

**Universität
Basel**

Wirtschaftswissenschaftliche
Fakultät

WWZ

November 2016

Prejudice in Naturalization Decisions: Theory and Evidence

WWZ Working Paper 2016/04

Dragan Ilić

A publication of the Center of Business and Economics (WWZ), University of Basel.

© WWZ 2016 and the authors. Reproduction for other purposes than the personal use needs the permission of the authors.

Universität Basel
Peter Merian-Weg 6
4052 Basel, Switzerland
wwz.unibas.ch

Corresponding Author:
Dr. Dragan Ilić
Tel: +41 (0) 61 207 33 57
Mail: dragan.ilic@unibas.ch

PREJUDICE IN NATURALIZATION DECISIONS: THEORY AND EVIDENCE

Dragan Ilić*

November 28, 2016

Abstract

Immigrant groups that are marginalized in their host countries are disproportionately more likely to have their citizenship applications rejected. It is not readily obvious whether this disparity is due to prejudice on the part of decisionmakers or due to applicant differences in meeting naturalization standards. To address this question, I develop a simple model of a council deciding whether to grant applicants citizenship. The model implies an empirical test for relative prejudice using average applicant group rejection rates. Using Switzerland as a case study, I apply the test to newly collected data from six large municipalities. In five municipalities, the test cannot reject the hypothesis of no relative prejudice with respect to country of origin. The rejection pattern of the sixth municipality is consistent with prejudice. The model illustrates that the underlying mechanism in the decisionmaking process has bearing on the inference of prejudice from empirical data.

*University of Basel, Faculty of Business and Economics, Peter Merian-Weg 6, 4002 Basel, Switzerland, and ETH Zurich, CER Center of Economic Research, ZUE F11, Zurichbergstrasse 18, 8092 Zurich (e-mail: dragan.ilic@unibas.ch). I am grateful to Georg Nöldeke, George Sheldon, Alois Stutzer, and Brigitte Guggisberg for their help on this project. I would like to thank all participating councils for their time and effort in providing the data. Funding for this research was generously provided by the WWZ Forum and the National Centre of Competence in Research "On the Move".

1 INTRODUCTION

For immigrants, naturalization can be described as a last stepping stone of settling in in their host country. Granting and receiving citizenship are not solely symbolic acts. Citizenship irreversibly opens the gate to a nation’s political and territorial association. Citizens can vote and have an indefinite right of stay. Citizenship also often comes with higher wages and improved access to the labor market (Fougère and Safi, 2009). The benefits of citizenship are not restricted to the naturalized; there is also evidence that naturalizations may improve the political and social integration of immigrants (Steinhardt, 2012; Hainmueller, Hangartner, and Pietrantonio, 2015). Naturalizations thus have the potential to affect a myriad of political, social, and economic settings in host countries. That potential is especially palpable in Europe, Northern America, and Oceania, where every tenth resident was not born where they live (United Nations, 2016).

Whenever minorities are at such mercy of decisionmakers, disputes about prejudice tend to crop up: housing, jobs, policing, and the justice system are just a few examples. The concern in naturalizations is whether the eligibility of marginalized applicants striving for citizenship is measured by the same yardstick. However, despite the gravitas of the decision, not much is known about systematic prejudice in granting citizenships. This contrasts with a well-established literature on discrimination in other arenas.¹ But

¹The following attempt gives a tiny glimpse into the literature. Unless otherwise noted, the cited studies investigate data from the United States. For discrimination in housing, see Zhao, Ondrich and Yinger (2006), who focus on real estate brokers, and the study by Ahmed and Hammarstedt (2008) for Sweden’s rental market. A comprehensive recent study by Ewens, Tomlin and Wang (2014) pinpoints to statistical discrimination rather than taste-based discrimination in the rental apartment market. For surveys about ethnic discrimination in labor markets, see Neumark (2013) and Lang and Lehmann (2012). Research on discrimination in policing spans from the seminal paper by Knowles, Persico and Todd (2001) until a very recent and controversial study by Fryer (2016) on police violence. Ayres and Waldfogel (1994) present early evidence for discrimination at the bail bond setting stage, and Alesina and La Ferrara (2014) reject the hypothesis of no prejudice in capital sentencing. Shayo and Zussman (2011) present evidence for judicial ingroup bias in Israel. Rich data is key in empirical analyses: More evidence on discrimination comes from experiments selling iPods in local online classified advertisements (Doleac and Stein, 2013) and from a job market field experiment that implemented changes in the cost of discrimination during required collaborations (Heddegaard and Tyran, 2016). Sports provide excellent data: Price and Wolfers (2010) study racial discrimination among basketball referees in the NBA, and Parsons et al. (2011) study prejudice among baseball umpires in the MLB.

there is no reason to expect that naturalizations are immune to prejudice, nor is there a lack of controversy. For example, an ongoing lawsuit by Muslim applicants in the United States raises complaints against Immigration Services.² This is just the latest tip of the iceberg. There has been a wave of lawsuits related to naturalizations in the United States in the last ten years.³ In Germany, the United Nations Committee on the Elimination of Racial Discrimination expressed its discomfort with application questions biased against Muslim applicants (CERD, 2008). In Switzerland, applicants from Turkey and former Yugoslavia, both minority groups, have long been suspected to be at a disadvantage by sometimes being rejected at will (Helbling, 2008). Yet more controversy could be looming. Within the scope of its immigration reform, the US government has proposed to provide a path to citizenship for its 11 million undocumented workers (The White House, Office of the Press Secretary, 2013).

One reason for the dearth of research on prejudice in naturalizations is scant empirical data. Countries usually only publish aggregate data of granted citizenships, at best stratified by country of origin. Unravelling prejudice, however, requires more data. Lower naturalization rates or higher rejection rates for certain groups, for instance, may have many accountable causes. For one, marginalized groups tend to face larger socio-economic disadvantages, so it is conceivable that they are less likely to be eligible and would fail to meet naturalization standards more often than other applicant groups. Accounting for such differences in patterns of eligibility is key if one wants to test rigorously for prejudice. With rich microlevel data one could attribute differences in naturalization outcomes to the effects of differences in measured applicant characteristics. Not all these characteristics, however, are usually observable to researchers.

An exception that managed to pursue this very strategy is the study by Hainmueller and Hangartner (2013) for Switzerland. Hainmueller and Hangartner collected data on

² *Arapi et al v. U.S. Citizenship & Immigration Services et al*, 4:2016cv00692, E.D. Mo May 18, 2016

³ For example, *Tarek Hamdi v. United States Citizenship and Immigration Service et al*, 5:2010cv00894, C.D. Cal, June 16, 2010. The online database of the Justia Corporate Center (dockets.justia.com) list several thousands lawsuits related to US naturalization applications since 2004.

closed ballot votings in hundreds of municipalities, administrative divisions of the state that have the subsidiary authority to grant Swiss citizenship to their local residents. By tediously codifying voting leaflet information of 2,400 anonymous naturalization referendums from 1970-2003, Hainmueller and Hangartner gathered evidence that supports the suspicion of disadvantaged minority applicants. However, the specific research design proves problematic because it is germane to that particular empirical setting. Courts have since outlawed the controversial policy. Municipalities can still decide locally, but Switzerland now employs, in line with most other countries, councils for evaluating naturalization applications to ensure privacy and justifiability for the applicants. In assessing their applications, councils usually also conduct personal interviews. Switzerland's regime change, then, exemplifies the general empirical difficulty in detecting prejudice in naturalizations. Not only does the evaluation policy render research access to microlevel data next to impossible, it also makes the omitted variable bias problem appear in the empirical analysis. Signals of eligibility in personal exchanges are difficult to observe, let alone easily quantifiable. A viable research design to test for prejudice in naturalizations has thus remained at large.

This paper addresses this gap and develops a simple model of a decisionmaking process in naturalizations that implies an outcome test for prejudice. The model posits that a principal's decision whether to grant citizenship will take into account any signal that is informative of an applicant's eligibility. That makes the rate of *rejected* applications indicative of the required expected eligibility for being granted citizenship. The model demonstrates how the rank order pattern of these rates across various groupings of principals and applicant groups can be tested for prejudice. More concretely, in the model principals weigh the benefit of naturalizing qualified applicants against the cost of mistakenly naturalizing unqualified applicants. In addition to country of origin, a principal observes a noisy but informative signal about the merit of each of its applicants. Qualified applicants are more likely than unqualified ones to send a favorable signal. In this set-

ting, the principal's optimal decision is characterized by setting a threshold of expected qualification. Only an applicant whose signal exceeds this standard will be granted citizenship. Different principals may use different standards, but the model implies that if principals are not prejudiced against applicants based on, say, some applicants' country of origin, the rank order of the rejection rates, grouped by country of origin, should be the same for every principal. It is this variation in strictness of granting citizenship that the model exploits (and requires) when empirically testing for prejudice. The next section shows that such patterns are often observable in naturalization data. A nice feature of the model is that it does not require knowledge of all the applicant characteristics that are taken into account by the decisionmaking councils; in particular those characteristics that are typically not available to researchers.

The model adds to the ongoing research on modeled outcome tests of discrimination, which originates back to Knowles, Persico and Todd (2001). In their paper, police officers and motorists strategically interact in the context of motor vehicle searches. To test prejudice among the police, Knowles, Persico, and Todd compare the search success rates of white, black, and Hispanic motorists. Anwar and Fang (2006) generalize that seminal model and drop two key assumptions; the requirement that motorists respond to the probability of being searched and the assumption of a monolithic police force. Discrimination in healthcare is the subject in Anwar and Fang (2013), who investigate error patterns in emergency room visit discharges. In their model, physicians have to decide whether to keep emergency room patients in the hospital for further investigation or whether to send them home. The test for prejudice against patients of different races relies on the comparison of their bounce back rates - the probability that a patient had been misdiagnosed and needed to revisit an emergency room shortly after being mistakenly discharged. Parole release data have also been exploited in modeled outcome tests. Mechoulan and Sahuguet (2015) and Anwar and Fang (2015) put forth models to test for prejudice using rates of recidivism grouped by prisoner race. Finally, Alesina and

La Ferrara (2014) adapt Anwar and Fang's (2006) model to investigate capital sentencings and let courts minimize type I and type II errors in their rulings. Their model implies that with unbiased courts, conditional on defendants race the error rate should be independent of victims race. What is common to all modeled outcome tests is that they rely on rational choice behavior and, on that account, manage to get by without microlevel data. In general, average grouped outcome data suffice for the empirical tests. The model in this paper falls into line and builds upon Anwar and Fang's (2006) framework.

Using Switzerland as a case study, I apply my model to a landmark decision in naturalization law. In 2003, the Swiss Federal Court declared closed ballot voting illegal and mandated that councils, usually staffed with locally elected politicians, are to assess the eligibility of the local applications. In another study, Hainmueller and Hangartner (2017) show that this large scale of institutional change raised the municipal naturalization rates in Switzerland dramatically, even in municipalities that did not switch regimes.⁴ I raise new application data from six large Swiss municipalities and exploit the within-municipality variation in rejection rate patterns before and after the 2003 ruling.⁵ That is, for each municipality, I compare its rejection rate rank order before and after 2003. The unexpected regime change ensures that possible differences in the quality of the applicant pool are not confounding the test results. The test for prejudice cannot reject the hypothesis of no prejudice in five of the six municipalities. One municipality shows evidence that is consistent with prejudice. Finally, the model has a key testable implication - as the number of applications rise, the rejection rate increases as well. This rests upon the assumption that a rise in applications is driven by less eligible applicants. The empirical results support this prediction.

This paper unfolds as follows. Section 2 describes international naturalization policies and highlights considerable variation in strictness even within countries, a data pattern required for an empirical implementation of the test. The model and the test for prejudice

⁴I will revert to the Swiss naturalization studies in more detail in the next section.

⁵No Swiss municipality publishes information on rejections on their own.

are presented in Section 3. Section 4 proceeds with the Swiss case study and describes the new data on rejection rates. Section 5 presents the results from the empirical analysis. Section 5 discusses and concludes.

2 NATURALIZATION POLICIES AROUND THE WORLD

In countries that do not follow the *ius soli* principle, obtaining citizenship usually requires an active engagement even on the part of those immigrants that were born in their host countries. Naturalization procedures are country-specific, and these policies vary in strictness. For instance, in the European Union, despite attaining corresponding rights in all member states upon naturalization, policies range from actively promoting naturalization to restrictive access. These degrees are not cast in stone but change over time, again depending on the country (Bauböck et al., 2006). Whereas Belgium and Germany have strived for more liberal practices in the recent past, Greece and Italy have adopted increasingly conservative attitudes. This cross-sectional variation in strictness is not confined to the country level. Helbling (2010) shows that within countries, naturalization rates and the strictness of the interpretation of the national citizenship laws differ greatly among administrative divisions. In Germany, this holds true at the *Länder* level (Dornis, 2001; Hagedorn, 2001; Hailbronner, 2006). Austria, too, shows variations at its regional level (Cinar and Waldrauch, 2006). For the United States, North (1987) concludes that the implementation of the naturalization procedures vary a great deal between district offices. Helbling and Kriesi (2004) and Hainmueller and Hangartner (2013) reveal striking differences among municipalities in Switzerland. And even notoriously centralized states like France do not seem to be immune from distinct implementations of its national law (Weil, 2004).

The Swiss case lends itself particularly well for empirical research on naturalizations because the political system is characterized by a emphatic devotion to federalism and

subsidiarity. Helbling (2008) and Hainmueller and Hangartner (2013) provide comprehensive overviews on Switzerland's citizenship policy. Here I only highlight the features which resonate with the model developed in the next section. Citizenship in Switzerland is regarded as an act of state that is to be primarily delegated to the lowest of the three political levels, the municipality. Interested immigrants have to apply in their municipality of residence. The requests are forwarded to the cantonal and federal level, which conduct formal checks whether the application fulfills the legal requirements.⁶ Upon positive feedback, the tricky part begins. The applicant has to convince the municipality's council of his or her merit. Among other things, the council's decision is based on legal background checks, the applicant's characteristics, and the personal impression during face to face interviews. Integration to Swiss life, as measured by familiarity with local laws, traditions, and customs, is the central pillar in the assessment. These factors are often proxied by language skills, job status, or even hobbies. It falls to each council to which extent these requirements need to be met in order to be granted citizenship. This autonomy helps explain the varying naturalization rates at regional levels.

Before 2003, in hundreds of municipalities these decisions were discretionary and untraceable because closed ballot votings were used to decide on naturalization applications. That is, Swiss citizens could cast their vote on any naturalization application in the municipality were living in. Only applicants with a popular majority of "yes" votes received Swiss citizenship. This controversial policy would turn out to be immensely useful for research. Hainmueller and Hangartner (2013) collected data on 2,400 recorded naturalization referenda held between 1970 and 2003 in 44 Swiss municipalities using such closed ballots. Their analysis shows that, despite accounting for a battery of applicant characteristics that are commonly deemed relevant for naturalization, the proportion of "no" votes were about 40% higher on average for applicants from former Yugoslavia and

⁶Articles 14 and 15 of the Swiss Naturalization law leave some room for interpretation but list the following basic requirements: Overall 12 years of residency, integration into and familiarity with Swiss customs and tradition, adherence to Swiss law, and lack of threat for the internal and external security of Switzerland.

Turkey in comparison to applicants from rich European countries like Germany or the UK. Lower socio-economic credentials in form of occupational skills, educational attainment, or length of residency matter as well, but not in equal measure. Hainmueller and Hangartner attribute the discriminatory treatment to the lack of accountability associated with the nature of closed ballots. This attribution is consistent with the results of their follow-up study, where they compare naturalization rates before and after 2003 (Hainmueller and Hangartner, 2017). That year the Swiss Federal Court declared closed ballot voting illegal for naturalization applications. The ban was based on two grounds. First, the right to appeal rejected applications was not ensured. Closed ballots, by definition, lack the basis for any contestable justification. Second, the court disapproved of the severe lack of privacy in the evaluation of the applicants, who had to reveal detailed background information to all eligible voters, usually in form of detailed leaflets (the very information that was harnessed for research). Hainmueller and Hangartner show that the switch to elected councils increased the naturalization rates of the formerly discriminated applicants on average by 50%, more so in municipalities with a high vote share of the conservative Swiss People’s Party (SVP), and even in municipalities that did not switch their naturalization regime. This is interpreted as evidence of discrimination on the assumption that once accountable legislators instead of anonymous voters are in charge of handling naturalizations, indulgence of discriminatory preferences become more costly and thus less likely.

Heterogeneity in strictness in the application of the law are a widespread feature of naturalization policies. In the next section I advance a model that exploits these variations to develop a test for prejudice. The model is based on the premise that in screening their applicants, councils take into consideration any signals that are informative of actual qualification for citizenship. Chiswick and Miller (2008) suggest that individual-level characteristics systematically relate to citizenship status. Duration of residence, command of the host country’s official language, and, to some extent, higher

levels of skill (indexed by education attainment and occupation), all link to naturalization, although Dronkers and Vink (2012) put the role of skill in perspective. The highest explanatory power for naturalization according to Chiswick and Miller are country of origin characteristics, even more so than individual characteristics. This is in accord with Dronkers and Vink (2012) and DeVoretz and Pivnenko (2005), who show that immigrants from poor or politically unstable countries are more likely to naturalize. My model assumes that deciding councils condense this individual and group information into an index of expected eligibility.

3 A MODEL OF NATURALIZATION DECISIONS

3.1 THE MODEL

Consider two councils that separately evaluate their resident immigrants which apply for naturalization. Let $c \in \{A, B\}$ denote the councils. In each council, there are continuums of applicants grouped into country of origin $e \in \{R, F\}$, where R and F stand for related and foreign, respectively.⁷ Suppose that among applicants of origin e , a fraction α_e is objectively unqualified for naturalization. I allow for the possibility that the unobservable characteristics among applicants from different countries of origin may differ. Councils therefore evaluate the merit for naturalization based on the applicant's origin and a myriad of observable applicant characteristics. A council may consider information such as gender, age, number of children, language skills, familiarity with local habits and law, duration of residence, employment status, or level of education and income. Importantly, a council also processes characteristics that are difficult to observe for a researcher, such as demeanor or congeniality during personal interviews. Let us assume that a council condenses all this information into a one-dimensional index $\theta \in [0, 1]$ which reflects the likelihood that an applicant is unqualified for naturalization. This index is randomly

⁷The model readily extends to n councils and m immigrant groups.

drawn from a continuous density function $f_u^e(\cdot)$ if the applicant is actually unqualified. If the applicant is qualified, the index is drawn from $f_q^e(\cdot)$. For this index to be indicative of qualification, the two densities are assumed to satisfy the strict monotone likelihood ratio property, that is, for $e \in \{R, F\}$, $f_u^e(\theta)/f_q^e(\theta)$ is strictly increasing in θ . This implies $F_q^e(\theta) < F_u^e(\theta)$ for all θ . In words, higher values of the signal θ are more likely if the applicant is unqualified. Some applicants may produce signals which will unambiguously lead to rejections, such as a criminal record. I thus assume an unbounded likelihood ratio: $f_u^e(\theta)/f_q^e(\theta) \rightarrow +\infty$ as $\theta \rightarrow 1$. After interviewing an applicant of country of origin e with signal θ , a council decides whether to grant citizenship or not. While councils only see imperfect signals during the course of the application, they do eventually realize whether their decision was correct. Councils derive the benefit $b(c, e)$ from naturalizing a qualified applicant. At the same time, councils bear the marginal cost t when mistakenly naturalizing an unqualified applicant. Note that the benefit can depend both on the deciding council and on the applicant's country of origin.⁸ Councils may have a taste to prefer applicants of certain origin to be naturalized (or conversely, dislike applicants of certain origin to become citizens). Based on this taste, a given council whose benefit depends on the applicant's country of origin is said to be prejudiced:

$$b(c, R) \neq b(c, F)$$

On the other hand, councils may derive different levels of benefit in general from naturalizing qualified applicants, levels that do not depend on country of origin. Such differences in benefits could stem, for instance, from varying identity preferences based on group distinction or high council standards associated with the merit of citizenship. Define councils to be heterogenous if

$$b(A, e) \neq b(B, e)$$

⁸For the sake of simplicity, I assume t to be independent of council and country of origin.

for some e . Councils that derive little benefit from naturalization are said to be strict. Likewise, councils that derive a lot of benefit are said to be lenient. It is easy to see that heterogenous councils do not imply prejudice against a given country of origin. By the same token, homogenous councils do not imply the lack of prejudice as both councils might equally prefer or dislike applicants from certain backgrounds.

3.2 THEORETICAL IMPLICATIONS

Denote by U the undesirable event that a naturalized applicant turns out to be unqualified. The probability of this event depends both on the signal θ observed during the interview and on country of origin. Following Bayes' rule, this probability is given by

$$\Pr(U | e, \theta) = \frac{\alpha_e f_u^e(\theta)}{\alpha_e f_u^e(\theta) + (1 - \alpha_e) f_u^e(\theta)}. \quad (1)$$

The monotone likelihood ratio property implies that this probability strictly increases in θ . Since the signal is informative, a higher level correctly reflects an increase in the mistake probability. Now consider the decision problem of a council faced with this information:

$$\max \{b(c, e) [1 - \Pr(U | e, \theta)] - t \Pr(U | e, \theta); 0\}$$

The first term describes the expected benefit from naturalizing a qualified applicant minus the cost of mistakenly naturalizing an unqualified applicant. Not naturalizing yields a benefit of zero. The costs associated with the naturalization process itself are considered fixed and are thus disregarded. Thus, the council does best to naturalize if and only if

$$b(c, e) [1 - \Pr(U | e, \theta)] > t \Pr(U | e, \theta),$$

in words, whenever the expected benefit of naturalizing outweighs the expected cost.

This naturalization condition reduces to

$$\Pr(U \mid e, \theta) < \frac{b(c, e)}{b(c, e) + t} \quad (2)$$

Intuitively, a high benefit of naturalization makes for a riskier behavior because the council accepts a higher probability of mistakes. On the other hand, high costs of making mistakes do not leave a lot of room for suspicious signals. Because $\Pr(U \mid e, \theta)$ is strictly increasing in θ , the naturalization condition described by (2) implies that the council grants citizenship if and only if

$$\theta \leq \theta^*(c, e),$$

where the naturalization threshold $\theta^*(c, e)$ is pinned down by

$$\Pr(U \mid e, \theta^*(c, e)) = \frac{b(c, e)}{b(c, e) + t}. \quad (3)$$

The applicant sending the signal $\theta^*(c, e)$ is called the marginal applicant who is deemed just worthy enough to be granted citizenship. Any applicant with a higher signal than this standard will be rejected. Likewise, any applicant with a lower signal will be naturalized. It is straightforward to see that $\theta^*(c, e)$ is strictly increasing in $b(c, e)$ and strictly decreasing in t . The higher the benefit of naturalization, the worse the qualification standard is allowed to become. In turn, higher costs imply a stricter expected quality conveyed by the signal. The signal threshold $\theta^*(c, e)$ determines the average rejection rate of council c against applicants of origin e in equilibrium:

$$r(c, e) = \alpha_e [1 - F_u^e(\theta^*(c, e))] + (1 - \alpha_e) [1 - F_q^e(\theta^*(c, e))] \quad (4)$$

This rejection rate is monotonically decreasing in $\theta^*(c, e)$: The worse the signal is allowed to become, the less applicants are rejected. Note that a council that is not prejudiced

might nevertheless use different signal thresholds for the applicant groups. An unprejudiced council does not aim to set equal signal thresholds but equal probabilities of qualification for the marginal applicants. If one applicant group is known to have a larger unqualified fraction or if the signal θ is distributed differently between the two applicant groups, councils take this information into account. This implication is an inherent part of models of statistical discrimination. Equation (1) illustrates that because the applicant's qualification is not perfectly observed, a council's optimal assessment about the qualification of a given applicant does not solely depend on that applicant's signal. The assessment also considers the fraction of qualified applicants in that group. Also note that if councils are homogenous they all derive the same benefit from a given applicant group. In that case, the threshold condition (3) implies that unbiased councils set the same signal threshold for this group. Consequently, the rejection rate defined by (4) against this particular applicant group would be the same for both councils.

Based on this insight, consider the following simple transitivity example. Assume $b(A, R) > b(B, R)$ applies so that council A derives a larger benefit from naturalizing related immigrants than council B does. If the two councils are not prejudiced, it must be true that $b(A, R) = b(A, F)$ and $b(B, R) = b(B, F)$. It immediately follows that $b(A, F) > b(B, F)$ so that council A also derives a larger benefit from naturalizing immigrants from group F . In other words, if the councils are heterogenous but not prejudiced, the ranking of $b(A, e)$ and $b(B, e)$ does not depend on country of origin. What can be said about the signal thresholds? We know that $\theta^*(c, e)$ is strictly increasing in $b(c, e)$, so above example implies $\theta^*(A, R) > \theta^*(B, R)$ and $\theta^*(A, F) > \theta^*(B, F)$. Moreover, because the rejection rate $r(c, e)$ is monotonically decreasing in $\theta^*(c, e)$, it also follows that $r(A, R) < r(B, R)$ and $r(A, F) < r(B, F)$. The rejection rates of the lenient council A will be lower for both applicant groups. To sum up, if councils are heterogenous but not prejudiced, the ranking of the rejection rates across councils does not depend on country of origin. If this rank order is violated, we can deduce relative prejudice among the councils. However,

because there is no objective rank order that defines impartiality, one cannot pinpoint the discriminating council.

3.3 AN OUTCOME TEST FOR PREJUDICE IN NATURALIZATION DECISIONS

The model provides a test for prejudice that is applicable with average outcome data. This test can be implemented even when researchers have no access to the signals which are observed by the councils when deciding whether to grant citizenship. The theoretical implications predict that under the null hypothesis of no relative prejudice among the councils, the rank order of the average rejection rates for a given country of origin e across councils $c \in \{A, B\}$ does not depend on country of origin $e \in \{R, F\}$. Heterogeneity across councils is thus a prerequisite for the empirical application of the test.

3.4 DISCUSSING THE MODEL

Outcome tests have notorious issues with infra-marginality. Generally one cannot infer disparate treatment from (average) outcome data. Instead, it is the outcome of marginal decisionmaking that is informative of animus.⁹ It is useful to elaborate on this distinction. Recall that the councils only naturalize applicants with a signal that is below the naturalization threshold (3) for that group. In other words, a council only naturalizes applicants who are deemed qualified enough. If a council is not biased, at the margin it requires the same probability of qualification no matter the country of origin. But we know that depending on the group-specific distributions of the signals and the fraction of unqualified applicants in that group, the average rejection rates may vary despite the same marginal decisionmaking process. Since empirical data only provides information on average outcomes, the infra-marginality issue poses a key obstacle for inferences of disparate treatment via outcome data.

⁹For an extended description of this issue see Becker (1993), Yinger (1996), or Ayres (2002).

The proposed outcome test in this paper circumvents this issue. The test does not directly compare the average rejection rates for a given country of origin across councils. Instead, it makes use of the rankings implied by the model, an indirect identification strategy so to speak. These rankings exploit the simple fact that under the model assumptions, the direction in which the average rejection rate moves is unambiguously determined by the direction in which the council adjusts the required marginal probability of qualification. A higher (lower) marginal probability of qualification always implies a higher (lower) average rejection rate. Put simply, the average moves with the marginal because they are strictly monotonically related. So although we cannot infer directly if the required marginal probabilities of qualification are equal when looking at the average rejection rates, we do know which rank order of the average rejection rates would reject the hypothesis that the marginal probabilities are equal.

This identification strategy bears a caveat. Like in Anwar and Fang's (2006) test for racial prejudice in motor vehicle searches, there is some leeway of prejudice in the average rejection rates due to the ordinal nature of the test. Imagine that one council is prejudiced against applicants from a certain country of origin, which raises the rejection rate. But the proposed test will fail to detect prejudice if this rate remains within the allowed range which is consistent with the null hypothesis of an independent rank order across councils. This is the case if the prejudice is not too strong, where strong is relative and depends on the magnitude of prejudice required for a violation of the rank orders and on the differences in benefit across councils. The larger these differences, the more leeway there is. In statistical terms, there is a high probability of a type-II error, not rejecting the null hypothesis of no prejudice. All the same, if the test does indicate prejudice, it does so with high confidence.

There also looms an empirical issue. The tests assumes that all councils face pools of applicants which are sampled from the same population. This assumption may not hold. For instance, if one municipality attracts particularly high-skilled immigrants from

Turkey, rejection rates in that municipality may be lower because of the higher eligibility, not because its council is more lenient. In the case study below, my identification strategy circumvents this pitfall by testing *within* municipalities over time. The sudden change of the naturalization regime in 2003 mandated by the Swiss Federal Court ensures that councils face pools of applicants which are sampled from the very same population. Because a naturalization application takes at least two years to go through, applicants could not have anticipated the regime change, so any observable changes in rejection rates cannot be explained by a change in the application pool prior to the ruling.

4 DATA

There are a lot of municipalities in Switzerland, and most of them are small. In 2016, with a population count of 8 million residents, there were 2'294 municipalities in Switzerland. The mean municipality had 3'600 residents, the median municipality had 1'400 residents. A quarter of the population has no Swiss passport. Still, naturalizations are not an everyday occurrence. Roughly 40,000 immigrants - two percent of the foreign population - are currently naturalized in Switzerland every year. The regional fragmentation and the naturalization rates suggest that rejection numbers by municipality and country of origin, are likely to have, on average, a low frequency.

In order to ensure a sufficient number of observations for the empirical test, I invited all Swiss municipalities in the German and French speaking parts with at least 20'000 citizens as of 2010 to list their total number of both granted and rejected citizenship applications by country of origin on an annual basis from 1998-2011. Of the 35 contacted municipalities, six managed to provide sufficient data, henceforth dubbed municipalities A to F. This seems like a low response rate, but almost no contacted municipality had the requested data at their immediate disposal. In contrast to granted citizenships, rejections are typically not systematically recorded and needed to be compiled manually

for the purpose of this survey. Depending on the available resources and the existing format of the records, some municipalities simply did not have the means to respond. Politics may have played another role. Naturalizations being a highly controversial issue in Switzerland, all municipalities were assured absolute anonymity. Even so, it is possible that some municipalities may have self-selected into abstaining.

The two most obvious marginalized countries of origin in Switzerland are Turkey and former Yugoslavia.¹⁰ Hainmueller and Hangartner (2013) have demonstrated that applicants from other countries have been shown to be less likely to be rejected. I follow their grouping of country of origin and classify two European comparison groups. Germany, France, and Great Britain comprise the richer northwestern group. I have narrowed the group to these countries because they provide a sensible number of observations in all municipalities. Italy, Portugal, and Spain make up the southern, less affluent European group. These two European groups were part of different immigration waves. The southern group settled in the 1970s in great numbers. The northwestern group has been particularly on the move since the 2000s. This aggregation by country of origin raises the number of categorical observations and facilitates the implementation of the empirical test. The combined rejections of the southern group allow for the juxtaposition to Turkish and Yugoslavian rejections. On their own, Italy, Portugal, and Spain do not provide enough data. Dearth of observations is also the reason why no other countries have not been included into the empirical analysis.

¹⁰The countries of former Yugoslavia are Bosnia and Herzegovina, Croatia, Kosovo, Macedonia, Montenegro, Slovenia, and Serbia. In what follows, I will simply write Yugoslavia.

Table 1: Municipal Rejection Rates in Percent, by Applicant Group and Year

Mun.	Group	Year													
		98	99	00	01	02	03	04	05	06	07	08	09	10	11
A	Yu	49	100	100	90	88	-	-	0	16	10	3	2	0	0
		(d)	(b)	(c)	(d)	(c)	(a)	(a)	(b)	(c)	(d)	(c)	(d)	(d)	(d)
	Turkey	100	100	100	83	-	-	-	-	0	0	-	0	0	0
		(b)	(b)	(b)	(b)	(a)	(a)	(a)	(a)	(b)	(b)	(a)	(b)	(b)	(b)
	South	26	0	43	67	42	-	-	0	0	0	0	0	0	0
		(c)	(b)	(b)	(c)	(c)	(a)	(a)	(b)	(b)	(b)	(b)	(b)	(b)	(b)
	North	-	-	-	50	0	-	-	-	0	-	0	0	-	0
		(a)	(a)	(a)	(b)	(b)	(a)	(a)	(a)	(b)	(a)	(b)	(b)	(a)	(b)
B	Yu	30	28	23	26	23	19	12	11	13	17	15	15	18	16
		(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(d)	(d)	(d)	(d)	(d)
	Turkey	19	7	7	11	18	6	0	6	9	6	7	3	7	12
		(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(d)	(c)	(d)	(d)	(d)
	South	8	7	4	0	4	0	4	0	7	6	3	5	5	6
		(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(d)	(c)	(d)	(d)	(d)
	North	0	0	0	0	0	0	0	0	0	13	0	14	13	10
		(b)	(b)	(b)	(b)	(b)	(b)	(b)	(b)	(b)	(b)	(b)	(c)	(c)	(c)
C	Yu	42	40	48	35	48	19	5	4	5	8	11	12	17	20
		(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)
	Turkey	27	29	25	17	23	10	15	13	15	18	15	19	23	25
		(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(d)	(d)	(d)	(d)
	South	14	12	7	8	8	0	0	0	0	0	0	4	3	5
		(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(d)
	North	0	0	0	9	0	0	0	0	0	0	0	0	0	0
		(b)	(b)	(b)	(c)	(b)	(b)	(b)	(b)	(c)	(c)	(c)	(c)	(c)	(d)
D	Yu	32	31	30	33	32	25	14	17	17	24	22	22	28	27
		(c)	(d)	(c)	(c)	(c)	(d)	(c)	(d)	(d)	(d)	(d)	(d)	(d)	(d)
	Turkey	20	25	18	23	22	14	0	8	7	17	13	14	15	15
		(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(d)	(d)	(d)	(d)	(d)
	South	9	13	12	10	14	6	0	0	5	4	11	11	12	11
		(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(c)	(d)	(d)	(d)	(d)
	North	0	0	0	0	0	0	0	0	0	0	0	7	0	0
		(b)	(b)	(b)	(b)	(b)	(b)	(b)	(b)	(b)	(b)	(b)	(c)	(c)	(c)

Table 1 Continued: Municipal Rejection Rates in Percent, by Applicant Group and Year

Mun.	Group	Year														
		98	99	00	01	02	03	04	05	06	07	08	09	10	11	
E	Yu	17	24	20	17	21	15	14	18	14	17	17	18	22	22	
		(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	
	Turkey	0	8	0	8	0	0	0	0	0	11	0	4	5	3	
		(b)	(c)	(b)	(c)	(b)	(b)	(c)	(c)	(b)	(b)	(c)	(c)	(c)	(d)	
	South	8	7	8	7	7	4	4	5	3	5	4	6	7	8	
		(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	
	North	19	18	21	21	19	14	17	16	18	20	23	23	24	24	
		(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	(d)	
	F	Yu	12	15	13	19	11	11	12	15	11	15	11	15	16	11
			(c)	(b)	(c)	(b)	(c)	(b)	(c)	(b)	(c)	(b)	(b)	(b)	(c)	(c)
		Turkey	7	14	0	0	8	0	8	0	0	7	0	0	10	6
			(c)	(b)	(c)	(b)	(c)	(b)	(c)	(b)	(c)	(b)	(b)	(b)	(c)	(c)
South		0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		(b)	(c)	(b)	(c)	(b)	(b)	(b)	(c)	(b)	(c)	(b)	(c)	(c)	(c)	
North		0	0	-	0	0	0	0	0	0	0	0	-	0	0	
		(b)	(b)	(a)	(b)	(b)	(b)	(b)	(b)	(b)	(c)	(b)	(a)	(b)	(b)	

Note: To preserve anonymity, the letters in the parentheses denote the categories within which the corresponding number of observations falls: a (no observation), b (1-10 observations), c (11-30), d (> 30).

Table 1 shows, in percent, the rejection rates for the six municipalities A to F for the years 1998 to 2011, classified by applicant group. For example, in municipality B in 2002 applicants from Yugoslavia were rejected at a rate of 23 percent. Figure A1 in the appendix and, by taking two examples, Figure 1 and 2 visualize these data. Table 2 conflates, for each applicant group and municipality, the combined rejection rates for entire period. For example, from 1998-2011 applications from the southern European group in municipality E got rejected at a rate of six percent. Taken together, in the six municipalities 1,516 out of 11,345 applications, or roughly 13%, were rejected during the 14 years of observation. The municipalities differ in their number of applications. Municipality E has the highest number of observations with over 5,000 recorded applications, accounting for almost half of the observations in the data set. Municipality A has less than 600 observations. By and large, Yugoslavians and Turks apply in greater

numbers than the two European groups. Their rates are different, too. Turks and Yugoslavians typically exhibit the highest rejection rates, broadly ranging from 10% to 40% over the years, with Yugoslavians consistently being rejected at a higher rate. Applicants from the northwestern European group usually do not have to fear rejections, except in municipality E, where they even join the ranks of the Yugoslavian rates. Because municipality E also provides the highest number of observations in the data set, the aggregated rejection rates over all municipalities for the northwestern group is higher than for the southern group. In general, however, the southern group is situated slightly above the northwestern group on an annual basis in the other municipalities. There seems to be another overarching ranking pattern in the rejection rates. There seems to be heterogeneity *between* municipalities, too. Across the board, municipality F does not reject often. Municipality D, in contrast, seems to be more restrictive in general. This dovetails with the conclusions from Section 2.

Table 2: Aggregated Rejection Rates in Percent, 1998-2011

Applicant Group	By Municipality						All Municipalities
	A	B	C	D	E	F	
Yu	32	18	20	25	18	14	21
	(g)	(g)	(h)	(g)	(h)	(g)	(3,632)
Turkey	59	8	20	16	3	5	13
	(c)	(g)	(f)	(g)	(f)	(f)	(1,537)
South	29	5	4	8	6	0	6
	(e)	(g)	(f)	(g)	(j)	(f)	(4,345)
North	10	7	5	1	20	0	16
	(b)	(e)	(f)	(e)	(i)	(d)	(1,831)
All Applicants	32	11	15	17	11	9	13
	(g)	(i)	(i)	(i)	(k)	(h)	(11,345)

Note: To preserve anonymity, the letters in the municipality-specific parantheses denote the categories within which the corresponding number of observations falls: a (no observation), b (1-10 observations), c (11-30), d (31-70), e (71-150), f (151-310), g (311-630), h (631-1270), i (1271-2550), j (2551-5110), k (> 5111). In the other parentheses, the number of observations are shown.

The time trends indicate an impact of the 2003 ruling. This is especially visible

in municipality, which enforced extreme rejection rates before 2003. Turkish and Yugoslavian applicants often got rejected altogether in large numbers, their rates barely indistinguishable from one another. All applications were put on hold in 2003 and 2004; afterwards hardly any applications got rejected anymore. Through particularly striking there, this general observation is not restricted to municipality A. In 2003 and 2004, rejection rates fell in most municipalities, a break exploited in the next section's test for prejudice. In the years after, the rates began to rise again. This later increase coincides with an increase in the number of applications, which started around 2007. In contrast, the earlier years show a fairly consistent number of applications per year for any given combination of municipality and applicant group.

5 EMPIRICAL ANALYSIS

The decision by the courts in 2003 made local naturalization practices the center of public attention. Municipalities, now under close scrutiny, suddenly risked sharp criticism for the handling of their naturalization applications. Rejecting applicants became more risky. In terms of the model, the cost of mistakenly naturalizing unqualified applicants decreased. In equation (2), this exogenous shock raised the signal threshold $\theta^*(c, e)$ and thus decreased the average rejection rate in equation (4). In turn, this implies an increase in the naturalization rate, a hypothesis supported by the findings in Hainmueller and Hangartner (2017). Against that background, the test for prejudice proposed in Section 3.3 predicts that if a municipality is unprejudiced, the exogenous shock should not change the ranking of its rejection rates by applicant group.

I first test whether there is a statistically significant break in rejection rates upon the decision of the courts. To that end, for each applicant group in every municipality I calculate the aggregate rejection rates from 1998-2002 and 2004-2006, respectively. After 2007, the number of applications began to rise, which could indicate a change of quality

in the pool of applicants. As in Knowles, Persico, and Todd (2001) and Anwar and Fang (2006), aggregating over time increases the number of observations for the empirical test to a sufficient frequency. Both the number of applications and the rejection rates are stable within these respective periods, which increases our confidence that there are no changes in the quality of the applicant pools that could be confounding the results.

I refrain from using a parametric test, which is only valid if the systematic component and distribution of the error component of the model are correctly specified. A simple Pearson χ^2 test of independence yields more robust results. I calculate the p -values under the null hypothesis that for each applicant group in every municipality the rejection rates before and after 2003 are equal. That is, I compare the proportions of granted to rejected applicants against the proportion that would be expected under the null hypothesis:

$$\sum_S \sum_Y = \frac{\left(\widehat{O}(e, Y) - \widehat{E}(e, Y)\right)^2}{\widehat{E}(e, Y)} \sim \chi^2(1)$$

$S \in \{granted, rejected\}$ describes the two possible application outcomes, $\widehat{O}(e, Y)$ is the estimated observed average frequency of (granted or rejected) applications, $Y \in \{1998 - 2002, 2004 - 2006\}$ denotes the two time periods, and $\widehat{E}(e, Y)$ is the estimated expected frequency under the null hypothesis of independence. In this 2x2 contingency table, the test statistic will be drawn from a χ^2 distribution with one degree of freedom. The fifth and tenth column in Table 3, " p -value", show the levels of statistical significance from this test. Except in municipality E, I drop the northern group from the tests because of its low frequencies. In all tested cases, the rates decreased or remained constant after the break. The null is rejected in 11 out of 19 tests, suggesting that, by and large, the rejection rates decreased significantly during the three years after the ruling.

For each municipality, I proceed to test for statistically significant rankings between the applicant groups before and after the break, respectively. For these rank tests, I also employ the non-parametric Pearson χ^2 test of independence. In Table 3, the subpanels

in each municipality panel first test the null hypothesis that all applicant groups in that municipality exhibit equal rejection rates; more precisely, that the rejection rates are independent of applicant group. For the 1998-2002 period, the null that the rates are equal is rejected in all municipalities. During the second period, the null is rejected in four of the six municipalities. Taken together, this makes it unlikely that rejections are independent of applicant group. The next three lines (in municipality E: six lines) in each municipality's subpanel test the pairwise rank orders of each combination of the applicant groups. In municipality D, for example, the null hypothesis that in the 2002-2004 period the rejection rates of Yugoslavian (16%) and Turkish applicants (5%) are equal is rejected at a five percent level of significance.



Figure 1: Annual Rejection Rates in Municipality D, by Country of Origin

Let us take a closer look at municipalities D and C. Figure 1 visualizes the annual rejection rates by applicant group for municipality D. Throughout the years, the rejection rates by applicants from Yugoslavia are the highest among all the depicted groups. Applicants from the southern group consistently show the lowest rejection rates. Turkish rejection rates fall in between. There is a distinct visual rank order of the rejection rates, an order which does not change with the drop in 2003. Statistics support the visual impression. The subpanel for municipality D in Table 3 rejects the null for all three pairwise comparisons of the applicant groups in both periods. On the basis of the conceptual framework outlined in Section 3 we therefore cannot reject the null hypothesis of no prejudice in municipality D.

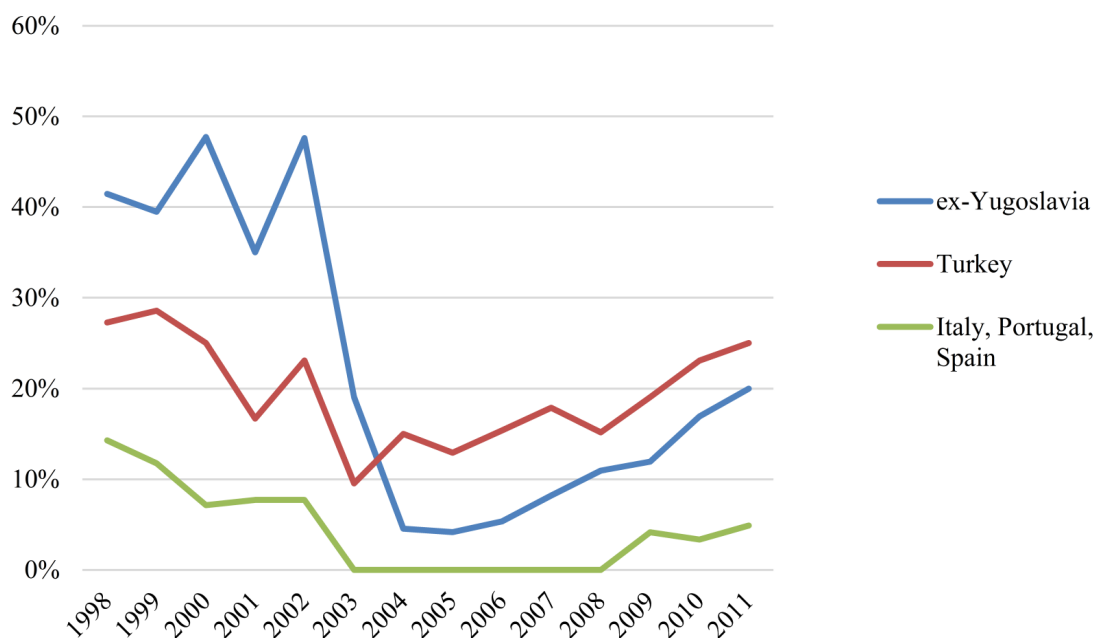


Figure 2: Annual Rejection Rates in Municipality C, by Country of Origin

Figure 2 pictures the development in municipality C. In contrast to the consistent rank order in municipality D, the figure reveals a reversal of the Yugoslavian and Turkish rank orders upon 2003. Table 3 shows that the rejection rate among applicants from Yugoslavia amounted to 42% before the break, only to drop to 5% afterwards. The Turkish rejection rate dropped as well, from 24% to 17%, but not enough to make up the sharp descent of the Yugoslavian rate. This rank reversal is statistically significant. The empirical test rejects the null that the rejections rates are equal before the break (after the break) at a one percent (five percent) level of significance. On that account, using the model we can conclude that, in this data set, there is evidence of prejudice in municipality C.

In a cross-sectional comparison of municipalities, we would neither be able to pinpoint the responsible municipality nor the discriminated applicant group. For the test for prejudice is a relative one, able to detect empirical patterns that are inconsistent with the assumption of unprejudiced municipalities. The within-municipality comparison upon the ruling of the courts overcomes both issues. The responsible municipality is obvious. And because the ruling raised the costs of prejudice, it is straightforward that, within the context of the model, municipality C used to be prejudiced against applicants from Yugoslavia before 2003.

The empirical results from the other municipalities do not point towards prejudice. Municipality A is an outlier in this analysis. The 2003 ruling practically brought about a complete reversal of the rejection rates, from all out to all in. But the data fail to provide a generally distinguishable rank order. Municipality B reveals an evident break only for the Yugoslavian rates, which consistently rank higher than the southern rates. In municipality E, we can conclude breaks for both the Yugoslavian and southern rate and can differentiate a combined rank order: Yugoslavian and northwestern applicants rank higher than Turkish and southern applicants in both periods. Finally, the Yugoslavian rate in municipality F ranks highest throughout. The Turkish rate hovers between the

Yugoslavian one and the southern one, the latter being zero in every year. Although rankable, none of them reveal a statistically significant break in 2003.

Table 3: Mean Rejection Rates in Percent and Rank Order Tests

Mun.	Group(s)	Rates and Ranks			Mun.	Group(s)	Rates and Ranks		
		98-02	04-06	p-value			98-02	04-06	p-value
A	Yu	82	11	<0.001	D	Yu	32	16	<0.1
	Turkey	90	0	<0.05		Turkey	22	5	<0.01
	South	43	0	<0.1		South	12	2	<0.05
	p-value			p-value					
	Yu=T=S	<0.001	0.68			Yu=T=S	<0.001	<0.01	
	Yu=T	0.51	0.71			Yu=T	<0.1	<0.05	
	Yu=S	<0.001	0.42			Yu=S	<0.001	<0.01	
T=S	<0.01	1		T=S	<0.1	0.24			
B	Yu	26	12	<0.05	E	Yu	20	15	<0.001
	Turkey	12	5	0.45		Turkey	4	0	0.27
	South	1	0	0.8		South	7	4	<0.05
	p-value			p-value					
	Yu=T=S	<0.001	0.17			Yu=T=S=N	<0.001	<0.001	
	Yu=T	0.10	0.32			Yu=T	<0.01	<0.05	
	Yu=S	<0.001	<0.001			Yu=S	<0.001	<0.05	
T=S	0.11	0.47		Yu=N	1	0.58			
C	Yu	42	5	<0.001	T=S	<0.1	0.24		
	Turkey	25	14	<0.1	T=N	<0.01	<0.01		
	South	10	0	<0.05	S=N	<0.001	<0.001		
	p-value			p-value					
	Yu=T=S	<0.001	<0.01		F	Yu	14	13	0.80
	Yu=T	<0.01	<0.05		Turkey	6	6	0.95	
	Yu=S	<0.001	<0.005		South	0	0	1	
T=S	<0.05	<0.01		p-value					
				Yu=T=S	<0.01	<0.1			
				Yu=T	<0.1	0.24			
				Yu=S	<0.01	<0.05			
				T=S	<0.01	<0.1			

Finally, the overall pattern of the data after the break displays a positive relationship between the number of applications and the rejection rate. Some years after the court ruling, more and more immigrants expressed interest in the Swiss passport. As the

number of applications rose, the rejection rate increased, too. This finding accords with the prediction of the model. The expansion of the applicant pool entailed a reduction in quality at the margin. With a fixed signal threshold for naturalization, this change implies a rise in the rejection rate. At the same time, it implies a rise in the naturalization rate; an implication which is consistent with the findings in Hainmueller and Hangartner (2017).

6 CONCLUSION

This paper proposes and empirically implements a test for prejudice in naturalizations based on a theoretical model of regionally deciding councils. The model uses rejection rates, classified by applicant group, as an indicator of prejudice. The test assesses the hypothesis that applicants striving for citizenship in their administrative division are measured by the same local yardstick. Using Switzerland as a case study, I collect an original data set on rejection rates from six large municipalities and exploit a landmark ruling in naturalization law in 2003, a ruling that exogenously raised the costs of rejecting applicants. According to the test, there is bias against applicants from former Yugoslavia in one of the six municipalities. In the other five, the test cannot reject the hypothesis of no prejudice. In addition, I provide empirical evidence consistent with a testable implication of the model, lending credibility to its descriptive validity.

One drawback in the existing research on naturalization is a lack of understanding about the mechanism driving the naturalization decision. That process has been looked on as a black box. My model aims to provide a generalizable comprehension of how rejection outcomes, and hence successful naturalizations, might be determined from the behavior of the deciding councils. The inference of eligibility for naturalization can take many forms, some of which are likely not observable to researchers. The implication is that even strong disparities in rejection rates when controlled for observable character-

istics may not prove prejudice. In contrast to a simple empirical analysis, the model helps understand how the observed outcomes have come about and provides a frame to interpret these data.

This may shed new light on existing conclusions of naturalization outcomes. Hainmueller and Hangartner (2017), for example, attribute the rise in naturalization rates in Switzerland after the 2003 ruling to discriminatory preferences by the the previous electorate. This conclusion emanates from the observation that for Turkish and former Yugoslavian applicants, the spike in naturalization rates was more pronounced in municipalities where the conservative Swiss People's Party (SVP) has higher vote shares. The model developed in this paper illustrates that this share might not necessarily be a "good proxy for xenophobic preferences" (p.12), preferences which manifest in the indulgence of prejudiced behavior against immigrants. Instead, the share could reflect a universally conservative state of mind. Just as one would hesitate to conclude that countries which are more reluctant to naturalize immigrants are prejudiced, the model allows for different levels of unprejudiced strictness among regions within a country. In the model this would imply that decisionmakers lower their required standard of qualification due to the increase in looming public accusation of prejudice. Because the conservative municipalities lower their bars from higher levels, they start letting through marginalized applicant groups at a higher rate than lenient municipalities, where the bulk of marginalized applicants had already cleared the bars. Other applicant groups would be less affected overall because their higher eligibility was never a real issue even in conservative municipalities. In this sense, this paper proposes an unprejudiced mechanism that is consistent with the same evidence.

In cross-sectional empirical implementations of the model, councils are not necessarily facing identical applicant pools. Researchers thus need to empirically verify that on average the pools of applicants faced by the different decisionmaking bodies are the same. This assumption could be ensured by testing variables that proxy for eligibility.

For example, data from the Program for International Student Assessment (PISA) would help answering the question whether students of a given country of origin have equal educational skill sets in different regions. Labor force survey or household panel data could be also used to verify similar eligibility.

I would also like to emphasize that, like in any empirical analysis with observational data, the test for prejudice proposed in this paper is only valid under certain assumptions. For example, in the model applicants do not interact strategically with the councils. One might object that the council's beliefs about the applicants might disincentivize marginalized groups to invest in integrational capital in the first place. Such models of self-fulfilling prophecy have been discussed by Arrow (1973) and Coate and Loury (1993). This is a limitation of this study. In ongoing research I address such an extension and show that under certain conditions, multiple equilibria may arise in which the branch of rank order tests to which my test belongs may not be valid (Ilić, 2016).

Finally, the conclusions about prejudice in this empirical analysis should be treated with mindful caution. The empirical data are novel and were raised manually, thus lacking rigorous validation and standardization. My hope is that the approach used in this paper will lead the responsible authorities to understand the types of data required to answer questions about prejudice in naturalization. Ideally, there would be a central register where naturalization applications, their subsequent decisions, and relevant individual information about the applicants are being filed. Another helpful instrument would be the introduction of complementary standardized naturalization achievement test nationwide, which would provide an additional measure for comparing the relative eligibility of the applicant groups.

7 APPENDIX

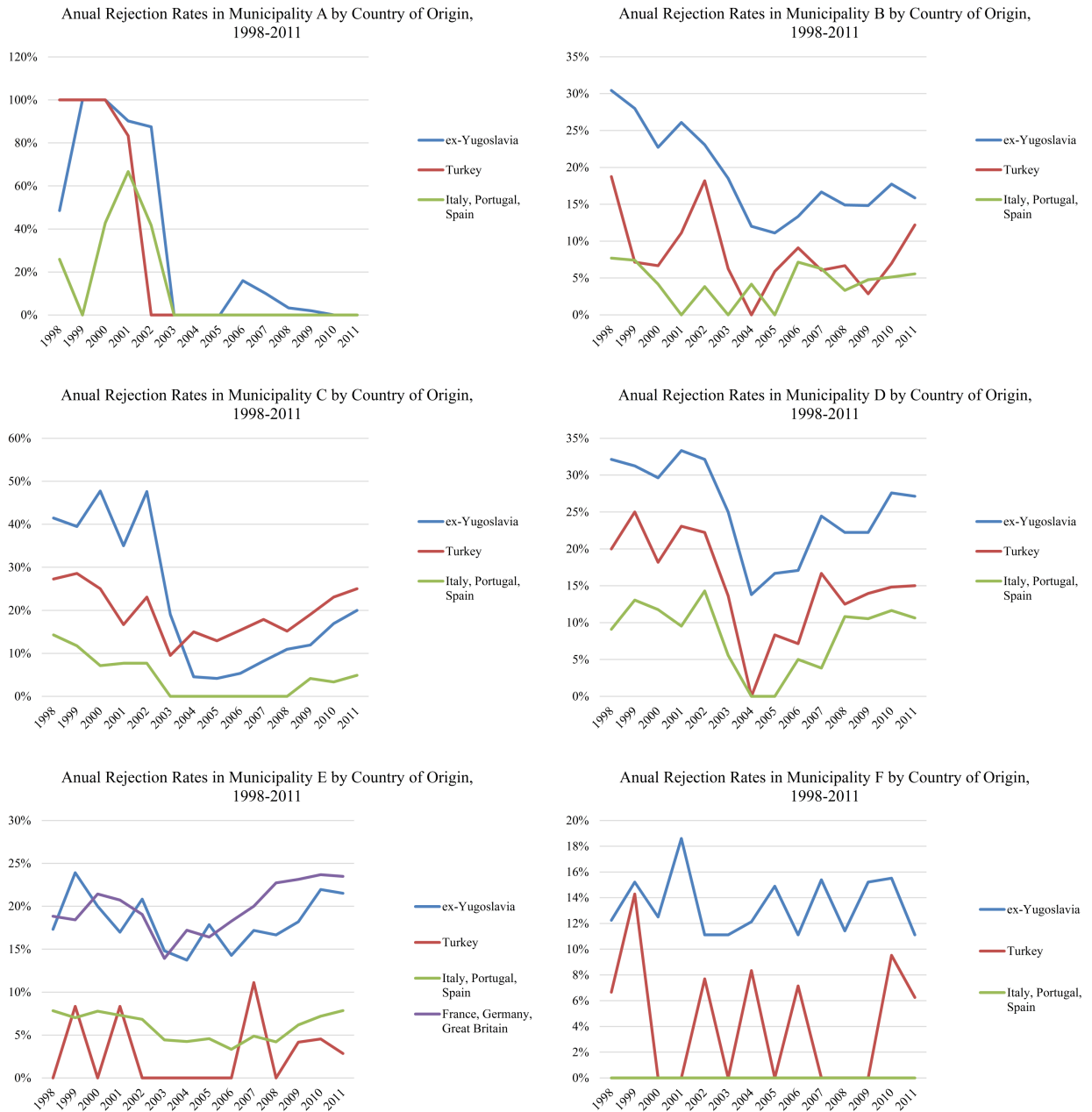


Figure A1: Annual Rejection Rates in all Municipalities, by Country of Origin

REFERENCES

- Ahmed, Ali M., and Mats Hammarstedt.** 2008. “Discrimination in the rental housing market: A field experiment on the Internet.” *Journal of Urban Economics*, 64(2): 362–372.
- Alesina, Alberto, and Eliana La Ferrara.** 2014. “A Test of Racial Bias in Capital Sentencing.” *American Economic Review*, 104(11): 3397–3433.
- Anwar, Shamena, and Hanming Fang.** 2006. “An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence.” *American Economic Review*, 96(96): 127–151.
- Anwar, Shamena, and Hanming Fang.** 2012. “Testing for the Role of Prejudice in Emergency Departments Using Bounceback Rates.” *The B.E. Journal of Economic Analysis & Policy*, 12(3): Article 4.
- Anwar, Shamena, and Hanming Fang.** 2015. “Testing for Racial Prejudice in the Parole Board Release Process: Theory and Evidence.” *Journal of Legal Studies*, 44(1): 1–37.
- Arrow, Kenneth J.** 1973. “The Theory of Discrimination.” In *Discrimination in Labor Markets.*, ed. O. Ashenfelter and A. Rees. Princeton, N.J.: Princeton University Press.
- Ayres, Ian.** 2002. “Outcome Tests of Racial Disparities in Police Practices.” *Justice Research and Policy*, 4: 131–142.
- Ayres, Ian, and Joel Waldfogel.** 1994. “A Market Test for Race Discrimination in Bail Setting.” *Stanford Law Review*, 46(5): 987–1047.
- Bauböck, Rainer, Eva Ersboll, Kees Groenendijk, and Harald Waldrauch.** 2006. *Acquisition and Loss of Nationality. Policies and Trends in 15 European States. Volume 1: Comparative Analyses.* Amsterdam University Press.

- Becker, Gary S.** 1993. “The Evidence Against Banks Doesn’t Prove Bias.” *Business Week*, 18.
- CERD.** 2008. “Concluding Observations of the Committee on the Elimination of Racial Discrimination.”
- Chiswick, Barry R., and Paul W. Miller.** 2008. “Citizenship in the United States: The Role of Immigrant Characteristics and Country of Origin.” *IZA Discussion Paper No. 3596*.
- Cinar, Dilek, and Harald Waldrauch.** 2006. “Acquisition and Loss of Nationality. Policies and Trends in 15 European States. Volume 2: Country Analyses.” , ed. Rainer Bauböck, Eva Ersboll, Kees Groenendijk and Harald Waldrauch, Chapter Austria, 19–62. Amsterdam.
- Coate, Stephen, and Glenn C. Loury.** 1993. “Will Affirmative-Action Policies Eliminate Negative Stereotypes?” *American Economic Review*, 83(5): 1220–1240.
- DeVoretz, Don J., and Sergiy Pivnenko.** 2005. “The economic causes and consequences of Canadian citizenship.” *Journal of International Migration and Integration*, 6(3): 435–468.
- Doleac, Jennifer L., and Luke C.D. Stein.** 2013. “The Visible Hand: Race and Online Market Outcomes.” *The Economic Journal*, 123(572): F469–F492.
- Dornis, Christian.** 2001. “Integrationspolitik in föderalistischen Systemen (Jahrbuch Migration).” , ed. Lale Akgün and Dietrich Thränhardt, Chapter Einbürgerung in Deutschland: Die Verwaltungspraxis in verschiedenen Regionen im Vergleich, 63–99. LIT.

- Dronkers, Jaap, and Maarten Peter Vink.** 2012. “Explaining Acces to Citizenship in Europe: How Citizenship Policies Affect Naturalization Rates.” *European Union Politics*, 13(3): 390–412.
- Ewens, Michael, Bryan Tomlin, and Liang Choon Wang.** 2014. “Statistical Discrimination or Prejudice? A Large Field Experiment.” *The Review of Economics and Statistics*, 96(1): 119–134.
- Fougère, Denis, and Mirna Safi.** 2009. “Naturalization and employment of immigrants in France (1968-1999).” *International Journal of Manpower*, 30(1/2): 83–96.
- Fryer, Roland G. Jr.** Forthcoming. “An Empirical Analysis of Racial Differences in Police Use of Force.” *Journal of Political Economy*.
- Hagedorn, Heike.** 2001. “Einbürgerungspolitik in Deutschland und Frankreich.” *Leviathan*, 29(1): 36–57.
- Hailbronner, Kay.** 2006. “Acquisition and Loss of Nationality. Policies and Trends in 15 European States. Volume 1: Country Analyses.” , ed. Rainer Bauböck, Eva Ersboll, Kees Groenendijk and Harald Waldrauch, Chapter Germany, 213–251. Amsterdam.
- Hainmueller, Jens, and Dominik Hangartner.** 2013. “Who Gets a Swiss Passport? A Natural Experiment in Immigrant Discrimination.” *American Political Science Review*, 107(1): 159–187.
- Hainmueller, Jens, and Dominik Hangartner.** forthcoming. “Does Direct Democracy Hurt Immigrant Minorities? Evidence from Naturalization Decisions in Switzerland.” *American Journal of Political Science*.
- Hainmueller, Jens, Dominik Hangartner, and Giuseppe Pietrantuono.** 2015. “Naturalization Fosters the Long-term Political Integration of Immigrants.” *PNAS*, 112(41): 12651–12656.

- Helbling, Marc.** 2008. *Practising Citizenship and Heterogenous Nationhood: Naturalisations in Swiss Municipalities*. Amsterdam University Press.
- Helbling, Marc.** 2010. "Switzerland: Contentious Citizenship Attribution in a Federal State." *Journal of Ethnic and Migration Studies*, 36(5): 793–809.
- Helbling, Marc, and Hanspeter Kriesi.** 2004. "Staatsbürgerverständnis und politische Mobilisierung: Einbürgerungen in Schweizer Gemeinden." *Swiss Political Science Review*, 10(4): 33–58.
- Ilić, Dragan.** 2016. "Self-Fulfilling Prophecies in Rank Order Tests." WWZ Working Paper 2016/05.
- Knowles, John, Nicola Persico, and Petra Todd.** 2001. "Racial Bias in Motor Vehicle Searches: Theory and Evidence." *Journal of Political Economy*, 109(1): 203–229.
- Lang, Kevin, and Jee-Yeon K. Lehmann.** 2012. "Racial Discrimination in the Labor Market: Theory and Empirics." *Journal of Economic Literature*, 50(4): 959–1006.
- Mechoulan, Stéphane, and Nicolas Sahuguet.** 2015. "Assessing Racial Discrimination in Parole Release." *Journal of Legal Studies*, 44(1): 39–74.
- Neumark, David.** 2013. "International Handbook on the Economics of Migration." , ed. Amelie F. Constant and Klaus F. Zimmermann, Chapter Ethnic Hiring, 193–213. Cheltenham, UK: Edward Elgar.
- North, David S.** 1987. "The Long Grey Welcome: A Study of the American Naturalization Program." *International Migration Review*, 21(2): 311–326.
- Parsons, Christopher A., Johan Sualaeman, Michael C. Yates, and Daniel S. Hamermesh.** 2011. "Strike Three: Discrimination, Incentives, and Evaluation." *American Economic Review*, 101(4): 1410–1435.

- Price, Joseph, and Justin Wolfers.** 2010. “Racial Discrimination Among NBA Referees.” *The Quarterly Journal of Economics*, 125(4): 1859–1887.
- Shayo, Moses, and Asaf Zussman.** 2011. “Judicial Ingroup Bias in the Shadow of Terrorism.” *The Quarterly Journal of Economics*, 126(3): 1447–1484.
- Steinhardt, Max Friedrich.** 2012. “Does Citizenship Matter? The Economic Impact of Naturalizations in Germany.” *Labour Economics*, 19: 813–823.
- The White House, Office of the Press Secretary.** 2013. “FACT SHEET: Fixing our Broken Immigration System so Everyone Plays by the Rules.”
- United Nations.** 2016. “Department of Economic and Social Affairs, Population Division. International Migration Report 2015 (ST/ESA/SER.A/384).”
- Weil, Patrick.** 2004. *Qu’est-ce qu’un Français?* Gallimard.
- Yinger, John.** 1996. “Why Default Rates Cannot Shed Light on Mortgage Discrimination.” *Cityscape: A Journal on Policy Development and Research*, 2(1): 25–31.
- Zhao, Bo, Jan Ondrich, and John Yinger.** 2006. “Why do real estate brokers continue to discriminate? Evidence from the 2000 Housing Discrimination Study.” *Journal of Urban Economics*, 59(3): 394–419.