 33(1), 1–12.
- Tajfel, H. (1978). *The Social Psychology of Minorities*. London: Minority Rights Group.
- Thistlethwaite, D. and D. Campbell (1960). Regression-Discontinuity Analysis: An Alternative to the Ex Post Facto Experiment. *Journal of Educational Psychology* 51(6), 309.
- Tolnay, S. E. and E. M. Beck (1992). Racial Violence and Black Migration in the American South, 1910 to 1930. *American Sociological Review* 57(1), 103–116.
- Waisman, G. and B. Larsen (2008). Do Attitudes Towards Immigrants Matter? *The Stockholm University Working paper 2008:5*.

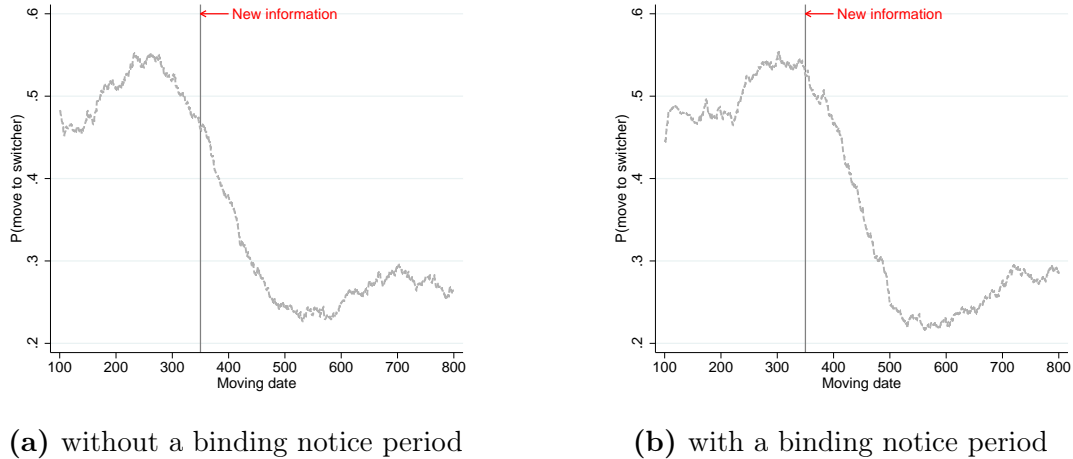
- Wesselmann, E. D., D. Bagg, and K. D. Williams (2009). "I Feel Your Pain": The Effects of Observing Ostracism on the Ostracism Detection System. *Journal of Experimental Social Psychology* 45(6), 1308–1311.
- Williams, K. (2009). *Advances in Experimental Social Psychology: Ostracism: A Temporal Need-Threat Model*, Volume 41. San Diego: Academic Press.
- Williams, K. D. and S. A. Nida (2011). Ostracism: Consequences and Coping. *Current Directions in Psychological Science* 20(2), 71–75.
- Wu, C. (1986). Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis. *The Annals of Statistics* 14(4), 1261–1295.

## A Simulation of a simple location search with overlapping search generations

How does the inflow of people with particular preferences into a municipality change when the municipality becomes less attractive for them? An *ex ante* understanding is necessary in order to choose appropriate econometric techniques when analyzing data about moving behavior to test hypotheses about location choices. This also holds in the context of our analysis concerning the effect of the revelation of negative attitudes towards foreigners on immigrants' location choices.

Whereas our theoretical model is about location decisions, our empirical data is restricted to information about moving dates. It is thus *ex ante* unclear how the incorporation of the new information about attitudes towards foreigners is reflected in moving patterns after the vote on the minaret initiative. As it can be considered random whether an individual makes his or her location choice immediately before or after the polling day, a regression discontinuity design could be applied regarding the decision to locate in a switcher municipality with the polling day as the threshold date. However, actual moves materialize only after a decision has been made, and we do not know when people made their location decisions. We therefore have to decide whether we would predict that an unexpected negative change in a municipality's attitudes towards foreigners (switcher municipality) would elicit a sharp drop in the probability of foreigners choosing to move to that municipalities or whether we, instead, predict that a gradual transition to a new equilibrium takes place.

To asses how the effect might materialize, we perform a simulation analysis of individual location choices. We simulate a simple location search with overlapping search generations to evaluate which incorporation pattern we would expect. We assume as given that only those individuals who search a location after the vote can incorporate the information about revealed attitudes into their choice. We generate a population of 3,500 observations and randomly assign a point in time at which they decide to relocate. We assume that individuals terminate the contract with their landlord as soon as they have decided to move and start to search for an alternative apartment. The probability of a searcher finding an apartment on each successive day after starting their search is set to 0,015, such that the average search length is 66,66 days. In the baseline setting, the probability that the apartment lies in a switcher municipality is 0.5. The new information enters the system at time  $t_0 = 350$  and reduces the probability of choosing a switcher municipality by  $\tau = -0.25$ . We simulate the location search, once with search frictions alone, i.e., individuals move at the point in time in which they find an apartment, and once considering a notice period of 3 months (or 90 days) assuming that individuals cannot move before 90 days, after starting their search even if they find housing before this limit. The two scenarios mark the extreme cases in terms of the transaction costs related to moving. While these

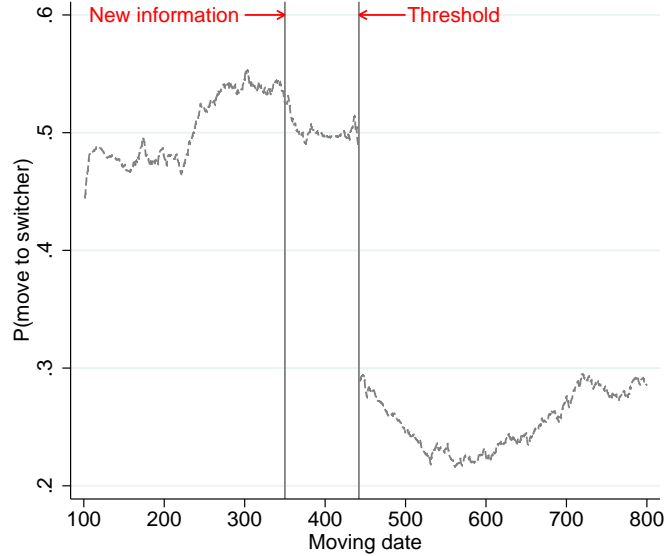


**Figure A.1:** Simulated probability of moving to a switcher municipality. Local linear smooth of the probability that an individual chooses to move to a switcher municipality over the moving date using a bandwidth of 60 days.

costs are assumed minimal in the first scenario, they are assumed to be maximal in the second. Movers cannot afford to rent two apartments at the same time and do not find follow-up tenants. Based on the two scenarios, we compare how the new information is incorporated and reflected in the pattern of the probability of moving to a switcher municipality over time.

Figure A.1a visualizes the resulting pattern without a notice period, estimated by a local linear smooth (LLS). The straight line indicates the point in time in which the new information enters the system. The effect seems to be incorporated gradually, and it takes some time until the probability of choosing a switcher municipality stabilizes at its new equilibrium. Figure A.1b shows the resulting picture if a notice period is taken into account. It again takes some time until the new equilibrium is reached. However, the new information is incorporated with a time lag and the pattern reveals a much steeper decline in the probability of moving to a switcher municipality. The effect of the information is seen to cause a sharp drop in the probability of moving to a switcher municipality, rather than a gentle decline as in the case when the information is gradually incorporated. As notice periods are very common in Switzerland (and the default option in rental agreements), we think that the second pattern of information incorporation is realistic with regard to our application, and in the choice of econometric techniques for the empirical analysis.

Assuming that it is random when individuals choose to relocate and find an apartment, a regression discontinuity framework is suitable to analyze the incorporation of the new information. The probability of moving to a switcher municipality seems to drop sharply once the new information takes effect on individuals' relocation behavior. This sharp drop, if it really exists, can be measured by estimating a jump or a discontinuity. However, the

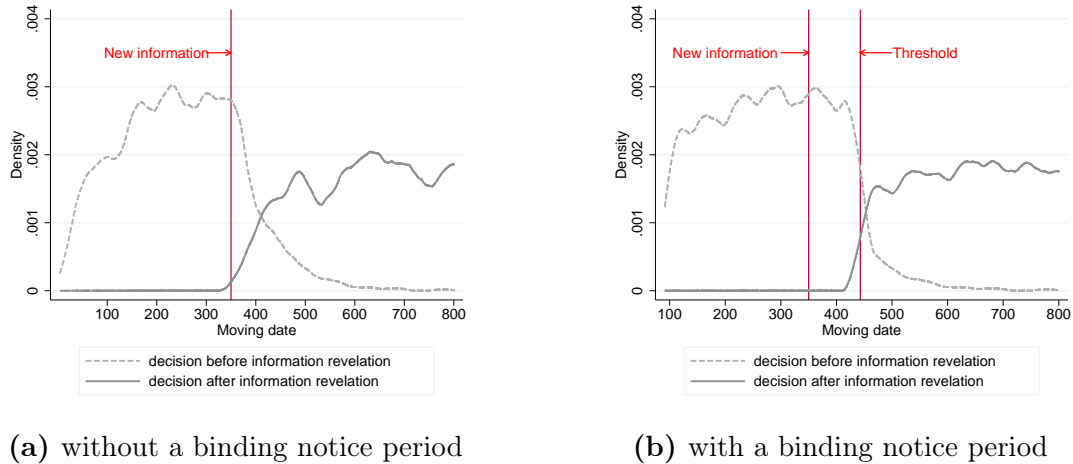


**Figure A.2:** Simulated probability of moving to a switcher municipality (with a binding notice period). Local linear smooth of the probability that an individual chooses to move to a switcher municipality over the moving date separately from both sides of the estimated threshold date, and using a bandwidth of 60 days.

exact date when the jump materializes is not known. We, therefore, apply a regression discontinuity design with unknown discontinuity points to test whether the relationship features a discontinuity and to determine its position. This method is discussed in detail in Section 4 on the empirical strategy. For the particular run of the simulation, there is a discontinuity at  $t = 442$  for the notice period scenario.

Re-estimating the relationship using LLS separately from both sides of the estimated threshold, we find in Figure A.2 that the probability of choosing a switcher municipality features a clear negative jump at the threshold value. To gain an idea why the jump mostly appears when a notice period is considered, we decompose the distributions of the moving dates. We separate between individuals who decided to relocate before the new information entered the system and those who decided to relocate afterwards. Figure A.3b shows the resulting densities for the scenario considering a notice period. We see that the estimated threshold date is temporally close to the point in time at which the movers who decided on the basis of the new information outweigh those who decided without it. Furthermore, the densities show steep slopes just before the threshold date. Thus, the change in the composition of the population regarding individuals with and without the new information occurs quite abruptly. The transition period in the baseline scenario, see Figure A.3a, is wider. The densities have a less steep slope in the period after the new information is released. This explains why the effect materializes much more smoothly than in the scenario with the notice period in most simulation runs.





**Figure A.3:** Composition of the pool of individuals moving regarding their information level. Kernel density estimates.

In a last step, we repeat the procedure of drawing a population and testing for the presence of a jump in the resulting moving pattern 500 times. Table A.1 summarizes the results of this simulation. We find a discontinuity in 82 percent of the draws in the notice period setting. In contrast, it is only found in about 26 percent of cases in the setting with search frictions alone.

**Table A.1:** Test results for discontinuities detected in simulated data

	Mean	SD	Min.	Max.	N
Without notice period					
Discontinuity detected	0.26	0.44	0	1	500
Threshold	399.04	31.01	360	499	128
$\tau$ at threshold	-0.16	0.11	-0.29	0.26	128
With notice period					
Discontinuity detected	0.82	0.38	0	1	500
Threshold	440.84	13.72	360	495	412
$\tau$ at threshold	-0.23	0.08	-0.42	0.36	412

*Note:* Summary of simulation results, once without notice period and once considering a notice period of 90 days.

The regression discontinuity method with unknown discontinuity points thus seems an appropriate econometric technique for analyzing information effects on moving patterns in a context with notice periods.

# B Results



Figure B.4: Campaign poster in support of the minaret initiative.

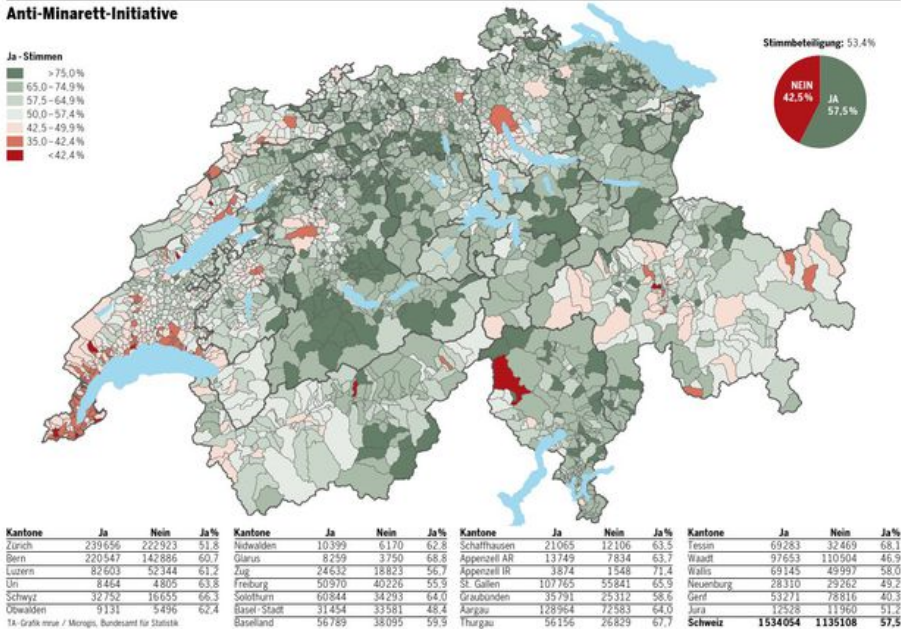
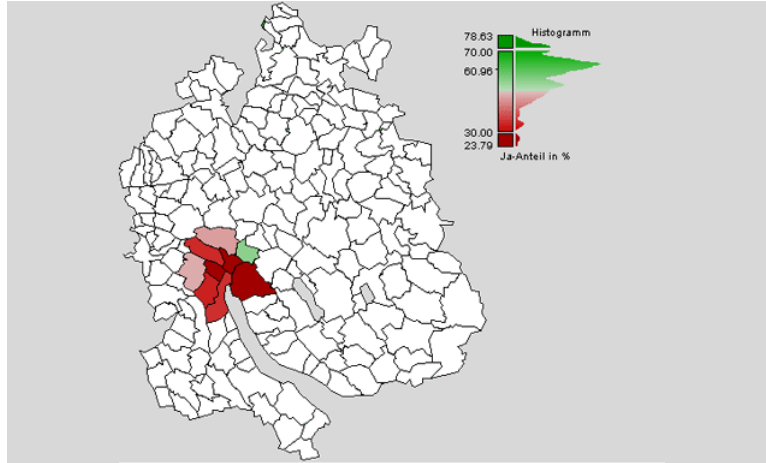


Figure B.5: Local media reporting about the vote on the minaret initiative. The excerpt of an article from the Swiss newspaper *Tages-Anzeiger* entitled "Land sagte Ja, Städte Nein" shows how the outcome of the vote on the minaret initiative was visualized in the press.



**Figure B.6:** Local media reporting about the vote outcomes at the voting district level. The graph was presented in the Swiss newspaper *Tages-Anzeiger* in the article entitled "Warum die Schwamendinger anders ticken als der Durchschnittszürcher", it depicts the outcome of the vote on the minaret initiative at the level of voting districts within Zurich.

**Table B.2:** Descriptive Statistics

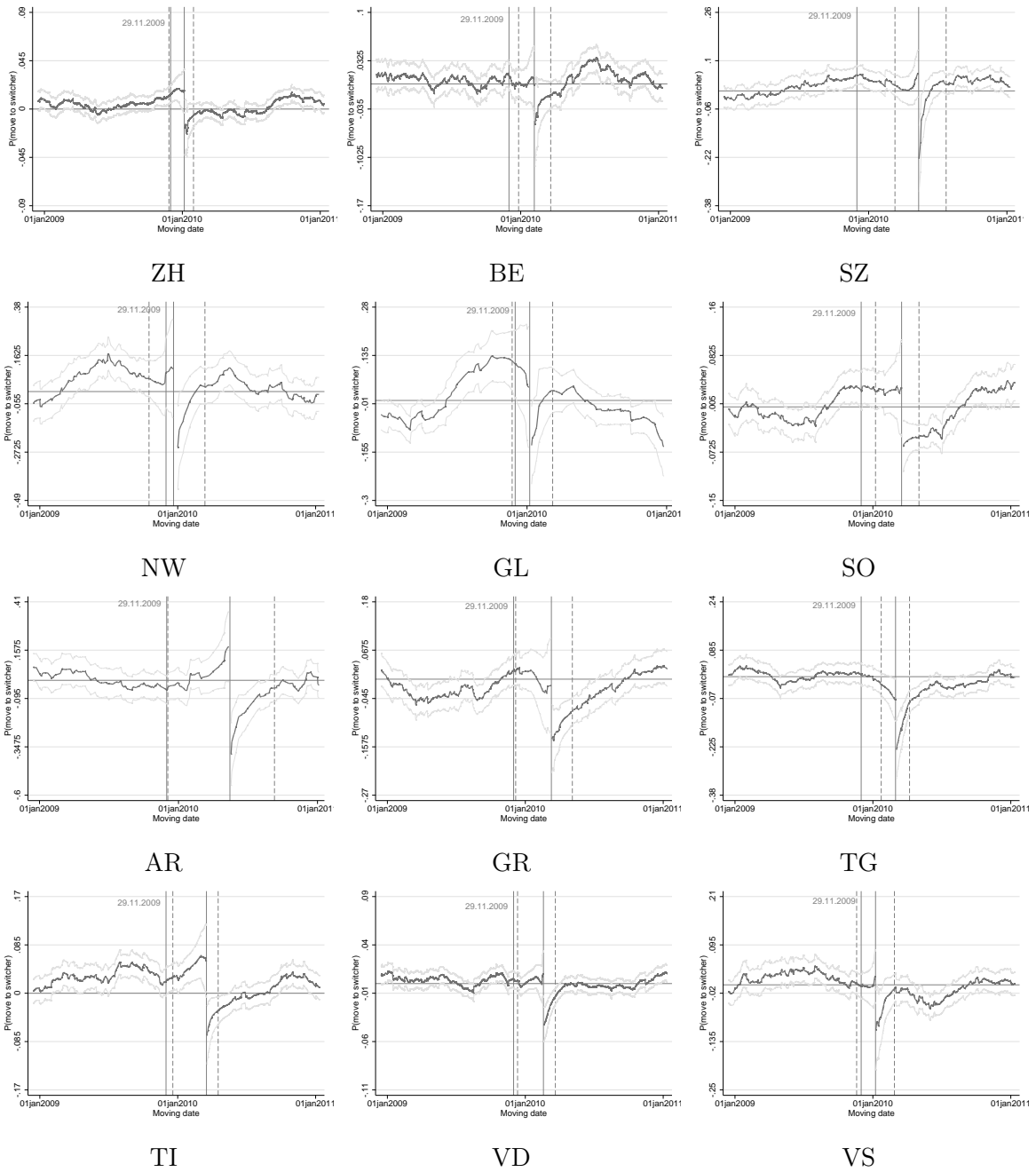
Variable	Mean	SD	Median	Min	Max	N
Switcher and Non-switcher municipalities in 2009						
Switcher	0.24	0.43	0	0	1	2,515
Mean vote shares $_{former}$						
All	50.82	12.6	52.48	13.98	85.33	2,515
Switcher	46.05	9.93	47.01	13.98	64.08	612
Non-switcher	52.35	12.98	54.82	14.8	85.33	1,903
Mean vote shares 2009						
All	63.26	10.71	64.1	25	100	2,515
Switcher	63.34	7.18	63.55	50	87.5	612
Non-switcher	63.24	11.62	64.3	25	100	1,903
$\Delta = (\text{vote share 2009} - \text{mean vote share}_{former})$						
All	12.44	7.57	12	-23.9	49.15	2,515
Switcher	17.29	7.25	17.14	3.25	49.15	612
Non-switcher	10.88	6.98	10.63	-23.9	39.08	1,903
rank change						
All	-0.3	48.39	-1	-294	299	2,515
Switcher	39.91	50.60	21.5	-8	299	612
Non-switcher	-13.23	39.81	-6	-294	131	1,903

*Notes:* Summary statistics for the main variables for the municipality classification.

**Table B.3:** Results for the specification test and the estimated threshold dates

Canton	$h_{test}$	$T_n$	$T_{n(90B)}^*$	$T_{n(95B)}^*$	$h$	Date $\hat{c}$
ZH	22.5	1.07	0.45	1.089	45	Jan 06, 2010
BE	22.5	2.245	0.559	1.702	45	Feb 06, 2010
SZ	37.5	4.768	0.796	1.751	75	May 13, 2010
NW	45	2.405	0.583	0.952	90	Dec 21, 2009
GL	45	0.777	0.38	0.764	90	Jan 08, 2010
SO	30	2.191	0.864	1.441	60	Mar 18, 2010
AR	45	3.327	0.837	1.398	90	May 16, 2010
GR	37.5	1.358	0.626	0.913	75	Mar 11, 2010
TG	30	2.624	0.272	0.796	60	Mar 02, 2010
TI	30	2.15	0.434	1.01	60	Mar 18, 2010
VD	22.5	1.866	1.46	2.442	45	Feb 18, 2010
VS	30	2.471	0.348	0.625	60	Jan 08, 2010

*Notes:* Summary of the results of the RDD with unknown discontinuity points. We report the resulting test value ( $T_n$ ), the bandwidth used in the specification test ( $h_{test}$ ), the critical values on the 10 and 5 percent level, the bandwidth used in the threshold search ( $h$ ) and the estimated threshold date ( $c$ ) for each canton. The bootstrap results are obtained using 1,500 resamples. The last column reports the estimated threshold date.



**Figure B.7:** Probability of moving to a switcher municipality in 12 cantons - RDD plots around the estimated threshold dates. Local linear smooth, using the same bandwidth ( $h$ ) as in the estimations in Table B.4, of the probability that a foreigner chooses to move to a switcher municipality over the moving date, separately from both sides of the threshold for all cantons in which a discontinuity was found. The dashed vertical lines indicate the testing region, the first solid line indicates the vote date, and the second the estimated threshold date.

**Table B.4:** Change in the probability that foreigners move to a switcher municipality - RDD estimates at the estimated thresholds for 12 cantons

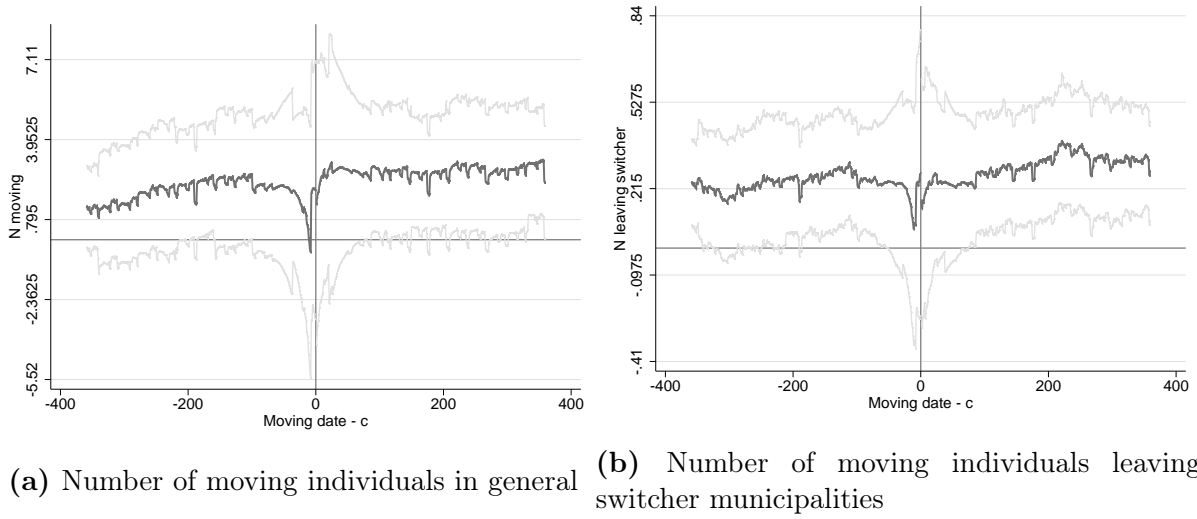
Dependent variable: Moving to a switcher municipality										
canton	ZH	BE	SZ	NW	GL	SO	AR	GR	TG	
$\tau$	-0.0345** (0.0156)	-0.0674* (0.0405)	-0.296*** (0.0689)	-0.437*** (0.153)	-0.187* (0.106)	-0.0953** (0.0483)	-0.611*** (0.121)	-0.121* (0.0683)	-0.168*** (0.0541)	
$P(\text{switch})_{before}$	0.07	0.17	0.20	0.36	0.27	0.14	0.39	0.22	0.27	
h	45	45	75	90	90	60	90	75	60	
N left	2,106	687	340	109	73	308	72	204	453	
N right	1,543	664	350	68	115	460	88	395	500	
Dependent variable: Moving to a switcher municipality										
canton	TI	VD	VS							
$\tau$	-0.132*** (0.0435)	-0.0536*** (0.0158)	-0.131** (0.0583)							
$P(\text{switch})_{before}$	0.21	0.05	0.27							
h	60	45	60							
N left	541	1,090	576							
N right	621	1,562	376							

*Notes:* Local linear sharp RDD estimates. Standard errors in parentheses.  $P(\text{switch})_{before}$  refers to the average probability of moving to a switcher municipality within a range of 60 days before the threshold date for the respective sample. Significance levels: \* .05 <  $p$  < .1, \*\* .01 <  $p$  < .05, \*\*\*  $p$  < .01.

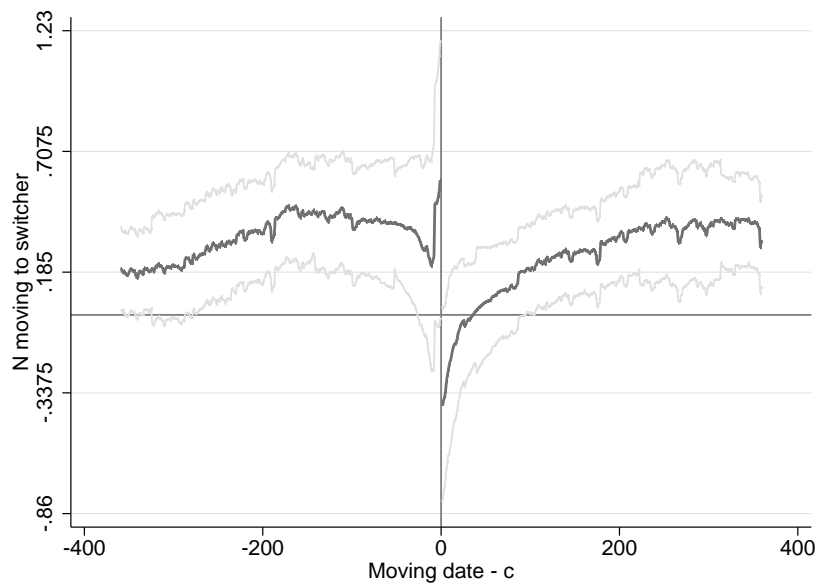
**Table B.5:** Results of the placebo specification test

Canton	$h_{test}$	$T_n$	$T_{n(90B)}^*$	$T_{n(95B)}^*$	$h$	Date $\hat{c}$
ZH	22.5	-0.695	0.294	0.655	45	-
BE	22.5	-2.227	0.248	0.672	45	-
SZ	37.5	-1.863	0.567	1.308	75	-
NW	45	-0.556	0.717	1.155	90	-
GL	45	-2.619	1.007	1.398	90	-
SO	30	1.229	2.265	2.752	60	-
AR	45	-0.249	0.674	1.228	90	-
GR	37.5	0.164	0.564	0.926	75	-
TG	30	-1.211	0.39	0.658	60	-
TI	30	-3.043	0.083	0.518	60	-
VD	22.5	-3.248	0.332	0.631	45	-
VS	30	0.369	0.437	0.76	60	-

*Notes:* Summary of the results of the RDD with unknown discontinuity points and a placebo testing period (testing period -365 days). We report the resulting test value ( $T_n$ ), the bandwidth used in the specification test ( $h_{test}$ ), the critical values on the 10 and 5 percent level, the bandwidth used in the threshold search ( $h$ ) and the estimated threshold date ( $c$ ) for each canton. The bootstrap results are obtained using 1,500 resamples. The last column would report the estimated threshold date if the null hypothesis could be rejected.



**Figure B.8:** Number of individuals moving around the threshold. Local linear smooth of the number of individuals who move over the moving date, separately from both sides of the threshold, using a bandwidth of 90 days.



**Figure B.9:** Number of individuals moving to a switcher municipality. Local linear smooth of the number of individuals who move to a switcher municipality over the moving date, separately from both sides of the threshold, using a bandwidth of 90 days.