

Refugee Roulette Revisited: Judicial Preference Variation and Aggregation on the Swiss Federal Administrative Court 2007-2012

Dominik Hangartner^{†‡}
Benjamin E Lauderdale[†]
Judith Spirig[‡]

[†]*London School of Economics* [‡]*University of Zurich*

ABSTRACT

Recent studies of asylum adjudication in several Western countries have found sizeable disparities between individual adjudicators. We contribute to this literature by exploiting a natural experiment from Switzerland, where all asylum appeals are handled by the Federal Administrative Court. Several features of the Swiss asylum appeal process conspire to offer an unusual opportunity to examine judges' revealed preferences and how they correlate with their party affiliation. First, the asylum cases have a common, uni-dimensional structure, as all decisions typically involve the appeal of an initial asylum decision. Second, the cases are assigned at random (conditional on language) to panels of judges, each of whom has a known party affiliation. As a result, we can test which of several decision- and game-theoretic theories of group decision-making seem to best fit the panel decisions as well as inferring the judges' individual preferences. We show that inconsistencies in decision-making due to panel composition were substantially reduced between 2007 and 2012, primarily because judges affiliated with the most liberal party converged towards the rest of the court.

Dominik Hangartner, Department of Methodology, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, and Institute of Political Science, University of Zurich, Affolternstrasse 56, 8050 Zurich. E-mail: d.hangartner@lse.ac.uk.

Benjamin E Lauderdale, Department of Methodology, London School of Economics and Political Science, Houghton Street, London WC2A 2AE. E-mail: b.e.lauderdale@lse.ac.uk.

Judith Spirig, Institute of Political Science, University of Zurich, Affolternstrasse 56, 8050 Zurich. E-mail: judith.spirig@uzh.ch.

The usual disclaimer applies.

I. INTRODUCTION

In 2007, among many other appeals the Swiss Federal Administrative Court had to decide two unconnected cases involving very similar applicants appealing the rejection of their asylum applications.¹ Both appellants were male Hazara Afghan, Shiite Muslims, from Kabul, and still had family living there. On July 4, a panel of three judges chaired by a member of the Social Democratic Party, ruled that the current situation in Kabul cannot be considered “reasonable” for removal and the asylum seeker has to be granted temporary protection. Two months later, on September 3, another panel consisting of three different judges chaired by an independent rejected the second appeal: the asylum seekers’ removal was deemed admissible and reasonable and he was thus given a short deadline by which he must leave Switzerland.

Based on the publicly available verdicts of the two appeals, there are no observable differences in the merits of the cases which could explain the different outcomes. Consistent adjudication is a central aim of any legal system², and so inconsistent decision-making across judges is a central concern for assessing the performance of such a system. Can the identity of the judges on the panel explain the differences in the two decisions? Do such disparities in asylum adjudication exist more generally? How do panel members aggregate their potentially diverse understanding of what the law requires, and does their decision-making become more consistent with repeated interactions? A sizeable body of literature has examined disparities in asylum adjudication at each stage of the process, from the initial application to the final appeal at the highest courts, in the U.S. (Ramji-Nogales, Schoenholtz and Schrag, 2007; Fishman, 2011) and Canada (Rehaag, 2007), but scholars still disagree about the prevalence of these disparities and measures to mitigate them.

Most of the asylum environments studied in past research have not featured random assignment of cases to asylum officers, judges or appeal courts.³ While the resulting studies often show large disparities in outcomes, the lack of quasi-random assignment of cases to decision-

¹Cf. FAC decisions E-3570/2006 and D-4576/2007.

²In Switzerland, for example, the due process clause of the Federal Constitution states that ‘very person has the right to be treated by state authorities (...) in a non-arbitrary manner’.

³One partial exception is the U.S. immigration court studied by Ramji-Nogales, Schoenholtz and Schrag (2007) where cases within a court are randomly assigned to judges, but assignment to one of the 53 courts is non-random and based on residence.

makers imposes important limitations on the inferences that we can draw. Another limitation is that many studies rely on data that is obtained with the permission of the governing institution, which typically replaces the name of the decision-maker with an anonymous identifier. Thus, correlating discrepancies in adjudication with judges' identities and characteristics (e.g. party membership) is often not possible. One previous study of asylum appeals (Fischman, 2011) leverages random assignment in asylum appeals at the US federal circuit court level to show that there is substantial inconsistency across judges, as well as evidence of suppressed dissents. We study heterogeneity in asylum appeals decisions, drawing upon a natural experiment from Switzerland that allows us to overcome some of the inferential challenges mentioned above and that illuminates the inconsistency of judicial decisions.

In Switzerland, an asylum seeker whose claim is initially rejected can lodge an appeal. Typically, appeals are lodged to challenge initial decisions of refugee status or pending deportation. Since 2007, all asylum appeals are handled by the Swiss Federal Administrative Court (FAC). Approximately thirty judges, belonging to two divisions of the FAC, deal exclusively with asylum appeals and collectively handle about 3,000 cases per year. Judges are elected to the FAC by the Swiss Parliament, have a publicly known party affiliation, and are typically backed by a party when running for office. Cases are assigned to panels of three judges in an automated procedure that renders the case merits independent of the judge's preferences, conditional on observable covariates. Cases are handled by one of two procedures. Under the "ordinary procedure", a case is jointly decided by a chair, a second, and a third judge. If no consensus can be reached, the outcome is decided by majority rule voting. Since 2008, the optional "simplified procedure" involves only the approval of the chair and the second judge. Under both procedures, we only observe the joint decision of the panel but never the judges' individual opinions. Drawing upon the FAC's online archive of all asylum appeal decisions, we collected a new dataset that contains case characteristics, panel composition and decision outcomes for the approx. 17,000 appeals from the court's inception in January 2007 to the end of 2012. We use these data to examine several questions: how panel composition affects the outcome of the appeal, the preferences of the judges' with respect to what threshold to apply to appeals, and the aggregation rule that combines their individual preferences into a joint

decision. We are then able to evaluate whether institutional efforts to improve consistency were successful over time.

The Swiss asylum appeal process has several features that allow us to overcome some of the inferential challenges typically associated with studying disparities in asylum adjudication (Ramji-Nogales, Schoenholtz and Schrag, 2007) and judicial behavior. First, all cases have a common, uni-dimensional structure since they exclusively deal with asylum issues.⁴ Second, the FAC processes all asylum appeal decisions in one federal office. Potential differences in the average merit of cases submitted to different regional offices typically undermine the comparability of estimated preferences across offices (e.g. Ramji-Nogales, Schoenholtz and Schrag (2007)). This cannot happen in our study where cases are sequentially assigned to the two asylum appeal divisions of the FAC. Third, cases are assigned to a panel of judges by an automated procedure that minimizes imbalances in judges' workload. As a consequence, the assignment of cases to judges is independent of the case's merit, conditional on the language of the asylum decision. The applicant's country of origin is the only, and readily observable, case characteristic that could be correlated with the language of the asylum decision if, for example, asylum seekers from the Ivory Coast are more likely to submit their application in a SEM processing center in the French-speaking part of Switzerland. After controlling for the appellants' country of origin, the appeals are effectively randomly assigned and differences in outcomes can be attributed to differences in panel composition. Fourth, there is no evidence that the distribution of cases changed when the optional "simplified procedure" was introduced in January 2008. This allows us to identify the effect of this change in the institutional rules and test whether the "simplified procedure" further increased the inconsistency in judges' decisions.

We find that in 2007, identical appellants face dramatically different grant rates depending on the panel of judges their case is assigned to. The heterogeneity in judges' preferences, which strongly correlates with their party affiliation in expected ways, does not affect all

⁴While the assumption that judges' preferences are dominantly unidimensional is often invoked in judicial politics, there is historical (Jeffries, 2001; Greenhouse, 2007) and statistical (Lauderdale and Clark, 2012) evidence that judges's preferences vary across areas of the law. In the context of our study, however, the unidimensionality assumption seems much more credible, as all decisions typically involve the appeal of an initial asylum or temporary admission decision.

appellants in the same way. For a number of appellants' countries of origin with typically low case merits, appeals are rarely granted, even by panels consisting of three liberal judges. For other countries, however the inconsistency of judges' decisions matters greatly. For asylum seekers from Afghanistan, a country with high grant rates over the study period, the expected probability of a successful appeal ranges between 15 % and 50 % solely as a function of panel composition. We estimate a variety of models using different aggregation rules to test which fits best. The best fitting simple aggregation rule is that only the chair judge's preferences matter, closely followed by a model in which decisions reflect the median judge's preference. We show that a game-theoretic model reflecting the decision procedure among the three judges fits better than either of these simple aggregation rules, indicating that the chair has some latitude to deviate from the median's preferences, so long as that deviation is not too large. We then fit models to examine the evolving pattern of decision-making on the court over time, after the introduction of the simplified procedure. We find little evidence that the introduction of the simplified procedure changed the patterns of decision-making, and some evidence that the chair's influence may have declined over time leading to a decrease in the proportion of inconsistent decisions.

Our study makes several contributions. First, our findings have important implications for the research that studies disparities in asylum adjudication. While most studies to date show sizeable disparities in asylum adjudication (Law, 2004; Ramji-Nogales, Schoenholtz and Schrag, 2007; Taylor, 2007; Fariss and Rottman, 2009) between decision-makers that face different cases, our results provide clear causal evidence that the identity of judges matters when facing cases with, in expectation, the same merits and that the estimated preferences strongly correlate with judges' party membership in expected ways. Second, our study adds to a growing literature on group decision-making (Sunstein, 2006; Fischman, 2007), preference aggregation (Fischman, 2011; Van Dijk, Sonnemans and Bauw, 2012) and the effects of social interactions of panel members (Kastellec, 2013; Fischman, 2013). Previous studies that infer individual preferences using IRT-type models rely on decision-makers' individual votes. Fischman (2011) presents a model of suppressed dissent that combines information about individual votes and random panel assignment. We explore a context where individual votes are not reported, but

leveraging randomization of panels we are able to recover individual preferences from joint decisions of panels by fitting a variety of aggregation rules. This framework has the potential for a wide range of applications since joint decisions without any recorded information on individual votes is the norm in a variety of contexts, e.g. search committees and management boards. Third, we add new empirical evidence to the growing literature on the effects of judges' identity on decisions (Ashenfelter, Eisenberg and Schwab, 1995; Sunstein, Schkade and Ellman, 2004; Peresie, 2005; Sunstein, 2006; Abrams, Bertrand and Mullainathan, 2008; Boyd, Epstein and Martin, 2010; Gazal-Ayal and Sulitzeanu-Kenan, 2010; Shayo and Zussman, 2011; Glynn and Sen, 2012; Posner, 2013) and show that the heterogeneity in preferences strongly correlates with party affiliation in expected ways.

II. ASYLUM APPEALS AND THE SWISS FEDERAL ADMINISTRATIVE COURT

Similar to most other Western countries, Switzerland grants asylum in accordance with the 1951 Geneva Convention relating to the status of refugees and the 1967 Protocol. All asylum applications are processed by the State Secretariat for Migration (SEM). Of the 26,000 decided asylum applications in 2012, only 15 % were granted (UNHCR, 2013), and more than 10 % of the rejection decision were appealed.⁵ If an initial asylum request is rejected by the SEM based on the merits of the case, the applicant may appeal within thirty days. Since its inception in January 1, 2007, all appeals are handled by the FAC.⁶ Decisions of the FAC are generally not appealable at the Swiss Federal Supreme Court: the FAC is the court of last resort in the Swiss asylum process.

When the FAC receives a new asylum appeal, it is assigned to one of the court's designated

⁵Viewed in a comparative perspective, Switzerland received a high share of asylum applications in recent years. In 2012 almost 26,000 applications were registered, making Switzerland the sixth largest receiving country (UNHCR, 2013) in terms of both the total number of applications (just below the U.K. but well above Canada) and relative to its population size (below Cyprus but above Luxembourg). This number, however, drops quite considerably if not the population size but the purchasing power parity-adjusted Gross Domestic Product per capita is used as denominator. With the numbers of asylum-seekers almost doubled between 2010 and 2012 (UNHCR, 2013), asylum policies have become heavily politicized in Switzerland, as right-wing parties like the Swiss People's Party have discovered the topic is an effective springboard for mobilizing voters against asylum-seekers and immigration more generally.

⁶Previously, decisions were handled by mostly the same judges but under the Swiss Asylum Appeal Commission which was part of the Federal Department of Justice and Police, the same department that also supervises the SEM.

asylum divisions and entered into the administrative system. All appeals are categorized according to the language of the asylum decision (one of the official Swiss languages German, French, or Italian) and forwarded to one of the division’s chambers. Subsequently, the case is assigned to a panel of three judges⁷—chair, second, and third—by a software program called ‘Bandlimat’—named after the first President of the FAC. The ‘Bandlimat’ assigns cases sequentially to judges, considering (i) language of the asylum decisions, (ii) the urgency of the appeal, (iii) judges’ language competencies, and (iv) current workload. The assignment of cases is completely mechanical and interference with the software’s assignment has to be justified, logged and entered by the president of the court. Each case receives an identical weight and the sole objective function of the software is to minimize the imbalance in workload created by the case assignment under the constraints (i)–(iii). Hence, we will treat this assignment procedure as if it were random conditional on the appellant’s country of origin which might be weakly correlated with the language of the decision.

In 2007, all cases were handled through the *ordinary* procedure that is characterized by the following structure: the chair judge receives the case files, makes additional investigations if necessary, assigns one of her clerks to draft a decision and forwards all materials to the second judge. The second judge reads the case files and the decision draft, either agrees or disagrees, proposes changes and forwards everything to the third judge. The third judge reads the case files and decision drafts and the second judges’ comments, either agrees or disagrees, proposes changes and returns the file and his comments to the chair. In case of disagreement, the panel further discusses and possibly revises the decision, if no consensus can be reached the outcome is decided by majority rule voting. Since January 1, 2008, the *simplified procedure* is applied to cases that are classified as either “clearly justified” or “clearly unjustified” by the chair judge according to the court’s guidelines. The initial assignment to a panel of three judges is the same as for the ordinary procedure, but if the chair judge invokes the simplified procedure and the second judge agrees with both classification and decision, the decision-making process ends here and the file is not forwarded to the third judge.

The approximately thirty judges belonging to the two divisions of the FAC handle on

⁷Cf. Article 21, Paragraph 1 of the Federal Law on the FAC

average about 3,000 cases per year. The judges are voted into office by the United Federal Assembly—a joint meeting of the two chambers of the Swiss Parliament—for a term of six years, with no term-limits. Most of the judges have a known party affiliation (a few run as independents) and in general, judge candidates are backed by a party when running for office. Judges’ party affiliation is not only visible for the members of the Swiss parliament but also clearly indicated on the FAC’s official website. Although not a written rule, it is customary for the party mix of judges to reflect the relative power of the different parties in the Swiss parliament, where the far-right Swiss People’s Party (SVP) has the most seats, followed by the left-wing Social Democratic Party (SP), the centrist Christian Democratic People’s Party (CVP), the center-right Liberals (FDP), and the leftist Green Party with only a few seats.⁸ Given its seat strength in the Unified Federal Assembly, the relative under-representation of right-wing judges is a recurring criticism raised by the SVP. In past years, media attention has frequently focused on the FAC’s asylum decisions and even individual judges⁹, either for being too liberal by right-wing newspapers¹⁰ or too restrictive by leftist pressure groups.¹¹

III. METHODOLOGY

To estimate individual preferences from the observed aggregate decisions of three judge panels, we need a framework for modelling preference aggregation. We adopt a 1D case-space model (Kornhauser, 1992), which allows us to theoretically describe the preferences of judges and to map different preference aggregation rules onto likelihood estimators. Each case j has *facts* that can be described as a location ψ_j . We treat smaller values of ψ_j as indicating stronger appeals (case facts), and larger values of ψ_j as weaker appeals. Each judge i has *preferences* that can be described as a cutpoint θ_i . Each judge, if deciding the case alone, would rule in favor of the appellant if and only if $\psi_j < \theta_i$. Thus, judges with lower cutpoints θ_i are inclined to grant fewer appeals, and judges with higher cutpoints are inclined to grant more appeals. An assumption of this unidimensional model is that all judges agree on the ranking of relative

⁸A few minor parties also have seats in the National Council or the Council of States but their vote share is too small to justify explicit representation in the Federal Courts

⁹<http://www.weltwoche.ch/ausgaben/2011-02/artikel-2011-02-zuwanderung-befangene-asylichter.html>

¹⁰<http://www.weltwoche.ch/ausgaben/2007-32/artikel-2007-32-richterliche-arr.html>

¹¹See e.g. http://www.freiplatzaktion.ch/static/media/medialibrary/2014/04/Rb_3-13.pdf

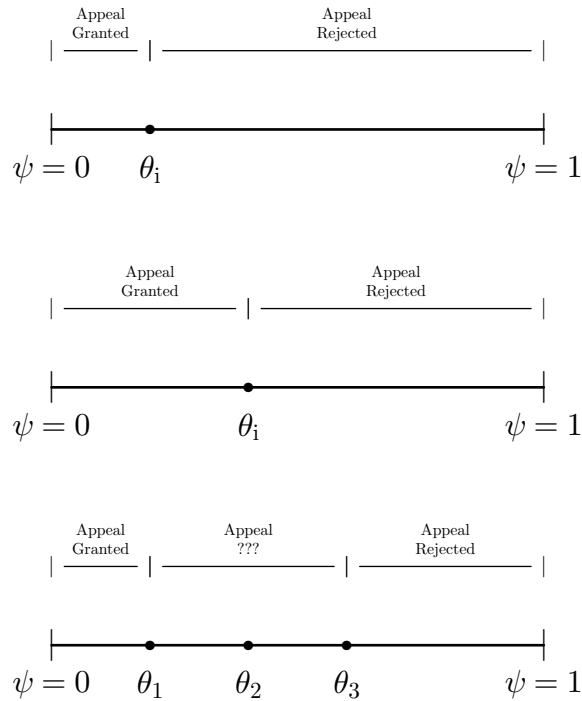


Figure 1: The top two axes illustrate the mapping between preferences and hypothetical single judge decisions. The bottom axis illustrates the three judge decisions that are actually observed, indicating the range of cases over which decisions depend on which preference aggregation rule most closely matches the court’s decision-making.

merits of appeals, and disagree only on threshold to apply. This has some implications for interpretation, which we discuss below.

The top two axes of Figure 1 show two different hypothetical judges, and the decisions they would make if they decided cases alone. However, as described earlier, the cases we are studying are decided jointly by three judges according to the procedures described in the preceding section, and so the resolution of cases in which the three judges disagree (bottom axis of Figure 1) depends on the aggregation rule that combines their preferences into a decision.

Let $i(j)$ be the indices of the judges hearing case j , so that $\theta_{i(j)}$ is a three component vector, with the first element $\theta_{i_1(j)}$ corresponding to the chair, the second $\theta_{i_2(j)}$ to the second judge, and the third $\theta_{i_3(j)}$ to the third judge. We consider only those aggregation rules that can be described by a function $f(\theta_{i(j)})$ that maps the preferences of the three judges into an

effective preference of the panel,¹² This allows us to define a generic likelihood function for the observable votes:

$$\mathcal{L}(\theta) = \prod_j p(\psi_j < f(\theta_{i(j)}))^{y_j} \cdot p(\psi_j > f(\theta_{i(j)}))^{1-y_j} \quad (1)$$

A. Decision-Theoretic Aggregation Rules

We consider a range of decision-theoretic preference aggregation models in our analysis. If we imagine the panel voting by majority rule internally, we expect θ_{med} to determine the outcome. If we imagine the panel voting with a requirement of a unanimity rule to grant an appeal, θ_{min} (the most conservative judge) determines the outcome. If instead unanimity is required to reject an appeal, θ_{max} (the most liberal judge) determines the outcome. If the chair's preferences dictate the outcome, we would expect $\theta_{1(j)}$ to determine the outcome. We fit models corresponding to these, as well as other (less plausible) aggregation functions, in our empirical analysis.¹³

B. A Game-Theoretic Aggregation Rule

We also consider a game theoretic preference aggregation rule that reflects the sequence of decision-making described earlier. As noted above, the chair of the panel initially receives the case files, reviews it, and sets out a provisional decision. Then, the second and third judges get the opportunity to review the file and provisional decision in turn, and only if there are unresolvable disagreements after further rounds of discussion is a decision by majority rule taken. This is clearly a costly process in terms of time and effort for all the judges concerned. The chair gets to frame the decision before the other judges even see the case, and there is an incentive for the second and third judges to follow the chair's decision rather than engaging in the effort necessary to determine if they disagree, let alone formulating an alternative to the

¹²It is difficult to make a substantive argument for the kinds of non-monotonic preference aggregation functions that could not be described as a mapping of the three judges' preferences into a single effective preference on the same scale.

¹³If we observe three judge panels, we cannot identify the preferences of certain judges under certain models. For example, if we assume the median judge's preferences determine the outcome, we cannot point identify the preferences of either the judge with the lowest or the highest threshold for asylum cases. For the minimum and maximum models, we cannot identify the two highest and the two lowest, respectively. However, we can identify which judges these are, and bound their θ with the next most extreme judge's position.

chair's provisional decision and potentially pushing the process to a majority vote.

One way of thinking about the incentives set up by this process is to consider what range of chair's decision rules $\theta_{i_1(j)}$ are close enough to θ_{med} to make deferring to the chair's decision attractive to the median judge. We assume that each judge must pay a cost δ in order to determine the case facts ψ_j , otherwise they have a standard uniform prior. We assume that each judge gets a payoff of -1 for decisions that are incorrectly decided according to their own preferences.¹⁴

We describe an equilibrium of this game in which the first judge (chair) pays the cost δ , writes a provisional decision based on the revealed value of ψ_j and her decision rule θ_g , and the second and third judges confirm that decision without paying the research cost δ . In this equilibrium, some of the decisions taken by the chair may be at odds with those of the median judge, but the other two judges do not know which provisional decisions they would like to reverse. Given that the court reverts to majority rule if there is sustained disagreement, to support such an equilibrium, the first judge must decide cases sufficiently similarly to how the panel median would decide cases in order to make the remaining judges not want to pay the cost of review. While we know qualitatively that in fact the second and third judges do occasionally overrule the chair, this simple oversight model and equilibrium highlight an important dynamic that enables the first judge to more strongly influence the outcome of the case, a dynamic that would be present in some form in a more nuanced oversight model as well.

Since we assume that all the judges know one another's relative preferences, the second and third judges will only consider a full review in the cases where the chair is more conservative than they are and proposes to reject the appeal, or where the chair is more liberal than they are and proposes to accept the appeal. If we denote the provisional decision of the chair as y_j^* , we can immediately state the condition that the chair's decision rule θ_g must satisfy in order

¹⁴Plausible alternative penalties for incorrectly decided cases, such as those that are increasing in the distance from the judge's preferred cutoff to the case facts, lead to equilibria that are observationally equivalent.

to support this equilibrium:

$$\delta > \begin{cases} p(\theta_g > \psi_j > \theta_{med} | \psi_j > \theta_{med}), & \text{if } \theta_1 > \theta_{med} \text{ \& } y_j^* = 1 \\ 0, & \text{if } \theta_1 = \theta_{med} \\ p(\theta_g < \psi_j < \theta_{med} | \psi_j < \theta_{med}), & \text{if } \theta_1 < \theta_{med} \text{ \& } y_j^* = 0 \end{cases} \quad (2)$$

That is, the probability that the present case would be decided differently under the equilibrium decision rule than under the median rule must be less than the added cost of disputing the chair's provisional decision. If we assume a uniform distribution of case facts ψ_j ,

$$p(\theta_g > \psi_j > \theta_{med} | \psi_j > \theta_{med}) = \frac{\theta_g - \theta_{med}}{1 - \theta_{med}} \quad (3)$$

$$p(\theta_g < \psi_j < \theta_{med} | \psi_j < \theta_{med}) = \frac{\theta_{med} - \theta_g}{\theta_{med}} \quad (4)$$

Since the chair would prefer to apply her own decision-rule to any given case, she will set $\theta_g = \theta_1$ if it satisfies this condition, and otherwise she will adopt the most extreme value that does. Thus, the equilibrium decision of the panel follows the decision rule:

$$\theta_g = \begin{cases} \theta_{med} - \delta(\theta_{med}), & \text{if } \theta_1 < \theta_{med} - \delta(\theta_{med}) \\ \theta_{med} + \delta(1 - \theta_{med}) & \text{if } \theta_1 > \theta_{med} + \delta(1 - \theta_{med}) \\ \theta_1, & \text{otherwise} \end{cases} \quad (5)$$

C. Case Facts Varying by Country of Origin

It is possible that non-random assignment by language might be related to the strength of a case. In order to account for this possibility, in some of our analysis we include country of origin as a covariate for predicting the strength of the case ψ_j . When we consider models that add covariate information, we continue to assume that $\psi_j \in [0, 1]$, however we model the distributions of the ψ_j on the unit interval. For such models, we assume a quadrilateral distribution on $x \in [0, 1]$ with probability density function $q(\psi|\xi) = 1 + \xi * (2\psi - 1)$, where $\xi_{C(j)}$ specifies a country-of-origin specific parameter determining the distribution of case facts from a given country relative to all other countries. The assumption of this distribution for case-attributes is not without loss of generality,¹⁵ however some functional form assumptions

¹⁵In particular, it constrains how extreme the distributions of case facts can be for countries relative to one another. For example, it cannot describe a situation where all the cases from one country are stronger than those from another country.

are necessary in order to allow the distribution of ψ_j to vary by country of origin for each case C_j . This quadrilateral distribution has the attractive theoretical feature that if the sum of the ξ weighted by the number of cases for each country is constrained to equal zero $\sum_C n_C \xi_C = 0$, the assumed distribution of case facts ψ_j across all cases remains uniform.

Since the country of origin is known to the judges, and immediately indicates something about the merits of the asylum case, specifying country-specific distributions of case facts also modifies equilibrium behaviour. Where $Q_{C(j)}(\psi) = \xi\psi^2 + (1-\xi)\psi$ is the cumulative distribution function of the quadrilateral distribution evaluated at ψ with parameter $\xi_{C(j)}$,

$$p(\theta_g > \psi_j > \theta_{med} | \psi_j > \theta_{med}) = \frac{Q_{C(j)}(\theta_g) - Q_{C(j)}(\theta_{med})}{1 - Q_{C(j)}(\theta_{med})} \quad (6)$$

$$p(\theta_g < \psi_j < \theta_{med} | \psi_j < \theta_{med}) = \frac{Q_{C(j)}(\theta_{med}) - Q_{C(j)}(\theta_g)}{Q_{C(j)}(\theta_{med})} \quad (7)$$

The equilibrium decision of the panel follows the decision rule:

$$\theta_g = \begin{cases} Q_{C(j)}^{-1}(Q_{C(j)}(\theta_{med}) - \delta Q_{C(j)}(\theta_{med})), & \text{if } \theta_1 < Q_{C(j)}^{-1}(Q_{C(j)}(\theta_{med}) - \delta Q_{C(j)}(\theta_{med})) \\ Q_{C(j)}^{-1}(Q_{C(j)}(\theta_{med}) + \delta(1 - Q_{C(j)}(\theta_{med}))) & \text{if } \theta_1 > Q_{C(j)}^{-1}(Q_{C(j)}(\theta_{med}) + \delta(1 - Q_{C(j)}(\theta_{med}))) \\ \theta_1, & \text{otherwise} \end{cases} \quad (8)$$

D. Measuring the Consistency of Decision-Making

Given the case space model, we can also calculate the extent to which random assignment of judges leads to inconsistency in the decisions that the court makes. Let $\tilde{\theta}_j$ be the consensus of the court, let $f(\psi)$ be the distribution of case facts, and let $\tilde{\theta}_j$ be the effective preferences of the judges hearing case j . Then, on average, the fraction of cases decided differently than how they would if the consensus were consistently applied, is:

$$\mathcal{E} = \frac{1}{M} \sum_{j=1}^M \left| \int_{\theta_j}^{\tilde{\theta}} f(\psi) d\psi \right| \quad (9)$$

To compute this quantity, we need a benchmark for what decisions an entirely consistent Court ought to make. We use the expected grant rate conditional on country-of-origin C_j where we include that information in the analysis, but *excluding* which judges heard the case:

$$\mathcal{E} = \frac{1}{M} \sum_{j=1}^M |\hat{\pi}_j - E[\pi | C_j]| \quad (10)$$

In analyses that do not use country-of-origin, $E[\pi|C_j]$ reduces to the mean appeal grant rate over all cases, and the inconsistency rate to the mean absolute error of the fitted values. This gives some intuition for why this measure captures inconsistency. If the set of judges hearing the case do not matter, then the fitted values (predicted probabilities of granted appeals) given the identities of the judges should not vary at all. To the extent they do, this indicates that the judges matter and cases are being decided differently depending on which judges hear those cases. This measure \mathcal{E} assumes again that judges only disagree about thresholds, not the relative merits of cases, but that means that \mathcal{E} provides a lower bound on the court’s inconsistency. If judges also disagree about merits, the true inconsistency of the court will be higher than what we estimate.

IV. DATA AND SAMPLE

Stipulated by Federal Law¹⁶, the FAC has to publish all its decisions. To construct our data, we wrote a script to automatically download all asylum appeal decisions from the FAC’s online database¹⁷ between 2007 and 2012. The next step consisted of extracting the verdict text using Python. In particular, we extracted the unique case id, the day of the decision, the panel composition and the role of the judges, the language of the case, the appellant’s country of origin and legal representation, and the verdict. We complement this database with personal information about the judges, most importantly their party membership, which we compile from the judges’ CVs on the official website of the FAC.

Overall, the sample—that is, the universe of Swiss asylum appeal decisions in the period under study—includes 16,967 decisions of which roughly 65 % are in German, 30 % in French and 5 % in Italian. To minimize coding error, we manually code the variables mentioned for a large set of verdicts so that we could cross-check the accuracy of our Python code. Based on several rounds of extensive testing, we are confident that the variables are correctly extracted and verdicts correctly classified. For more than 97 % of cases the binary classification of the verdict as either “granted” or “rejected” is straightforward. In the few remaining cases appeals are only partially granted with the consequence that the appellant does not get full refugee

¹⁶Cf. Article X, Paragraph X of the Federal Law on the FAC

¹⁷<http://www.bvger.ch/publiws/pub/search.jsf>

	LL	df	AIC
Game	-642.7	29	1343.5
Chair	-649.2	28	1354.4
Median	-655.8	28	1367.6
Min	-656.6	28	1369.3
Max	-666.5	28	1389.0
Third	-679.3	28	1414.6
Null	-711.1	1	1424.2
Second	-688.6	28	1433.3

Table 1: Table of fit statistics for MLE estimates of judge’s preferences in 2007 under different aggregation rules, sorted by AIC.

status but at least temporary protection. Since the asylum seeker does not face deportation in those cases, we decided to code them as “granted” as well.¹⁸

In 2007, the year when the court came into existence, the then 27 judges processed a total of 1,471 cases, all under the “ordinary procedure”. From 2008 onwards, the number of processed cases increases to about 3,000 in 2012. This performance increase is at least partially due to the introduction of the “simplified procedure” which is used to handle about 53 % of cases by 2012. From the asylum seeker’s vantage point, the average probability of winning an appeal is quite small, 19.5 % in 2007, and steadily declining throughout the study period to less than 12.0 % in 2012. There are 38 unique judges that serve during the period due to replacements.

V. RESULTS

A. 2007

We begin by fitting models to the data from 2007 using a variety of aggregation rules: minimum preference, median preference, max preference, chair’s preference, second judge’s preference, third judge’s preference, and the game theoretic model. Some of these rules are intended as placebo tests (we do not expect them to fit well) and therefore they provide us a check that our estimation approach has power against implausible alternatives. All of these models have 1 degree of freedom per judge, and the game-theoretic model has one additional degrees of freedom for the cost δ .

As Table 1 indicates, the best fitting simple aggregation rule is the chair judge, followed by

¹⁸If we change the coding for these cases to “rejected”, the estimates remain similar.

the median model; however the game-theoretic model fits better than these simple aggregation rules, even taking the additional degree of freedom into account. The estimate cost parameter $\hat{\delta} = 0.133$. Translated into the logic of the game, this implies that second and third judges are implicitly willing to accept 1 case decided contrary to their preferences so long as this allows them to avoid independently researching $1/0.133 \approx 7.53$ (or more) cases.

This is a theoretically compelling result, given the structure of the decision procedure followed by the court. Because the chair sees the case first, and writes the initial draft of the decision, he or she has an opportunity to frame the decision, while the second and third judges have an incentive to not investigate the case as thoroughly as they would if they were the chair. However, we do find clear evidence that the preferences of the second and third judge matter to some extent. If the other judges exerted no constraint on the chair, we would estimate a large value of δ and the chair model would fit as well as the game in terms of log-likelihood and better by AIC. A likelihood ratio test of the two models has $\Lambda = 12.93$, $df = 1$, and $p=0.0003$.¹⁹

In Figure 2, we show the estimated preferences of the judges under the game theoretic model both via MLE as described above and also via Bayesian mean posterior with a hierarchical beta prior on the judge’s preferences.²⁰ This Bayesian estimation approach is analogous to treating the judges’ preferences as random effects rather than fixed effects, which has the effect of reducing spurious variation in the judges’ preference estimates that might arise due to the limited number of cases that each judge decides. The judges are labeled by the party that nominated them, showing a clear political association between estimated preferences about how many asylum appeals to grant and political affiliation. The judges nominated by the Swiss People’s Party (SVP) are among the least favorable towards asylum seekers, while the judges nominated by the Social Democratic Party (SP) are among the most favorable towards asylum seekers, consistent with these parties’ general stances on asylum and immigration issues.

¹⁹However the assumptions of the test are not met with respect to the parameter δ , as null model that the chair determines the decision corresponds to large values of δ , for which the likelihood is flat with respect to changes in that parameter.

²⁰The hyperpriors on the parameters of the beta prior for the judges preferences are improper positive uniform. The prior on δ is standard uniform. We estimate the model via Hamiltonian Monte Carlo (HMC) using Stan (Stan Development Team, 2014).

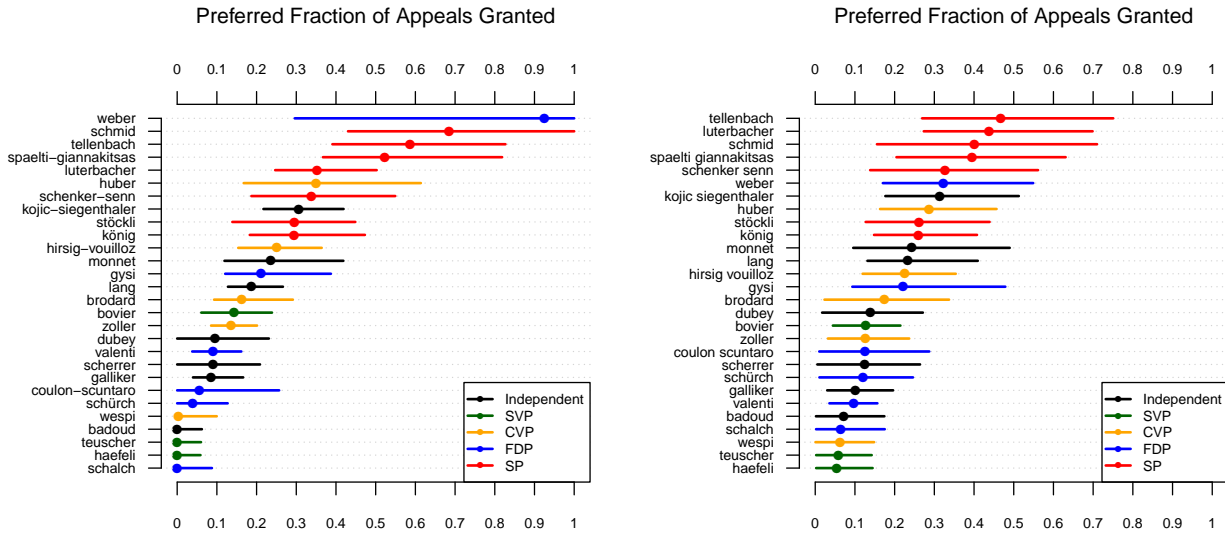


Figure 2: The estimated preferences of the judges in 2007, under the chair-median mixture model, estimated by MLE (left) and Bayesian mean posterior under a hierarchical prior (right).

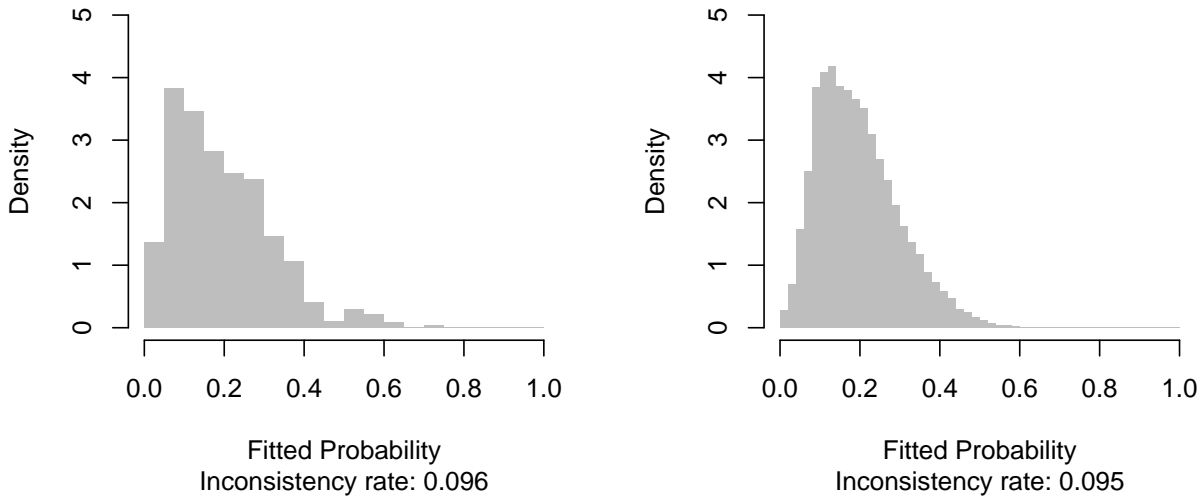


Figure 3: Predicted probability of successful appeal for each case under the game theoretic model, given the empirical compositions of the judicial panels in 2007, estimated by MLE (left) and Bayesian mean posterior under a hierarchical prior (right).

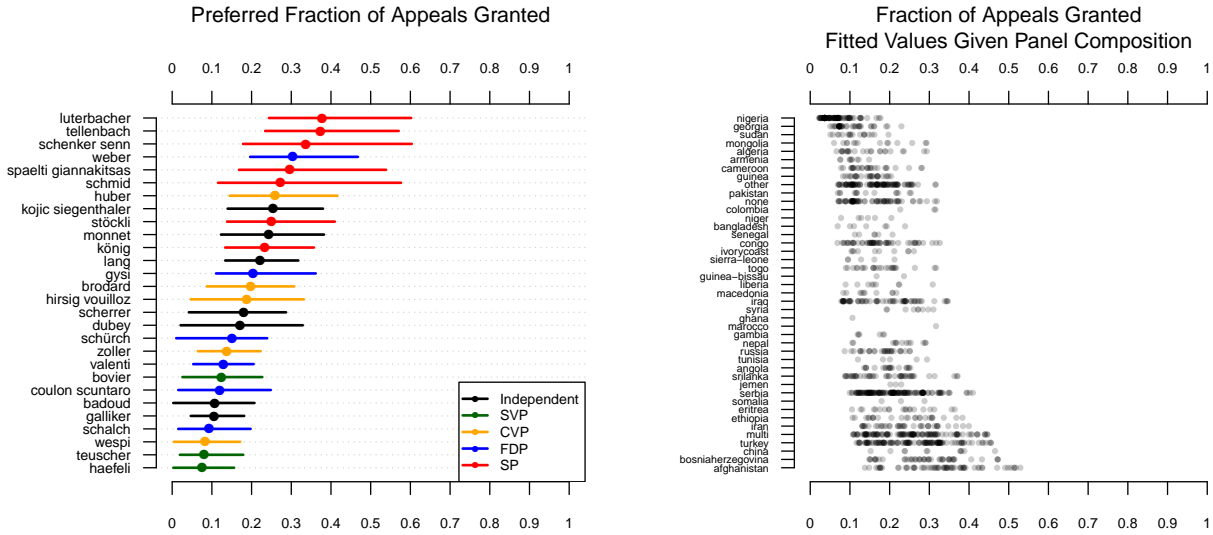


Figure 4: The left panel shows the estimated preferences of the judges when country of origin effects are estimated, and the right panel shows the variation in fitted values by country of origin.

In Figure 3 we show the fitted value distribution along with the estimated inconsistency rate for the court in 2007. The randomization of judges, combined with the preference variation estimated above, left some appeals with an expected probability of success of 5%, while others due to more favorable judges had a 60% chance of success. In total, this distribution implies that about 9.5% of cases heard by the court are being decided differently than they would if the median preferences of all 28 judges determined all cases.

To address the possibility that a language by country-of-origin interaction is leading to a violation of our as-if-random assignment assumption, we also estimated a version of the model where we estimate the population of case facts ψ_j for each country. To estimate this model, we exclusively use the Bayesian estimation procedure, as the number of observations for some countries is small and estimating the country distributions by MLE is problematic. As noted earlier, we assume $\psi_j \sim Q(\xi_{C(j)})$ where ξ_C are country-of-origin specific parameters determining the distribution of case facts from a given country C relative to all other countries, and are constrained such that the aggregate distribution of case facts ψ_j remains uniform.

As Figure 4 shows, the level of variation across judges is somewhat reduced when we introduce country-of-origin variation in the strength of cases, and we recover evidence of substantial variation in the success rates of appeals across countries. The mean posterior estimate for the fraction of cases decided differently than the median is 6.9%, with a 95% credible interval from 4.6% to 9.7%. The overall degree of variation in judges' preferences is somewhat lower than without controlling for country-of-origin, suggesting that some of the variation across judges was related to their language preferences and the relative strength of cases brought by applicants from different countries. The estimated inconsistency rate thus declines versus the MLE estimates presented earlier.

B. 2007-2012

After 2007, the introduction of the simplified decision procedure means that we do not observe the identity of the third judge for all cases. The cases where we fail to observe the third judge are a highly non-random subset: they are the cases the judges themselves deemed to be easy cases, almost all of which are rejections. However, our identification strategy relies on random assignment of cases to panels, and so to continue to analyze all cases, we need to find a symmetric approach to dealing with cases where either two or three judges' identities are observed.

In order to consistently estimate the preferences of the judges, we focus on the second-best fitting model from 2007, where the case is determined entirely by the chair's preferences. This model is a reasonable approximation of the game-theoretic model given the observed extent of preference variation and the estimated cost parameter δ . This model still focuses attention on the extent to which the court is able to achieve consistent outcomes despite having different judges deciding different cases.

Identifying change over time is difficult though, because both the strength of asylum appeals and the preferences of the judges could be changing over time. We fit a model where the preferences of the judges vary from year to year, as do the distributions of cases from each country. This means that we cannot make absolute comparisons of judges' preferences across years: the estimates from this model are with respect to the distribution of cases at a particular

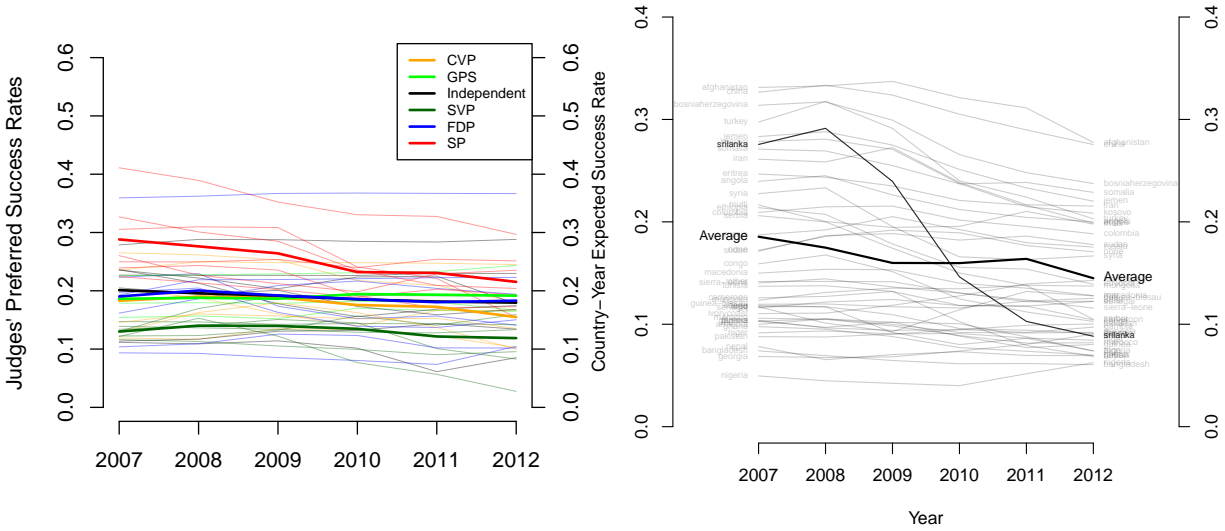


Figure 5: Preference estimate trajectories for all judges serving on the court from 2007-2012, with party mean trajectories (left). Expected probability of success by country of origin (right).

moment in time. Nonetheless, this enables us to determine whether consistency improved or worsened over time, although we will not be able to distinguish the hypothesis that the cases become systematically easier/harder to decided consistently from the hypothesis that the judges converged/diverged in their behavior. We put a random walk prior on the judges' preferences and the country-level parameters to smooth changes over time, and estimate this model on all data from 2007-2012 by HMC.

The left panel of figure 5 shows the trends in the judges' preference estimates over time. The general decline in asylum appeal success rates appears to be largely driven by the decline in success rates in cases chaired by judges from the Social Democratic Party, who initially were by far the most favorably to applicants but by 2012 were only slightly more favorable than the judges of the other parties aside from the Swiss People's Party.

The right panel of figure 5 shows the trends in the expected success rates by country over time. There is some decline and some convergence over time. The only country of origin exhibiting a large change in success rates is Sri Lanka, where success rates decline from about 25-30% to about 10% after the decisive defeat of the Liberation Tigers of Tamil Eelam (LTTE)

in 2009.

Figure 6 shows the distributions of fitted values relative to the court consensus, by year. Consistent with the convergence seen in the chairs' preference estimates, there is a general downward trend in the degree of inconsistency in court decisions from 2007 to 2012, with inconsistency rates dropping from 6.3% to 4.4% of all cases. This trend suggests that the court has either induced the more liberal judges of the Social Democratic Party to defer to the consensus of the court, or the increasingly acrimonious politics of immigration in Switzerland has achieved the same effect, with the result that the court has improved the consistency of its decision-making over time.

VI. CONCLUSION

Many studies of asylum adjudication are limited in their inferences by non-random assignment of cases to officials or courts and the unknown identity of the decision-maker. Using the population of asylum appeal decisions of the Swiss FAC where judges identity and party affiliation is known, our analysis demonstrates that judges' preferences in fact vary dramatically regarding the merits of cases and that they correlate strongly with party affiliation in expected ways. In particular, we show that in the year of the court's inception identical appellants faced dramatically different grant rates depending on the panel of judges their case was assigned to. Since cases were conditionally randomly assigned to judges, the disparities between judges cannot be explained by differences in case merits, and the ideal point estimates do not change when we control for appellant's country of origin. This disparity violates the very essence of the due process clause of Article 9 of the Swiss Constitution that stipulates that "every person has the right to be treated by state authorities in good faith and in a non-arbitrary manner".²¹ That said, the magnitude of the variation in expected success rates by panel composition is far smaller than has been found in previous studies in the US.

More concretely, for 2007 we find that an appellant whose case is assigned to a panel consisting of three of the most conservative judges from, say, the Swiss People's Party faces a dramatically higher rejection rate than if her case had been assigned to a panel of the three

²¹Cf. Article 9 of the Federal Constitution of the Swiss Confederation.

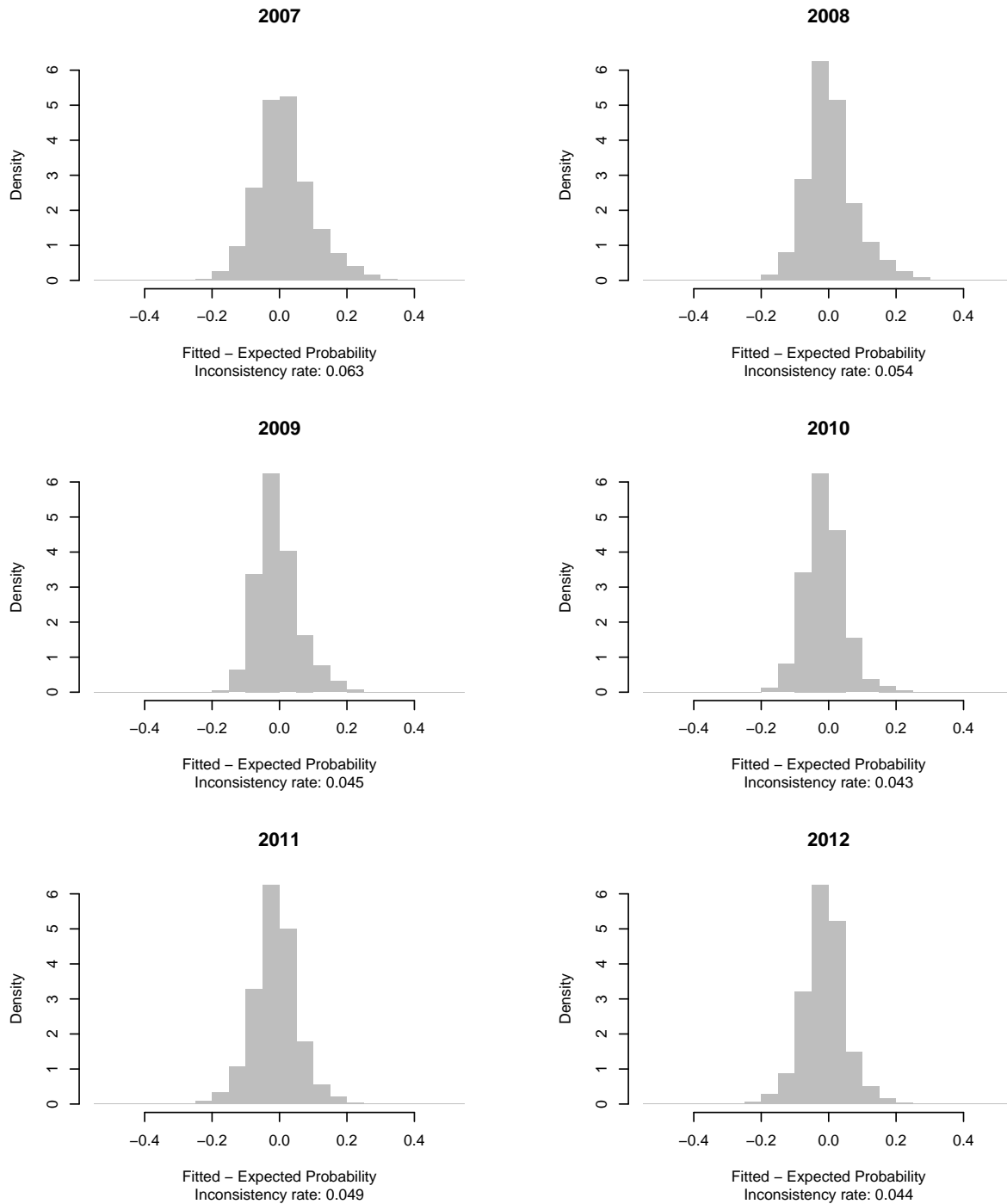


Figure 6: The distribution of fitted values minus expected values given country of origin, by year, and the corresponding inconsistency rate estimates.

most liberal judges from, say, the Social Democrats. This heterogeneity does not affect all appellants in the same way. For some countries, e.g. Nigeria, Sudan and Georgia, appeals are extremely rarely granted, even by the most liberal judge. For other countries, however, the inconsistency in judges decision matters more: for asylum-seekers from China (mostly of Tibetan ethnicity) and Afghanistan, the fitted success probability of an appeal ranges between 15% and 50% solely as a function of panel composition. Looking into how panels of judges make decisions and aggregate their preferences, we find that the chair of the panels is able to leverage the fact that they see each case before the other judges to have greater influence than their colleagues.

The court's institutional policy change to allow for effectively single judge decisions in post-2007 years did not further increase the heterogeneity in outcomes. To the contrary, we find a moderate decrease in disparities for the years 2007 – 2012. The inconsistency rate, measuring deviations from the court's average threshold, more than halves from 6.3% to 4.4% over this six year period. We have found no evidence that this is due to any qualitative reform of the decision-making process. Two features of the FAC are explicitly designed to increase consistency in decision making: (i) from time to time, it issues land-mark rulings that have *stare decisis*-like implications for subsequent decisions on similar appeals and (ii) judges from different chambers are assigned to serve on the same panel in non-urgent cases. However, the more straightforward explanation for the observed change is the moderation of the most liberal judges from the Social Democratic Party. These judges became less favourable to asylum applicants during a period when the right-wing Swiss People's Party was increasingly making immigration and asylum a major political issue.

REFERENCES

- Abrams, David S, Marianne Bertrand and Sendhil Mullainathan. 2008. Do Judges Vary in Their Treatment of Race? In *American Law & Economics Association Annual Meetings*. bepress p. 93.
- Ashenfelter, Orley, Theodore Eisenberg and Stewart J Schwab. 1995. "Politics and the judiciary: The influence of judicial background on case outcomes." *J. Legal Stud.* 24:257.
- Boyd, Christina L, Lee Epstein and Andrew D Martin. 2010. "Untangling the causal effects of sex on judging." *American journal of political science* 54(2):389–411.
- Fariss, Christopher J and Andy J Rottman. 2009. "The Political Economy of Decision Making Bias: An Analysis of Asylum Merit Hearings in US Immigration Court." *Available at SSRN 2080286* .
- Fischman, Joshua B. 2007. Decision Making Under a Norm of Consensus: A Structural Analysis of Three-Judge Panels. In *American Law & Economics Association Annual Meetings*. bepress p. 58.
- Fischman, Joshua B. 2011. "Estimating preferences of circuit judges: A model of consensus voting." *Journal of Law and Economics* 54(4):781–809.
- Fischman, Joshua B. 2013. "Interpreting Circuit Court Voting Patterns: A Social Interactions Framework." *Journal of Law, Economics, and Organization* p. ews042.
- Gazal-Ayal, Oren and Raanan Sulitzeanu-Kenan. 2010. "Let My People Go: Ethnic In-Group Bias in Judicial Decisions: Evidence from a Randomized Natural Experiment." *Journal of Empirical Legal Studies* 7(3):403–428.
- Glynn, Adam and Maya Sen. 2012. "Identifying Judicial Empathy: Does Having Daughters Cause Judges to Rule for Womens Issues?" *Working Paper* .
- Greenhouse, Linda. 2007. *Becoming Justice Blackmun: Harry Blackmun's Supreme Court Journey*. Macmillan.
- Jeffries, John Calvin. 2001. *Justice Lewis F. Powell, Jr.* Fordham Univ Press.
- Kastellec, Jonathan P. 2013. "Racial diversity and judicial influence on appellate courts." *American Journal of Political Science* 57(1):167–183.
- Kornhauser, Lewis A. 1992. "Modeling collegial courts. II. Legal doctrine." *Journal of Law, Economics, & Organization* pp. 441–470.
- Lauderdale, Benjamin E and Tom S Clark. 2012. "The Supreme Court's many median justices." *American political science review* 106(04):847–866.
- Law, David S. 2004. "Strategic Judicial Lawmaking: An Empirical Investigation of Ideology and Publication on the US Court of Appeals for the Ninth Circuit."

- Peresie, Jennifer L. 2005. "Female judges matter: Gender and collegial decisionmaking in the federal appellate courts." *Yale Law Journal* pp. 1759–1790.
- Posner, Richard A. 2013. *The behavior of federal judges: a theoretical and empirical study of rational choice*. Harvard University Press.
- Ramji-Nogales, Jaya, Andrew I Schoenholtz and Philip G Schrag. 2007. "Refugee Roulette: Disparities in asylum adjudication." *Stanford Law Review* pp. 295–411.
- Rehaag, Sean. 2007. "Troubling patterns in Canadian refugee adjudication." *Ottawa L. Rev.* 39:335.
- Shayo, Moses and Asaf Zussman. 2011. "Judicial ingroup bias in the shadow of terrorism." *The Quarterly Journal of Economics* 126(3):1447–1484.
- Stan Development Team. 2014. "Stan: A C++ Library for Probability and Sampling, Version 2.5.0."
URL: <http://mc-stan.org/>
- Sunstein, Cass R. 2006. *Are judges political?: an empirical analysis of the federal judiciary*. Brookings Institution Press.
- Sunstein, Cass R, David Schkade and Lisa Michelle Ellman. 2004. "Ideological voting on federal courts of appeals: A preliminary investigation." *Va. L. Rev.* 90:301.
- Taylor, Margaret H. 2007. "'Refugee Roulette' in an Administrative Law Context: The 'Déjà vu' of Decisional Disparities in Agency Adjudication." *Stanford Law Review* pp. 475–501.
- UNHCR. 2013. "Asylum Levels and Trends in Industrialized Countries – 2012." *Working Paper* .
- Van Dijk, Frans, Joep Sonnemans and Ed Bauw. 2012. "Judicial error by groups and individuals."