

Learning to be Unbiased: Evidence from the French Asylum Office

Mathilde Emeriau

London School of Economics and Political Science

Abstract: *What determines whether some asylum seekers are granted refugee status while others are rejected? I draw upon archival records from a representative sample of 4,141 asylum applications filed in France between 1976 and 2016 to provide new evidence on the determinants of asylum decisions. I find that applicants who are Christian (rather than Muslim) are more likely to be granted refugee status, controlling for all other individual characteristics available to the asylum officers making the decisions. However, linking archival records to detailed administrative data, I also show that bureaucrats at the French Asylum Office stop discriminating after about a year on the job. These findings have implications for strategies to curtail discrimination in courtrooms and administrations.*

Verification Materials: The materials required to verify the computational reproducibility of the results, procedures, and analyses in this article are available on the *American Journal of Political Science* Dataverse within the Harvard Dataverse Network, at: <https://doi.org/10.7910/DVN/HNVBZG>.

The arrival of over three million refugees in Europe between 2015 and 2019¹ triggered an unprecedented political crisis in the European Union (EU). Anti-immigrant parties campaigning on exclusionary policies are gaining ground across Europe. Several European countries reintroduced border controls, leading to the de facto suspension of the open-border Schengen area. The conflict among countries over the distribution and processing of the influx of asylum seekers is at the root of this crisis. All EU member states are bound by the Convention Relating to the Status of Refugees (hereafter the Geneva Convention), but despite decades of efforts to harmonize the European asylum sys-

tem, member states continue to handle asylum applications very differently (European Commission, 2016).

At the center of the asylum process, decision-makers in asylum courts and offices face the difficult task of determining whether asylum seekers are providing truthful and substantiated claims of persecution for reasons of “race, religion, nationality, membership of a particular group or political opinion.”² For asylum seekers, these decisions are very consequential: those granted refugee status will be able to stay in Europe, while those who are not have to return to their country of origin. However, the subjectivity of the definition of persecution and the possibility that applicants are falsifying claims

Mathilde Emeriau, Department of Government, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK (m.emeriau@lse.ac.uk).

I would like to thank Jens Hainmueller, Dominik Hangartner, David Laitin, and Stephane Wolton for their invaluable guidance in designing and implementing this study. I am also grateful to Pascal Brice, Aline Angoustures, members of OFPRA’s scientific committee, and the French National Archives for granting me access to the archives of the French Asylum Office. Adélaïde Choynet, Marie Gonzalez, Duncan Lawrence, Jean-Paul Levi, and Corinne Pallastrelli provided the technical support that made the data collection possible. Catherine Anderson, Astrig Atamian, Guillaume Beaud, Diane Belkheraz, Charles Bena, Lucie Benaiteau, Cécile Bénoliel, Sonia Bilon, Lisa Cavaliere, Emmanuel Charreau, Thomas Emeriau, Elyne Etienne, Jeanne Huybrechts, Anna Kauffmann, Shushan Kerovpyan, Maxime Krumholz, Alice Ledoux, Marion Lotin, Mathilda Loussert, Charles Maillard, Camille Urvoy, and Alice Zivohlava all provided excellent research assistance. This research is based upon work supported by National Science Foundation Grants No. 1624048 and 1627339 (Principal Investigator: David D. Laitin) the French National Research Agency (ANR) and the French government under the “Investissements d’Avenir” program LABEX LIEPP (ANR-11-LABX-0091, ANR-11-IDEX-0005-02) and the IdEx Université Paris Cité (ANR-18-IDEX-0001), the Stanford Immigration Policy Lab, The Europe Center, and the France Stanford Center for Interdisciplinary Research. Kirk Bansak, Sarah Brierley, Darin Christensen, Jennifer Cryer, Jane Esberg, Vicky Fouka, Saad Gulzar, Justin Grimmer, Jeremy Weinstein, and Yang-Yang Zhou all provided valuable comments.

¹Eurostat. “Asylum applications (non-EU) in the EU-27 Member States, 2008–2019.”

²Article 1(A)2 of the Geneva Convention.

American Journal of Political Science, Vol. 67, No. 4, October 2023, Pp. 1117–1133

© 2022 The Authors. *American Journal of Political Science* published by Wiley Periodicals LLC on behalf of Midwest Political Science Association. DOI: 10.1111/ajps.12720

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

gives judges and bureaucrats significant discretion to determine the veracity of their claims. On what basis do asylum officers decide? Is asylum granted based on the credibility of the claims of persecution, or do bureaucrats discriminate in the attribution of refugee status?

The lack of microlevel data has thus far limited researchers' ability to analyze the determinants of asylum decisions. In this study, I overcome previous data limitations by taking advantage of an unprecedented effort by the French Office for Refugee Protection and Stateless Persons (hereafter the French Asylum Office) to increase transparency. In 2009, the office opened its administrative archives and made nearly 1.5 million asylum applications available to researchers, which included hours of filmed interviews with some of the main actors in the French asylum system. I digitized a total of 4,141 asylum applications, a representative sample of all applications filed at the French Asylum Office between 1976 and 2016.³ I collected all the information recorded on the application form (e.g., age, gender, marital status, education, employment, and religion) and transcribed applicants' personal narratives, in which they explain why they need political asylum. Crucially, for each application, I also know the decision and the anonymized identifier of the bureaucrat who decided the case.

By comparing accepted and rejected applicants, I analyze how individual characteristics affect the probability of obtaining asylum in France, holding the country of origin and year of arrival constant. Since I had access to the same information as the bureaucrats who decided the cases, omitted variable bias is unlikely to be a concern. I employ three different strategies to control for the effect of the personal narrative. First, I combined hand coding and supervised machine learning to measure and predict whether the narrative contains credible claims of persecution. Second, I developed a list of substantive features of the text that *should* be relevant to the decision. Third, I applied the supervised Indian Buffet Process to a random subset of the data to identify the text features that *are* the most relevant in explaining the decision, which I then predict in the rest of the sample (Fong and Grimmer 2016).

This study yields two main findings. First, I demonstrate evidence of discrimination in the attribution of refugee status in France since 1976. While asylum decisions should only be based on assessments of the applicant's claims that they will be persecuted if forced

to return to their country of origin, I instead find that Christians are substantially more likely to be granted refugee status than Muslims. I show that these differences cannot be fully explained by variations in other individual characteristics or by differences in their personal narratives.

Moreover, these gaps are unlikely to be driven by differences in unobservable characteristics revealed during the interview process or by the selective assignment of particular cases to certain bureaucrats. Second, I show that French bureaucrats stop discriminating based on these characteristics after about a year on the job: Muslims are much more likely than Christians to be discriminated against when an inexperienced bureaucrat is assigned to their case than if an experienced case officer makes the decision. The pattern is similar, though less robust, for educational attainment and skill level. Overall, this suggests that bureaucrats learn on the job not to discriminate.

This study makes three main contributions. First, it contributes to the "refugee roulette" literature by analyzing the first microlevel dataset on a representative sample of asylum applications.⁴ Several studies have examined country-level data and concluded that both humanitarian and strategic interests explain variation in acceptance rates (Fariss and Rottman 2009; Holzer, Schneider, and Widmer 2000; Keith, Holmes, and Miller 2013; Neumayer 2005; Rosenblum and Salehyan 2004; Rottman, Fariss, and Poe 2009; Salehyan and Rosenblum 2008; Schneider and Holzer 2002; Toshkov 2014). However, these country-level analyses are merely suggestive since they do not account for the possibility that variation in the composition of asylum seekers between countries of origin could confound these results.

Second, this study provides credible evidence that on-the-job experience can mitigate discrimination using a fine-grained measure of experience. Arnold, Dobbie, and Yang (2017) uncover a similar pattern when comparing discrimination by bail judges across different courts in the United States who have different average levels of experience. Third, the study broadens our understanding of discrimination within bureaucracies. It adds to a growing number of empirical studies which demonstrate that bureaucrats discriminate on the basis of ethnicity and religion (Butler and Broockman 2011; Hemker and Rink 2017; McClendon 2016; Neggers 2018; Olsen, Kyhse-Andersen, and Moynihan 2020; White, Nathan,

³The French National Archives granted me access on October 9, 2015 (derogation JM/2015/1504) to 10,000 applications filed up to December 31, 2014 and on August 30, 2017 to 500 additional applications filed in 2015 and 2016. The names of the applicants and their family members were redacted to ensure anonymity.

⁴Holzer, Schneider, and Widmer (2000) and Holzer et al. (2000) control for some individual characteristics in their analyses, but these are limited to applicants' age, gender, and marital status. Keith and Holmes (2009) also analyze individual-level data but for a nonrepresentative subset of asylum applications.

and Faller 2015). Most notably, by linking self-reported religious affiliation with administrative decisions, a first in the French context, this dataset provides a unique opportunity to study religion-based bureaucratic discrimination in France using real-world data.

Asylum Process in France

To apply for refugee status in France, asylum seekers first need to fill out a standardized application form that elicits its demographic and socioeconomic information and to provide a personal narrative that describes, in French, their motives for seeking political asylum. Applications submitted to the French Asylum Office in Paris are directed to the relevant geographic division and assigned to a bureaucrat. Before 2006, this official decided whether to interview the applicant after reading her application; starting in 2006, interviews became mandatory. The bureaucrat advises his supervisor on whether to grant the applicant refugee status, and the supervisor makes the final decision. Those who obtain asylum receive a 10-year renewable residency permit (or a 1-year residency permit in the case of subsidiary protection, described below). Those who are rejected can appeal this decision to the National Court of Asylum, a three-judge panel that publicly reexamines asylum cases. If their appeal is rejected, asylum seekers can resubmit an application to the French Asylum Office if they have new information to provide.

What determines whether a person will receive political asylum in France? The Geneva Convention requires all signatories, including France, to grant asylum to individuals with a well-founded fear of being persecuted for reasons of “race, religion, nationality, membership in a particular group or political opinion.”⁵ In 1998, France introduced another form of protection called *territorial asylum*, which was replaced in 2003 by subsidiary protection, which is granted to those who do not meet the Geneva Convention’s definition of persecution but who would be subject to the death penalty, torture, or indiscriminate violence in the context of an internal or international armed conflict in their country of origin.

Bureaucrats working at the French Asylum Office have substantial discretion to decide who receives refugee status for two main reasons: (1) subjective interpretations of the Geneva Convention’s definition of persecution and (2) applicants who may be economic migrants and are falsifying persecution claims. As early as 1970, the asylum office’s annual activity report noted that “the

pace of arrivals remains high, although the Office is striving to exclude some elements — Yugoslav in particular — who are in reality, economic refugees, in search of better life and employment conditions.” Since the 1980s, when large-scale fraud by applicants from the Republic of Zaire and countries in South East Asia was revealed, bureaucrats at the French Asylum Office have been advised to use caution when assessing claims of persecution. A former director of the French Asylum Office (1996–2000) illustrated the challenge associated with making asylum decisions: “You are gold diggers. There is a huge stack of rocks. In this stack, there are a few gold nuggets. You have to find them, but there are a lot of rocks.”

Moreover, asylum officers often lack the time, space, training, and documentation needed to make informed decisions. The 1978 activity report accessed for this study describes their disastrous working conditions: “Four officers usually worked in a 12 square meter office with two typists typing, while the other two officers try to assist asylum seekers in the remaining space.” Information provided in the 1986 activity report suggests that asylum officers decided on average 3.5 cases per day, twice as many as their German counterparts at the time resolved. While their working conditions have improved over time, these bureaucrats still report working under very stringent time constraints. As recently as 2013, an asylum officer described how a lack of time and documentation significantly impaired her ability to “discover *the truth*” (Aho Nienne 2013).

Employment at the French Asylum Office has also been precarious. Until 1993, it had no permanent employees, relying mostly on temporary workers to deal with frequent fluctuations in the number of applications (Figure D.1, p. 18, in the online supporting information). In 1993, the unions successfully negotiated the conversion of temporary contracts into permanent positions. The share of temporary workers dropped radically afterward, but in the early 2000s, the office hired more to deal with a sudden uptick in applications.

In many respects, officials at the French Asylum Office resemble Michael Lipsky’s notion of street-level bureaucrats, who often respond to tight time constraints by developing “routines of practice and psychologically [simplify] their clientele and environment” (Lipsky 2010, xii) — a process that he argues can give rise to favoritism and stereotyping. While there is evidence from multiple contexts that bureaucrats in various capacities discriminate on the basis of race and ethnicity, this research overwhelmingly relies on experimental methods (Butler and Broockman 2011; Hemker and Rink 2017; McClendon 2016; Olsen, Kyhse-Andersen, and Moynihan 2020; White, Nathan, and Faller 2015);

⁵ Article 1(A)2 of the Geneva Convention.

very few studies have identified discrimination using real-world data (Neggers 2018).

In this study, I leverage unique access to archival records from France to study discrimination (i.e., unequal treatment) on the basis of religion, education, skill level, and proficiency in French of otherwise equal applicants. Prior research has shown that citizens in Europe and the United States tend to prefer immigrants who are Christian rather than Muslim, high skilled over low skilled, more educated, and who are fluent in the host-country language (Adida, Laitin, and Valfort 2016; Adida, Lo, and Platas 2019; Bansak, Hainmueller, and Hangartner 2016; Hainmueller and Hiscox, 2007, 2010; Hainmueller and Hopkins 2015; Helbling and Trautmann 2020; Valentino et al. 2017). Yet, these characteristics should not be relevant to decisions about whether to grant asylum, since such determinations should only be based on an assessment of the credibility of the claims of persecution.

Sources of Discrimination

Economists distinguish between two sources of discrimination to explain group-based differential treatment. Differential treatment arises in *taste-based* discrimination because decision-makers derive utility from favoring one group over another (Becker 1971). *Statistical* discrimination stems from decision-makers' beliefs about the distribution of the relevant outcome within a certain group (Arrow 1973). Bohren et al. (2019) and Phelps (1972) further distinguish between cases in which the latter is based on accurate beliefs (which they label *accurate statistical discrimination*) or inaccurate beliefs (*inaccurate statistical discrimination*).

I examine how these different sources of discrimination might affect French Asylum Office decisions as follows. A bureaucrat evaluates the level of persecution faced by an asylum seeker who has observable characteristics (religion, skill level, educational attainment, and proficiency in French) $g = \{A, B\}$ and faces persecution ω . The bureaucrat cannot directly observe the level of persecution ω that the asylum seeker faces at home; he must rely on the applicant's written narrative. The credibility of the narrative, $q = \omega + \epsilon$, is related to the level of persecution ω , but with some noise ϵ introduced by (1) the fact that the applicant needs to write about her experiences in French (which is unlikely to be her native language) and (2) the fact that the asylum seeker can get help from family, friends, and volunteers from migrants' associations. For simplicity, I assume $\epsilon \sim \mathcal{N}(0, \sigma^2)$ is an independent random shock. When reading the narrative,

the bureaucrat extracts a noisy signal of its credibility, $s = q + \eta$. The noise η comes from the fact that he can only imperfectly assess the narrative's credibility. I model the bureaucrat's ability η to do so as a random noise drawn from a normal distribution $\mathcal{N}(0, \sigma^2)$. The bureaucrat receives $-(v - (\omega - c_g))^2$ as the payoff from evaluation $v \in \mathbb{R}$ for an asylum seeker with characteristic g and persecution ω . The introduction of a taste parameter, c_g , in the bureaucrat's payoff function allows for the possibility that the official derives utility from favoring one group over another. Normalizing $c_A = 0$, a bureaucrat has a taste-based partiality for group A if $c_B \geq 0$. In that case, the bureaucrat has higher standards for asylum seekers from group B than for those from group A , even when they face similar levels of persecution. This opens up the possibility of taste-based discrimination.

Statistical discrimination instead arises from *how* bureaucrats reach evaluation v . I assume that bureaucrats have prior beliefs about the level of persecution faced by group g , which are distributed according to the normal distribution $\mathcal{N}(\hat{\mu}_g, \sigma_\omega^2)$. The bureaucrat has a belief-based partiality toward group A if he considers asylum seekers from that group to be more persecuted than those from group B ($\hat{\mu}_A \geq \hat{\mu}_B$). These beliefs can be accurate, in which case they are equal to the true mean μ_g , but they can also be inaccurate ($\hat{\mu}_g \neq \mu_g$).

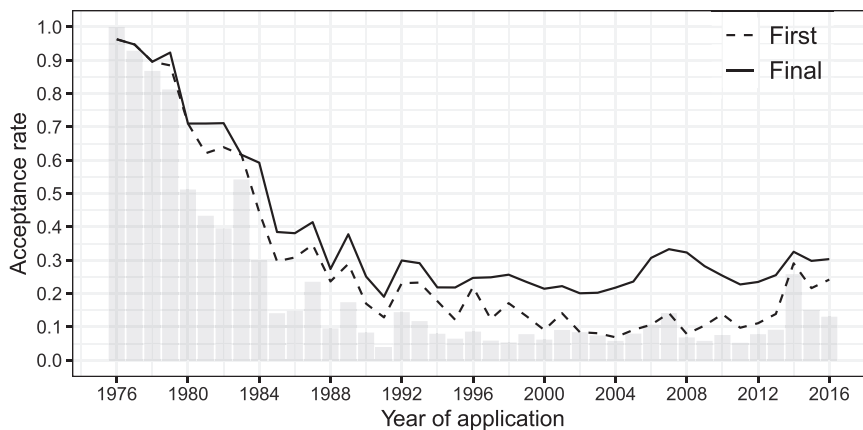
The bureaucrat updates his beliefs about the level of persecution faced by the asylum seeker after reading their narrative and chooses the evaluation v that maximizes his expected payoff with respect to this updated belief (see Appendix A, p. 3, in the online supporting information) for more details on the derivation of Equation 1).

$$v(s, g) = \frac{\sigma_\omega^2}{\sigma_\omega^2 + \sigma_\epsilon^2 + \sigma_\eta^2} s + \frac{\sigma_\epsilon^2 + \sigma_\eta^2}{\sigma_\omega^2 + \sigma_\epsilon^2 + \sigma_\eta^2} \hat{\mu}_g - c_g \quad (1)$$

Equation (1) provides an important insight for the empirical strategy. To estimate the extent of discrimination involved in decisions about whether to grant refugee status, it is important to control for the signal s that the bureaucrat receives to distinguish between the signal effect and the discrimination effect. In practice, however, I can only estimate q , the credibility of the narrative, since I do not observe η . Moreover, the estimation inevitably introduces some noise λ . As a result, I control for $s' = q + \lambda$. However, as long as λ and η are uncorrelated with group characteristics, I can recover the effect of individual characteristics on the decision. I return to this point in Section 5.

Discrimination occurs when two asylum seekers with the same signal, one from group A and one from

FIGURE 1 Acceptance Rate at the French Asylum Office (1976–2016)



Notes: This figure plots the acceptance rate on the basis of the Geneva Convention and the subsidiary protection after the first examination at the Asylum Office (dashed) and after examination by the appeals court and subsequent reexaminations by the Asylum Office (solid). The gray bars plot the acceptance rate based on the Geneva Convention among applicants in the sample.

group B , receive different decisions. Assuming $c_A = 0$,

$$\begin{aligned} D(s) &= v(s, B) - v(s, A) \\ &= \frac{\sigma_\varepsilon^2 + \sigma_\eta^2}{\sigma_\omega^2 + \sigma_\varepsilon^2 + \sigma_\eta^2} (\hat{\mu}_B - \hat{\mu}_A) - c_B. \end{aligned} \quad (2)$$

Equation (2) illustrates that disparate decisions, for example, in favor of group A , can arise from two sources: taste-based discrimination, which results from the official's preference for group A ($c_B \geq 0$), or statistical discrimination, the result of the bureaucrat's belief that group A is more persecuted on average than group B ($\hat{\mu}_B - \hat{\mu}_A \leq 0$).

Data

The sample for this study consists of 4,141 asylum applications filed at the French Asylum Office between 1976 and 2016. The sampling design was complicated by the fact that the office regularly destroys a subset of rejected applications to free up space in the archives. To deal with this issue, I proceeded in two steps. The French Asylum Office first extracted a random sample of 100,000 applications of all those filed since 1952—25,118 of which had already been destroyed. I then selected applications from among those remaining for in-depth data collection using the inverse of the probability of still being in the archives as the selection probability. To correct for remaining imbalance between the data collected

and the administrative sample, I used entropy balancing to reweight the sample for the analyses (Hainmueller 2012). Appendix B (pp. 4–9 in the online supporting information) contains more information on the sampling design and the construction of the weights.

The outcome of interest is whether an applicant was granted refugee status upon first examination at the French Asylum Office. On average, 15.3% of applicants in the sample were granted political asylum in France based on the Geneva Convention between 1976 and 2016 (Figure 1, gray bars). By construction, this rate is lower than the overall first-time acceptance rate at the Asylum Office (Figure 1, dashed line) for two reasons. First, I exclude cases of family unification and resettled refugees from the sample because these cases are not evaluated exclusively on the basis of their application. Second, I focus on the decision to grant refugee status based on the Geneva Convention, which presents a unique feature crucial to the research design. Indeed, the Geneva Convention is applied in France based solely on examining individual claims of persecution rather than on belonging to a particular group (Cohen 2000). This practice rules out the possibility that the observed group differences may be driven by the fact that bureaucrats know that some groups are more persecuted, since in theory this information should not matter. The same assumption does not hold for the attribution of the subsidiary protection since it is based, among other things, on the security situation in the applicant's region of origin, which helps determine whether she needs protection.

In France, asylum cases are examined sequentially, thus only those that did not qualify for protection under the Geneva Convention are examined for subsidiary protection. I thus recode applicants who received refugee status under subsidiary protection as having been rejected for the Geneva Convention protection.

Table 1 lists the main independent variables used in the analysis. In this study, I focus on identifying discrimination on the basis of religion, education, skill level, and language proficiency. Overall, close to 80% of applicants identify as either Christian or Muslim, but there is significant variation over time: Muslims represented 60% of applicants in 2016, up from 5% in 1976. Half of all applicants had started secondary education, and 14.2% reported starting university. Highly skilled jobs including academics, doctors, engineers, high-level executives, lawyers, journalists, and students represent 16.6% of the applicants. Blue-collar workers, civil servants, mechanics, farmers, guards, and other such employees represent 39.7% of applicants (*Middle*), and drivers, hairdressers, sales clerks, and homemakers 18.6% (*Low*). An applicant was coded as proficient in French if she listed French as a native or spoken language (26.8% of applicants did). In all specifications, I also control for age, gender, marital status, and time spent in France at the time of the application.

For each application, I also transcribed the personal narrative, the bulk of which were handwritten. Of the 4,141 applications included in the final sample, 93.4% submitted a narrative in French. Table 2 provides summary statistics on a number of features of the narratives that, based on my informal discussions with bureaucrats at the French Asylum Office, should affect how credible they are perceived to be. The number of words (including stop words) and the number of dates and locations mentioned (extracted using Stanford CoreNLP) are proxies for the narrative's level of detail. The narratives varied greatly in length, but on average were 777 words long and mentioned seven dates and seven locations. To understand the extent to which the narratives are personal and individualized, I counted the number of first-person pronouns used ("je," "j'ai," "me," "mon," "mes," and "moi"). On average, the narratives used first-person pronouns 35 times. To control for the originality of the narrative, I also computed each narrative's average Euclidean distance to the narratives of other asylum seekers from the same country of origin. Moreover, to gain insight into the topics covered in the narratives, I also estimated a structural topic model with 20 topics using country of origin as a covariate (Roberts et al. 2014). See Appendix C (pp. 9–17 in the online supporting information) for more information on the construction of these features.

TABLE 1 Summary Statistics on Dependent and Independent Variables

	Mean	SD	Min	Max
Got refugee status	0.153	0.360	0	1
Provided a narrative	0.934	0.248	0	1
Gender (female)	0.310	0.462	0	1
> 1 year in France	0.150	0.358	0	1
Age				
Under 20	0.091	0.287	0	1
Between 20 and 40	0.772	0.420	0	1
Over 40	0.138	0.345	0	1
Marital status				
Single	0.582	0.493	0	1
Married	0.418	0.493	0	1
Religion				
Christian	0.375	0.484	0	1
Muslim	0.397	0.489	0	1
Other	0.122	0.328	0	1
None/Missing	0.105	0.307	0	1
Education				
University	0.142	0.349	0	1
Secondary	0.495	0.500	0	1
Primary	0.112	0.315	0	1
None/Missing	0.252	0.434	0	1
Skill level				
High	0.166	0.372	0	1
Middle	0.397	0.489	0	1
Low	0.186	0.389	0	1
None/Missing	0.251	0.434	0	1
Speaks French				
Yes	0.268	0.443	0	1
No/Missing	0.732	0.443	0	1

Notes: All variables, except the first two, are self-reported. The number of observations is 4,141, except for marital status (N = 4,140).

For confidentiality reasons, I was not authorized to collect any individual-level information about the bureaucrats working at the French Asylum Office. However, I was able to construct a fine-grained measure of their experience using administrative data on a large random sample of asylum applications. To do this, I complement the list of 100,000 I used to sample asylum applications for archival data collection with an additional random sample of 500,000 applications filed between 1989 and

TABLE 2 Summary Statistics on Narrative Features

	N	Mean	SD	Min	Max
Level of detail					
Number of words	3,932	777.358	766.394	1	7,169
Number of dates	3,932	7.278	8.314	0	95
Number of location	3,932	6.540	7.511	0	84
Degree of personalization					
Number of personal pronouns	3,760	34.564	33.135	0	445
Originality					
Distance	3,927	47.125	23.355	6	347
Topic proportion > .2					
Family members	3,930	0.088	0.283	0	1
Bangladesh	3,930	0.038	0.192	0	1
Student protest	3,930	0.074	0.261	0	1
Escape	3,930	0.092	0.289	0	1
Sri Lanka	3,930	0.040	0.196	0	1
RDC	3,930	0.043	0.203	0	1
Life in France	3,930	0.224	0.417	0	1
Former Yugoslavia	3,930	0.078	0.269	0	1
China	3,930	0.040	0.197	0	1
Ethnic minorities	3,930	0.052	0.223	0	1
Kurds in Turkey	3,930	0.067	0.250	0	1
Political instability	3,930	0.049	0.215	0	1
Political opposition	3,930	0.055	0.228	0	1
Court hearing	3,930	0.020	0.140	0	1
Narrative	3,930	0.150	0.357	0	1
Angola	3,930	0.056	0.231	0	1
Salutation	3,930	0.059	0.235	0	1
Family at risk	3,930	0.043	0.202	0	1
Zaire	3,930	0.034	0.181	0	1
Religion	3,930	0.021	0.143	0	1

Notes: This Table provides summary statistics on the narratives' substantive features used as controls in Table 3, column 2.

2014. Together these two administrative samples provide information on 502,997 unique asylum applications filed between 1989 and 2014 — more than half of all applications filed at the Asylum Office during that period. The information contained in this database is scarce, but for the 309,913 applications filed after 2000, these records include the decision, as well as the date and the identifier of the bureaucrat who made the decision. After excluding bureaucrats who made at least one decision in 2000 (and for whom I cannot rule out the possibility that they started working for the office before that year) and those who made their first decision in 2015 or after, I can infer the start date of each remaining bureaucrat using the date of their first decision and computing the order in

which they made each of their decisions. I am then able to match this information with the main sample, allowing me to use the number of past decisions made by the bureaucrat in charge of the case as a proxy for his level of experience at the time of the decision.

Research Design

This unique individual-level dataset allows me to estimate discrimination in the attribution of refugee status in France. I regress an indicator variable for whether an applicant was granted asylum on her characteristics, while holding constant the credibility of her narrative,

her country of origin, and the year of application. The selection on observables assumption is well supported because I have access to the same information as the bureaucrats making the decisions. However, the fact that interviews became mandatory in 2006 raises the concern that some decisions could have been based on unobservable characteristics revealed during the interview. Fortunately, close to 40% of the applicants in the sample were not interviewed, which allows me to check that the results hold in this subsample. Moreover, while we do not know the exact mechanism used to assign cases to bureaucrats, we have good reason to believe it is based on observable characteristics: Cases are first dispatched to different divisions, each of which handles a different set of nationalities. The division head then assigns cases to the officials in their division on a monthly basis. Since this assignment happens before the interview takes place, we can assume that this assignment is based exclusively on observable characteristics. To address this concern, I show that the results hold when controlling for bureaucrat fixed effects. Finally, one necessary condition for the estimation is overlap, which is often difficult to achieve with highly collinear variables like religion and country of origin. I check for overlap between religion and country of origin in the data by plotting countries of origin in the sample as a function of the share of applicants from that country who are Christian on the x -axis, and the proportion who are Muslims on the y -axis (Figure D.2, pp. 19, in the online supporting information). While a number of countries cluster on the bottom-left and upper-right corners of the graph, a nonnegligible proportion is spread out on the diagonal, suggesting substantial overlap.

The efficacy of this research design also hinges on the ability to control for the signal s that the bureaucrat receives from the narrative, as outlined in Section 3. Moreover, the descriptive patterns reveal that the narratives differ significantly by religion, education, skill level, and proficiency in French, further emphasizing the need to control for the narrative's effect on the decision (Tables C.2, C.3, C.4, and C.5, pp. 11–12 in the online supporting information). I use three different methods to control for the narrative. First, in the main specification, I control for the credibility of the narrative, which I predict using a combination of human coding and automated classification. Due to the limitations of this method (outlined below), I check that the results are robust to two alternative strategies: controlling for substantive features of the narrative that *should* be relevant in predicting the decision (listed in Table 2) and controlling for features that *are* relevant to the decision (using the supervised Indian Buffet Process). It is important to note here that I use these measures to control for the credi-

bility of the narrative, q , not the applicant's level of persecution, ω . But this is not a concern since bureaucrats also only observe a signal of persecution, not the persecution itself, before making their decision. I now discuss all three methods in turn.

The first method I use to control for the narrative relies on a combination of hand coding and supervised learning to measure the credibility of the personal narrative — that is, the extent to which the narrative (1) meets the criteria outlined in the definition of refugee in the Geneva Convention and (2) sounds authentic and convincing in its claims of persecution. Three coders read 350 unique narratives, 59 of which were triple coded. These coders were selected from the pool of research assistants who helped with the digitization process and who had a master's degrees in law or political theory. This selection process ensured the confidentiality of the narratives, a first-order parameter in this study, and that coders were knowledgeable about the Geneva Convention. By the time they read and coded these narratives, they had each transcribed several hundred, an important requirement for reliable hand coding (Krippendorff 2004). Importantly, they were not informed of the decision reached by the bureaucrat.

For each narrative, they first coded whether it mentioned any form of persecution, and if so, what type (race, political opinion, religion, nationality, or social group). Second, they assessed whether the claims were believable, convincing, detailed, individualized, and coherent, using a 4-point scale for each. Third, they coded whether the narrative referred to a historical event and whether the applicant mentioned members of her family who were already living in France. Finally, research assistants were asked to determine whether the applicant made a reasonable claim to political asylum, as defined by the Geneva Convention, by answering the question, “In your opinion, is this person entitled to claim the right to asylum according to the Geneva Convention?” (*Disagree/Somewhat disagree/Somewhat agree/Agree*). On average, coders “agreed” with the statement 15% of the time, which is comparable to the average acceptance rate in the sample and “somewhat agreed” 35% of the time (Table C.6, p. 13, in the online supporting information). I coded narratives as credible if research assistants answered either “Agree” or “Somewhat agree” because the intercoder reliability, as measured by the Krippendorff's α coefficient,⁶ is much higher (0.48) when combining these two responses than when considering narratives to

⁶Krippendorff α is computed as $1 - (\text{number of disagreements actually observed}) / (\text{number of disagreements observed by chance})$. This ratio is used to determine intercoder reliability, where 0 indicates perfect disagreement and 1 perfect agreement. For more in-

be credible only when they were coded as “Agree” (0.19) (Table C.7, p. 13).

Using the set of hand-coded narratives, I then compare the performance of three classification algorithms to predict the credibility of the narrative on a left-out sample. Using Random Forest, I am able to accurately predict the credibility measure 75% of the time in the left-out sample of 85 narratives, a substantial reduction in error compared to a baseline of 48% (Table C.9, p. 1, in the online supporting information), and to ensure that the predicted probabilities are well calibrated (Figure C.1, p. 16). This predicted measure correlates as we would expect with the features of the narrative (Table C.10, p. 17) and, as detailed in Section 6, with the decision as well, which suggests that this measure is capturing the credibility of the narrative, even if imperfectly. Far from claiming that research assistants are better than civil servants at their jobs, or even that we should replace bureaucrats with computers, this procedure is an attempt to summarize, in a single variable, important variation among applicants’ narratives that could explain officials’ decisions whether to grant or withhold refugee status.

Although this procedure inevitably introduces some noise (λ), the first key identification assumption is that this noise is uncorrelated with the characteristics of asylum seekers I examine (religion, education, skill level, and proficiency in French). This would be problematic, for example, if research assistants were more likely to find the same narrative more credible if the asylum seeker is Christian than if she is Muslim. To help mitigate this concern, I check the robustness of the results by controlling for the effect of the narratives in two additional ways. First, I directly control for features of the narratives that *should* be relevant in explaining the decision (listed in Table 2). Second, I use the supervised Indian Buffet Process developed in Fong and Grimmer (2016) to discover the features of the text that *are* relevant in explaining the decision. The model takes a document-term matrix as its input and learns from a training set (30% of the narratives chosen at random [$N = 1180$]), a set of latent binary features that are predictive of both the text and the outcome. I set the number of features to eight and search using a range of parameters to select the model that ranked the highest on a quantitative measure of model fit (Fong and Grimmer 2016). I use this model to infer the latent treatments for the test set (the remaining 70% of narratives), which I use as additional controls in the analysis. By using a split-sample design, this procedure solves the

identification and estimation problems that arise from using the same documents to discover treatments and estimate causal effects (Egami et al. 2018).

The second identifying assumption is that the bureaucrat’s specific noise that proxies for his ability (η) — which I do not observe — is uncorrelated with the characteristics of asylum seekers I examine. This would be a concern if bureaucrats were better able to assess the credibility of the same narrative when the asylum seeker belongs to one group over another — if, for example, the bureaucrat was more knowledgeable about one group than another such that the same narrative would appear more credible if it were written by an applicant from a particular group. The fact that bureaucrats specialize by nationality rather than any of the characteristics that I examine helps mitigate this concern for education, skill level, and proficiency in French. This, however, is not implausible for religion, since this is also one of the motives defined in the Geneva Convention. To address this concern, I show that the results on religion are robust to restricting the analysis to the sample of asylum seekers who did not claim persecution on the basis of religion.

Results

Discrimination in the Attribution of Refugee Status

What are the determinants of asylum decisions in France? In Table 3, I analyze how the applicants’ religion, educational attainment, skill level, and proficiency in French affect asylum decisions. In addition to country-of-origin and year-of-application fixed effects, all specifications include a limited set of controls: age, gender, marital status, time spent in France prior to applying for refugee status, and the credibility of the narrative. This analysis reveals that religion is an important predictor of asylum decisions (column 1). Muslim applicants are 6.2 percentage points less likely to obtain refugee status than Christians who are similar across all other characteristics. This represents a substantial difference (41%) given that the average acceptance rate was 15% during the study period. The effect of educational attainment and skill level on the decision is large as well. Compared to those who reported starting university, those who reported secondary- or primary-level education were 2.4 and 6.3 percentage points less likely to be granted refugee status, respectively, though only the difference between primary and postsecondary education is statistically significant. Those with middle- or low-skill levels were both significantly much less likely (6.1 and 8.2 percentage points, respectively) to obtain refugee status. Compared

formation, see Krippendorff (2004, pp. 221–27). Agreements ranging from 0.4 to 0.6 are considered a sign of moderate agreement between coders (Landis and Koch 1977).

TABLE 3 Discrimination in the Attribution of Refugee Status

	Full sample	Substantive text features	Discovered text features	Not interviewed	Bureaucrats fixed effects	Minority status	Persecuted for religion
Credible Narrative	0.058** (0.013)	0.025 [†] (0.015)	0.062** (0.016)	0.049* (0.022)	0.065** (0.015)	0.056** (0.013)	0.031* (0.014)
Religion							
Muslim	−0.062** (0.020)	−0.056** (0.019)	−0.051* (0.023)	−0.081* (0.032)	−0.060* (0.024)	−0.084** (0.023)	−0.045 [†] (0.024)
Other	0.061* (0.027)	0.044 (0.027)	0.079* (0.034)	−0.016 (0.038)	0.034 (0.032)	0.092** (0.028)	0.054 [†] (0.029)
None/Missing	−0.017 (0.022)	−0.007 (0.022)	−0.024 (0.027)	−0.047 (0.030)	−0.009 (0.025)	0.002 (0.021)	0.021 (0.025)
Education							
Secondary	−0.024 (0.020)	−0.016 (0.020)	−0.010 (0.022)	−0.053 (0.034)	−0.024 (0.023)	−0.022 (0.020)	−0.019 (0.021)
Primary	−0.063* (0.028)	−0.064* (0.028)	−0.062* (0.031)	−0.067 (0.047)	−0.068* (0.033)	−0.059* (0.028)	−0.067* (0.030)
None/Missing	−0.033 (0.022)	−0.024 (0.023)	−0.023 (0.025)	−0.035 (0.039)	−0.040 (0.027)	−0.031 (0.022)	−0.045 [†] (0.024)
Skill level							
Middle	−0.061** (0.021)	−0.048* (0.021)	−0.059* (0.026)	−0.046 (0.031)	−0.051* (0.023)	−0.063** (0.021)	−0.057* (0.024)
Low	−0.082** (0.022)	−0.070** (0.022)	−0.077** (0.027)	−0.044 (0.032)	−0.063* (0.025)	−0.083** (0.022)	−0.079** (0.024)
None/Missing	−0.025 (0.022)	−0.020 (0.022)	−0.021 (0.025)	0.005 (0.032)	−0.012 (0.024)	−0.032 (0.022)	−0.025 (0.025)
Speaks French							
No/Missing	−0.033 [†] (0.018)	−0.027 (0.018)	−0.038* (0.019)	−0.027 (0.024)	−0.022 (0.021)	−0.036* (0.018)	−0.022 (0.019)
Observations	3,930	3,754	2,772	1,481	3,930	3,898	3,103
R ²	0.360	0.386	0.337	0.642	0.510	0.367	0.394
Mean Acceptance Rate	0.153	0.153	0.153	0.175	0.153	0.153	0.153

Notes: Point estimates and standard errors in parentheses. The dependent variable is a dummy variable indicating whether the applicant received refugee status upon first examination at the French Asylum Office. All regressions include demographic characteristics (age, gender, and marital status), an indicator variable for whether they spent more than 1 year in France before applying, country of origin, and year of application fixed effects. Reference category for “Religion” is Christian. Reference category for “Education” is university. Reference category for “Skill Level” is high. Reference category for “Speaks French” is yes.

[†] $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

to those who speak French, not speaking French has a small negative effect (3.3 percentage points) on the decision. However, when restricting the sample to applicants who were not interviewed, even though the estimates of the coefficients are similar in sign and magnitude, only the difference between Christians and Muslims remain statistically significant (Table 3, column 4).

To alleviate concerns that the difference between Christians and Muslims is driven by omitted variables, I perform four sets of robustness checks. First, I show in Table D.1 in the online supporting information (p. 21) that this result is robust to including additional covariates reported in the application form (column 1), including the number of children, whether the application was expedited, whether the applicant declared a passport or a diplomatic laissez-passer, whether the applicant has a family member who obtained refugee status in France, whether the applicant completed military service and whether her entry to the territory was legal (see Table B.1, p. 9) for summary statistics on these additional variables). I also show that this result is robust to including country of origin-year of application interactions (column 2), to pruning the sample using coarsened exact matching (Iacus, King, and Porro 2012) (column 3), and to reestimating effect sizes using a logistic regression model (Table D.2, p. 22). Finally, in column 4, I show that the coefficient is very similar when the credibility measure is omitted, which suggests that the results do not hinge entirely on this particular measure.

Second, I show that the result is robust to alternative ways of controlling for the effect of the personal narrative. In columns 2 and 3 of Table 3, in addition to the credibility of the narrative, I control, respectively, for substantive text features (described in Table 2) and for text features discovered by the supervised Indian Buffet Process in the left-out one-third of the sample. The coefficient is less precise in this last specification, which could be a function of the smaller sample size. Overall, the result is not sensitive to these alternative ways of controlling for the narrative's effect on the decision.

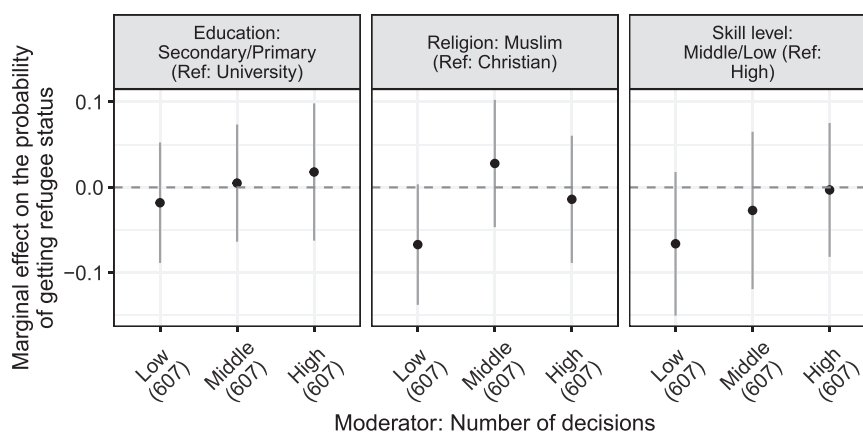
Third, by showing that the result holds after controlling for bureaucrat fixed effects, I reject the possibility that these differences result from a process in which asylum applications submitted by Muslims are systematically assigned to stricter bureaucrats. This analysis is complicated by the fact that information about bureaucrats is incomplete for applications filed before 2000, both in the administrative database and in the paper applications I digitized. Overall, the identifier of the bureaucrat who decided the case is missing for 17% of the total sample and for 53% of applications filed in 1999 or earlier. To deal with these missing values, I binned all ap-

plications for which the identifier of the bureaucrats was missing into an additional "missing" category (Table 3, column 5).

Finally, I perform a set of robustness checks to ensure that the difference between Christians and Muslims is not entirely driven by the fact that asylum officers have access to background information regarding the persecution experienced by these different groups. To rule out the possibility that Muslims are less likely to be granted refugee status because they are, in fact, less persecuted than Christians, I first show that the results hold when I control for whether the applicant belongs to a minority religious group in her home country (Table 3, column 6). The World Religion dataset of the Association of Religion Data Archives estimates the percentage of the population that has identified with Christianity or Islam since 1945 for most countries in the world. Using these estimates, I generate a binary variable that indicates, for each applicant in the sample, whether she belongs to a religious group that comprises less than 20% of the population in her home country.⁷

I then show that differences between Christians and Muslims also hold in the subset of applicants who did not claim persecution on religious grounds. The rationale for this test is that if Christian applicants are more likely to receive asylum than Muslims because they are more likely to be persecuted than Muslims, then the Muslim gap should not hold in the subsample of applicants who did not claim persecution on religious grounds. Only 6% of narratives claimed persecution on religious grounds according to the coding completed by the research assistants on a representative sample of applicants (Table C.6, p. 13 in the online supporting information), and the proportion of the text dedicated to the topic *Religion* calculated by the structural topic model exceeds 20% in only 2% of the narratives (Table B.1, p. 9). But identifying applicants who claim religious persecution is not straightforward. About 1,021 of the narratives (22%) mention at least one religious keyword from a relatively short list. To identify narratives that claim persecution on religious grounds (one of the five motives listed in the Geneva Convention), I read 292 narratives, a random sample of the 1,021 narratives that contained at least one religious keyword, and for each I coded whether the applicant was claiming persecution on religious grounds. Restricting the sample to narratives that either did not contain a religious keyword or contained one but was not coded as religious persecution, I find that the Muslim penalty

⁷A total of 165 observations did not merge due to unequal overlap in coverage by country. For example, Armenia, Bosnia and Croatia appear in the World Religion data only after 1995.

FIGURE 2 Discrimination by Bureaucrats' Experience

Notes: This figure shows the estimated conditional marginal effect, along with the 95% confidence intervals based on standard errors clustered by bureaucrats, and the number of observations in each bin in parentheses. The specifications include covariates and fixed effects for year of application and country of origin.

holds (Table 3, column 7). Overall, these additional tests suggest that the difference between Christians and Muslims is unlikely to be explained entirely by the fact that Christians are more persecuted than Muslims.

In short, these results show that applicants who are Christian are much more likely to be granted asylum than Muslim applicants. In a series of robustness tests, I rule out the possibility that this difference is entirely explained by observable or unobservable omitted variables by the selective assignment of cases to bureaucrats or by the fact that Muslims are less persecuted than Christians.

Bureaucrats' Experience Mitigates Discrimination

I next examine the influence of bureaucrats' experience (measured as the number of past decisions) on discrimination on the basis of religion, education, and skill level. I restrict the sample to the 1,821 asylum applications filed between 2000 and 2014 that were decided by a bureaucrat who started working between 2001 and 2014 for whom I know the number of past decisions. In this sample, the average number of past decisions is 457, the median 312, and the 90th percentile is 1,103 decisions. To estimate the conditional marginal effects, I use the binning estimator proposed by Hainmueller, Mumolo, and Xu (2019), which has two main advantages over the classical linear multiplicative interaction model. First, the binning estimator allows conditional marginal effects to vary across bins of the moderator by relaxing

the linear interaction-effect assumption. Second, the binning estimator ensures common support to reliably estimate the conditional marginal effects by constructing bins based on the support rather than on values of the moderator.

I thus first divide the number of past decisions into three equally sized groups — a low (1–184), middle (185–486) and high (487–3,220) number of past decisions — and pick the median number of past decisions as the evaluation point within each bin. I then estimate a model that includes, in addition to the covariates from the main specification (Table 3, column 1), interactions between the indicator variables for each of the three bins, the individual characteristics of interest, and the number of past decisions minus the evaluation point picked within each bin, as well as triple interactions of these. I conduct this analysis separately for religion (comparing Muslims to Christians), education (comparing secondary and primary to postsecondary) and skill level (comparing middle and low to high). Figure 2 plots the conditional marginal effects of these three characteristics at low, middle, and high values of the moderator, with 95% confidence intervals. These analyses reveal a consistent pattern for all three characteristics: the marginal effect is the most negative in the lowest bin, and it is reduced in subsequent bins, suggesting that discrimination is most pronounced at lower levels of experience. The pattern is similar when constructing bins based on the values of the moderator (Figure D.3, p. 24, in the online supporting information) and when estimating the conditional marginal effects using generalized additive models (Figure D.4, p. 25).

TABLE 4 Discrimination by Bureaucrats' Experience (Religion)

	Granted Refugee Status			
	(1)	(2)	(3)	(4)
More than 184		−0.011 (0.023)	0.011 (0.032)	−0.102 [†] (0.055)
Religion				
Muslim	−0.027 (0.040)	−0.078 (0.042)	−0.032 (0.052)	−0.067 (0.073)
Other	0.111* (0.043)	0.024 (0.038)	0.023 (0.062)	0.043 (0.116)
None/Missing	0.016 (0.045)	0.034 (0.058)	0.128 (0.095)	0.108 (0.128)
Interactions				
More than 184 × Muslim		0.097** (0.035)	0.126** (0.048)	0.172* (0.069)
More than 184 × Other		0.158* (0.072)	0.094 (0.084)	0.109 (0.129)
More than 184 × None/Missing		−0.032 (0.050)	0.192 (0.085)	−0.053 (0.117)
Bureaucrats fixed effects	No	No	No	Yes
R ²	0.251	0.273	0.310	0.501
Observations	1,190	1,190	659	659

Notes: This table shows point estimates and standard errors clustered by bureaucrats in parentheses from ordinary least squares (OLS) regressions with individual covariates and fixed effects for year of application and country of origin. The sample is restricted to the first 486 decisions of bureaucrats in columns 1 and 2 and to the first 486 decisions of bureaucrats who made at least 486 decisions in columns 3 and 4. Reference category for "Religion" is Christian.

[†] $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

To test whether this reduction in discrimination is statistically significant, I further restrict the sample to the first 486 applications (second tercile of the moderator) of each bureaucrat in the sample and report in Table 4 estimates from the main specification (column 1) and from the binning estimator described above (columns 2 to 4). While Muslims in this subsample are overall 2.7 percentage points less likely to be granted asylum (*SE*: 4.0 pp.) compared to Christians (column 1), I find a substantial and statistical difference depending on whether the application was decided by an inexperienced or experienced bureaucrat. Muslims are 7.8 percentage points (*SE*: 4.2 pp.) less likely to receive refugee status than Christians when their case is decided by an inexperienced bureaucrat. But while switching from an inex-

perienced to an experienced bureaucrat does not improve Christians' chances of receiving asylum (coefficient: −0.011, *SE*: 0.023), it increases Muslims' chances by 9.7 percentage points (*SE*: 3.5 pp.), such that the difference between Christians and Muslims is small and insignificant among experienced bureaucrats (column 2). To ensure that this pattern is not driven by the changing composition of bureaucrats over time, I further restrict the sample to applications examined by bureaucrats who decided at least 486 applications (columns 3 and 4 in Table 4). The results are robust to restricting the sample in this way and to including bureaucrat fixed effects (column 4).

The pattern is similar for educational attainment and skill level (Table 5). Compared to highly educated

TABLE 5 Discrimination by Bureaucrats' Experience (Education and Skill Level)

	Education				Skill Level			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
More than 184		−0.007 (0.044)	0.017 (0.071)	−0.150 (0.100)		0.005 (0.056)	0.046 (0.090)	−0.130 (0.127)
Characteristics								
Medium	0.004 (0.026)	−0.017 (0.036)	−0.025 (0.054)	−0.081 (0.072)	−0.044 (0.035)	−0.067 (0.047)	−0.086 (0.075)	−0.074 (0.112)
Low	0.009 (0.038)	−0.030 (0.045)	−0.073 (0.075)	−0.173 (0.117)	−0.073* (0.035)	−0.098* (0.042)	−0.085 (0.073)	−0.122 (0.110)
None/Missing	0.011 (0.038)	−0.070 (0.041)	−0.114 (0.058)	−0.198* (0.095)	−0.047 (0.033)	−0.063 (0.041)	−0.071 (0.070)	−0.078 (0.104)
Interactions								
More than 184 × Medium		0.034 (0.049)	0.032 (0.073)	0.097 (0.094)		0.048 (0.066)	0.030 (0.094)	0.074 (0.136)
More than 184 × Low		0.057 (0.063)	0.095 (0.100)	0.226 (0.142)		0.054 (0.065)	0.032 (0.101)	0.143 (0.135)
More than 184 × None/Missing		0.156* (0.061)	0.131 (0.088)	0.254 (0.132)		0.034 (0.064)	0.004 (0.094)	0.104 (0.126)
Bureaucrats fixed effects	No	No	No	Yes	No	No	No	Yes
R ²	0.251	0.270	0.295	0.498	0.251	0.263	0.297	0.500
Observations	1,190	1,190	659	659	1,190	1,190	659	659

Notes: This table shows point estimates and standard errors clustered by bureaucrats in parentheses from 8 OLS regressions with individual covariates and fixed effects for year of application and country of origin. The dependent variable is whether the applicant was granted refugee status. The sample is restricted to bureaucrats' first 486 decisions in columns 1, 2, 5 and 6 and to their first 486 decisions if they made more than that in columns 3, 4, 7 and 8. "Characteristics" refers to "Education" in columns (1) to (4) and to "Skill Level" in columns (5) to (8). Reference category for "Characteristics" is high.

[†] $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

and highly skilled asylum seekers, those with lower levels of education and skills are less likely to be granted refugee status when their case is examined by an inexperienced bureaucrat (the relevant coefficients are all negative although not all reach statistical significance at conventional levels), but, consistent with a reduction in discrimination among more experienced bureaucrats, the coefficients of the interaction terms are all positive (although only one reaches statistical significance). In Table D.3 in the online supporting information (pp. 23), I show that applications decided by experienced and inexperienced bureaucrats do not differ systematically, which allows me to mitigate concerns that this result reflects the fact that bureaucrats are assigned different types of decisions during the course of their tenure.

Overall, these patterns suggest that bureaucrats discriminate less as they get more experience, but this reduction is stronger and more robust for religion than for education and skill level. To investigate the time frame during which this reduction is taking place, I also report in Table D.3 in the online supporting information (pp. 23), for each group of applications, the average number of months the bureaucrat who decided the case had spent on the job at the time of the decision, simply counting the number of months between each decision and his first decision. The reduction in discrimination seems to be taking place over the course of roughly a year: in the first bin, applications are being decided by bureaucrats who had spent roughly 9 months on the job, while in the second bin they had spent at least 2 years.

Discussion

What explains this reduction in discrimination as bureaucrats get more on-the-job experience? If discrimination results primarily from the fact that bureaucrats have a preference for different types of applicants, then the reduction could be explained by a change in bureaucrats' preferences on the job, for example, if c_B is positive for inexperienced bureaucrats but close to zero for experienced bureaucrats. Contact theory hypothesizes that personal contact across social lines can reduce prejudice (Allport 1974). According to this logic, as bureaucrats interview more applicants, they could become less prejudiced against Muslims, leading to a reduction in taste-based discrimination. It is important to note, however, that the interview (which is the only point of contact between the bureaucrat and the asylum seeker) does not fit the traditional assumptions of contact theory, which maintain that contact should be cooperative, have a shared goal, and be endorsed by communal authority — and that participants should have equal power status. Contact during an interview quite strongly violates all of these assumptions, with the possible exception of communal authority. Indeed, recent research suggests that while collaborative contact reduces own-caste favoritism in the context of cricket leagues, adversarial contact does not lead to a similar reduction in favoritism (Lowe 2021).

If discrimination is instead driven by bureaucrats' prior beliefs about the level of persecution faced by different groups, a change in these beliefs (or in the weight they place on them) over time could also explain the reduction in discrimination observed. Several first-hand accounts of decision-makers at the French Asylum Office and the appeals court suggest that the narratives are relatively uninformative for bureaucrats when they start working at the Asylum Office, but that officials hone their ability to identify credible narratives over time. Reflecting on her time as an asylum officer, Aho Nienne reports: "My colleagues assured me that, with experience, I would acquire a gift that is crucial to our profession: intimate conviction. That indescribable feeling when an asylum seeker lies" (2013). Bureaucrats working at the appeals court confirm that they learn how to infer a sharp signal over time. "Being experienced means to be able to recognize when an applicant is lying or when he is a 'fake refugee,'" one of them testified. As a result, she continued, "At first you do not know what to look for, what to base [decisions] on" (Greslier 2007, p. 8). The intuition is simple: it takes time to be able to single out truly authentic narratives that distinguish themselves through original storytelling and language. As they read more

and more applications, bureaucrats become better able to distinguish fake narratives from true ones. A bureaucrat working at the French Asylum Office powerfully illustrated this intuition in anonymous testimony: "When you see for the 80th time the same story written with the same font, the same line spacing, in which just a few details change... Sometimes, there is even the name of another asylum seeker in the story at one place" (Aubry and Le Löt 2019). Bureaucrats' experience thus increases the precision, that is, decreases the variance σ_η^2 of the signal they receive when reading narratives. According to Equation (1), as σ_η^2 decreases, bureaucrats place more weight on the signal ($\frac{\sigma_\omega^2}{\sigma_\omega^2 + \sigma_\epsilon^2 + \sigma_\eta^2}$), than on their prior beliefs ($\frac{\sigma_\epsilon^2 + \sigma_\eta^2}{\sigma_\omega^2 + \sigma_\epsilon^2 + \sigma_\eta^2}$).

Conclusion

Are European countries abiding by the Geneva Convention and granting asylum to those who fear persecution for reasons of "race, religion, nationality, membership of a particular group or political opinion"? Until now, the lack of microlevel data on asylum decisions has limited researchers' ability to answer this question. Yet 10 years ago the French Asylum Office opened its archives. I exploited this unprecedented degree of transparency to digitize a representative sample of more than 4,000 asylum applications filed over the last 40 years to compile the first in-depth, microlevel dataset of asylum decisions. This dataset facilitates an in-depth examination of whether France is indeed granting asylum to those in need of protection. I compared accepted and rejected applicants to identify the effect of applicants' personal characteristics on the probability of being granted asylum in France.

This study provides empirical evidence that Christian applicants are much more likely to be granted refugee status than Muslim applicants. Moreover, I show that these effects are unlikely to be driven by either unobserved differences revealed during the interview process or by the selective assignment of cases to bureaucrats. Importantly, the findings also reveal that bureaucrats can self-correct their established discriminatory behaviors. As they gain more on-the-job experience, they discriminate less on the basis of religion, and to a lesser extent based on education and skill level. Anecdotal evidence suggests that bureaucrats learn over time to distinguish authentic narratives, which is consistent with statistical discrimination. While the narratives provided by applicants are relatively uninformative at first, bureaucrats gradually learn how to assess their credibility, allowing them to place more weight on the narratives than on

their prior beliefs. However, the available data do not allow me to rule out the possibility that the reduction in discrimination instead comes from a change in bureaucrats' preferences.

The results of the study have important policy implications for strategies to reduce discrimination at the French Asylum Office. For instance, exposure to past narratives during bureaucrats' training could help them learn to better assess the credibility of persecution claims once they are faced with actual applicants. Moreover, increasing bureaucrats' tenure could reduce the overall level of discrimination. But more research is needed to test the effectiveness of these interventions. This research also opens up a new avenue of study on the effectiveness of possible interventions to reduce the influence of bureaucrats' beliefs on decisions in order to curtail discrimination within administrations.

These findings unambiguously call for increased scrutiny in the attribution of refugee status in France, and within the EU more broadly. Since the Geneva Convention constitutes the common framework for granting asylum across Europe, asylum officers elsewhere on the continent are afforded similar levels of discretion as those in the French office. However, we currently lack the data needed to compare working and training conditions across European asylum offices. Similar data collection in other countries, either by researchers or even the offices directly, should be undertaken to determine whether the discrimination patterns identified in France apply more broadly.

References

- Adida, C. L., D. D. Laitin, and M.-A. Valfort. 2016. *Why Muslim Integration Fails in Christian-Heritage Societies*. Harvard University Press.
- Adida, Claire L., Adeline Lo, and Melina R. Platas. 2019. "Americans preferred Syrian refugees who are female, English-speaking, and Christian on the Eve of Donald Trump's election." *Plos One* 14(10): 1–18.
- Aho Nienne, C. 2013. *Pourquoi il ne faut pas demander l'asile politique en France*. <https://blogs.mediapart.fr/celine-ahonienne/blog/291013/pourquoi-il-ne-faut-pas-demander-l-asile-politique-en-france>, Allport, G. (1974): The Nature of Prejudice.
- Ariel R., White, Noah L. Nathan, and Julie K. Faller. 2015. "What do I need to vote? Bureaucratic discretion and discrimination by local election officials." *American Political Science Review* 109(1): 129–42.
- Arnold, D., W. Dobbie, and C. S. Yang. 2018. "Racial Bias in bail decisions." *The Quarterly Journal of Economics* 133(4): 1885–932.
- Arrow, Kenneth J. 1973. "The Theory of Discrimination." In: *Discrimination in Labor Markets*: 3–33.
- Aubry, A., and K. Le Loët. 2019. "Il y a des gens qui craquent. Et ils ne font pas semblant de craquer." *Les Jours* Episode 6. <https://lesjours.fr/obsessions/immigration-mineurs/ep6-ofpra-decision/>
- Bansak, Kirk, Jens Hainmueller, and Dominik Hangartner. 2016. "How economic, humanitarian, and religious concerns shape European attitudes toward asylum seekers." *Science* 354(6309): 217–22.
- Becker, G. S. 1971. *The Economics of Discrimination*.
- Bohren, J. A., K. Haggag, A. Imas, and D. G. Pope. 2019. "Inaccurate statistical discrimination: An identification problem (No. w25935)." *National Bureau of Economic Research*.
- Butler, Daniel M., and David E. Broockman. 2011. "Do politicians racially discriminate against constituents? A field experiment on state legislators." *American Journal of Political Science* 55(3): 463–77.
- Cohen, D. G. 2000. "Naissance d'une nation: Les personnes d'eplacées de l'après-guerre, 1945–1951." *Genèses* 38: 56–78.
- Egami, N., C. J. Fong, J. Grimmer, M. E. Roberts, and B. M. Stewart. 2018. How to Make Causal Inferences Using Texts. arXiv preprint arXiv:1802.02163.
- European Commission. 2016. Proposal for a Regulation of the European Parliament and of the Council Establishing a Common Procedure for International Protection in the Union and Repealing Directive 2013/32/EU.
- Fariss, C. J., and A. J. Rottman. 2009. The Political Economy of Decision Making Bias: An Analysis of Asylum Merit Hearings in U.S. Immigration Court. Available at SSRN. https://www.researchgate.net/publication/255725660_The_Political_Economy_of_Decision_Making_Bias_An_Analysis_of_Asylum_Merit_Hearings_in_US_Immigration_Court
- Fong, C., and J. Grimmer. 2016. "Discovery of treatments from text corpora." *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*: 1600–9.
- Greslier, F. 2007. "La commission des Recours des Réfugiés ou "l'intime conviction" face au recul du droit d'asile en France." *Revue européenne de migrations internationales* 23(2): 107–33.
- Hainmueller, Jens. 2012. "Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies." *Political Analysis* 20(1): 25–46.
- Hainmueller, Jens, and Michael J. Hiscox. 2007. "Educated preferences: Explaining attitudes toward immigration in Europe." *International Organization* 61(2): 399–442.
- Hainmueller, Jens, and Michael J. Hiscox. 2010. "Attitudes toward highly skilled and low-skilled immigration: Evidence from a survey experiment." *American Political Science Review* 104(1): 61–84.
- Hainmueller, Jens, and Daniel J. Hopkins. 2015. "The Hidden American Immigration Consensus: A Conjoint Analysis of Attitudes toward Immigrants." *American Journal of Political Science* 59(3): 529–48.
- Hainmueller, J., J. Mummolo, and Y. Xu. 2019. "How much should we trust estimates from multiplicative interaction

- models? Simple Tools to Improve Empirical Practice” *Political Analysis* 27(2): 163–92.
- Helbling, M., and R. Traunmüller. 2020. “What is Islamophobia? Disentangling citizens’ feelings toward ethnicity, religion and religiosity using a survey experiment.” *British Journal of Political Science* 50(3): 811–28.
- Hemker, J., and A. Rink. 2017. “Multiple dimensions of bureaucratic discrimination: Evidence from German welfare offices.” *American Journal of Political Science* 61(4): 786–803.
- Holzer, T., G. Schneider, and T. Widmer. 2000. “Discriminating decentralization federalism and the handling of asylum application in Switzerland, 1988–1996.” *Journal of Conflict Resolution* 44(2): 250–76.
- Iacus, Stefano M., Gary King, and Giuseppe Porro. 2012. “Causal inference without balance checking: coarsened exact matching.” *Political Analysis* 20(1): 1–24.
- Keith, L. C., and J. S. Holmes. 2009. “A rare examination of typically unobservable factors in US asylum decisions.” *Journal of Refugee Studies* 22(2): 224–41.
- Keith, Linda Camp, Jennifer S. Holmes, and Banks P. Miller. 2013. “Explaining the divergence in asylum grant rates among immigration judges: an attitudinal and cognitive approach.” *Law and Policy* 35(4): 261–89.
- Krippendorff, K. 2004. *Content Analysis: An Introduction to its Methodology*.
- Landis J. Richard, and Gary G. Koch. 1977. “The measurement of observer agreement for categorical data.” *Biometrics* 33(1): 159–74.
- Lipsky, M. 2010. *Street-Level Bureaucracy*.
- Lowe, M. 2021. “Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration” *American Economic Review* 111(6): 1807–44.
- McClendon, Gwyneth H. 2016. “Race and responsiveness: an experiment with South African politicians.” *Journal of Experimental Political Science* 3(01): 60–74.
- Neggers, Y. 2018. “Enfranchising your own? Experimental evidence on Bureaucrat diversity and election bias in India.” *American Economic Review* 108(6): 1288–321.
- Neumayer, Eric. 2005. “Asylum recognition rates in Western Europe: Their determinants, variation, and lack of convergence.” *Journal of Conflict Resolution* 49(1): 43–66.
- Olsen, A. L., J. H. Kyhse-Andersen, and D. Moynihan. 2020. “The unequal distribution of opportunity: a national audit study of bureaucratic discrimination in primary school access.” *American Journal of Political Science* 66(3): 587–603.
- Phelps, E. 1972. “The Statistical Theory of Racism and Sexism.” *The American Economic Review* 62(4): 659–61.
- Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, and David G. Rand. 2014. “Structural topic models for open-ended survey responses.” *American Journal of Political Science* 58(4): 1064–82.
- Rosenblum, Marc R., and Idean Salehyan. 2004. “Norms and interests in US asylum enforcement.” *Journal of Peace Research* 41(6): 677–97.
- Rottman, Andy J., Christopher J. Fariss, and Steven C. Poe. 2009. “The path to asylum in the US and the determinants for who gets in and why.” *International Migration Review* 43(1): 3–34.
- Salehyan, Idean, and Marc R. Rosenblum. 2008. “International relations, domestic politics, and asylum admissions in the United States.” *Political Research Quarterly* 61(1): 104–21.
- Schneider, G., and T. Holzer. 2002. *Asylpolitik auf Abwegen: nationalstaatliche und europäische Reaktionen auf die Globalisierung der Flüchtlingsströme*.
- Toshkov, Dimitar Doychinov. 2014. “The dynamic relationship between asylum applications and recognition rates in Europe (1987–2010).” *European Union Politics* 15(2): 192–214.
- Valentino, Nicholas A., Stuart N. Soroka, Shanto Iyengar, Toril Aalberg, Raymond Duch, Marta Fraile, Kyu S. Hahn, Kasper M. Hansen, Allison Harell, Marc Helbling, Simon D. Jackman, and Tetsuro Kobayashi. 2017. “Economic and cultural drivers of immigrant support worldwide.” *British Journal of Political Science* 49(4): 1201–26.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix A: Model

Appendix B: Sampling

Appendix C: Measuring the credibility of the narrative

Appendix D: Additional tables and figures