

The Burden of Proof in Immigration Bond Decisions: An Impact Study of *Brito v. Barr*

Anthony J. DeMattee,* Hallie Ludsin,† Grace Shrestha,‡
Grace Gerenday,§ Devon Thurman,¶ Jeffrey K. Staton||

Abstract

Detained individuals subject to deportation have the right to a bond hearing in immigration court similar to that of detained individuals accused of a crime. Unlike criminal law, immigration law places the burden of proof on detained people rather than the government. We analyze the impact of a federal court decision that shifted the burden of proof to the government via a synthetic control study and a qualitative research design grounded in a new theoretical analysis of immigration courts that focuses on judicial decision-making and prosecutorial discretion. The evidence suggests significant limits on the federal courts' ability to change bond outcomes merely through changing the burden of proof.

Keywords: separation of powers, judicial decision-making, prosecutorial discretion, observational study, treatment effects, synthetic control, immigration courts

Data availability statement. Replication materials are available at the *Journal of Law and Courts*' Dataverse archive.

Acknowledgments: We want to thank those who participated in this study. Their insights and expertise in navigating the immigration system greatly enriched our research. We appreciate the thoughtful feedback from our three peer reviewers and the editorial team at the *Journal of Law and Courts*, whose comments helped refine our analysis and improve the paper. The views and any errors remain our own.

Forthcoming at the *Journal of Law and Courts*

*Data Scientist, Democracy Program, The Carter Center

†Visiting Assistant Professor of Practice, Emory University School of Law

‡Department of Political Science, Emory University

§Department of Political Science, Emory University

¶Ph.D. Candidate, Department of Political Science, Emory University

||Professor, Department of Political Science, Emory University

1 Introduction

Between October 2018 and December 2019, Immigration and Customs Enforcement (“ICE”) detained an average of 49,000 individuals per day who were subject to removal from the country for alleged violations of U.S. immigration law. Immigration detention is akin to pretrial detention in the criminal law context, and like the criminally accused, individuals detained in immigration proceedings can seek bonded release pending the resolution of their removal case. However, unlike the criminally accused, the immigration detained carry the burden of proving that they are neither a flight risk nor dangerous. They must prove their case to the “satisfaction” of an Immigration Court Judge (IJ), a standard that allows considerable discretion on the part of judges directly accountable to the Attorney General of the United States. In November 2019, the United States District Court for the District of Massachusetts found that this framework violated the due process clause of the 5th Amendment. In *Brito v. Barr*, the court ordered that the burden of proof be assigned to the government and established that the government must prove that a respondent is a flight risk by the preponderance of evidence, and dangerous by clear and convincing evidence.¹ The decision put into practice common views of immigration advocates as well as several legal scholars who have argued that placing the burden of proof on the detained fails to provide sufficient protections for the due process rights of individuals subject to the jurisdiction of the United States (Brady, 2022; Holper, 2016). We study whether and if so how the *Brito* decision changed the outcomes of the bond process in Boston.

There are simple theoretical reasons to be both optimistic and skeptical about *Brito*’s ability to shape immigration court outcomes, and so the case invites us to be careful and clear about the theoretical mechanisms through which *Brito* should have worked. *Brito* also raises two common methodological challenges that complicate analyses of judicial impact. First, the behavioral effects of judicial decisions are not necessarily constrained to the primary targets of the decision. The *Brito* judgment targeted the judges of the Boston Immigration Court, but it could have influenced the initial bond decisions of the ICE officers in anticipation of different IJ treatment under the new standards. This process has implications for what we should expect to observe from IJs in the wake of *Brito*. Second, while the decision creates an opportunity to evaluate the effects of changing the burden of proof in immigration bond hearings, it is not immediately clear against what we should compare the Boston jurisdiction.

¹*Brito v. Barr*, 415 F. Supp. 3d 358, 271.

We tackle these challenges with an interdisciplinary research team that brings together the fields of law, political science, and public administration. These traditions jointly suggest alternative views about the likely effects of *Brito*. Assuming that *Brito* could have been at least minimally influential on Boston IJs, we develop a new theoretical model of the immigration bond process. The model suggests that if the decision had no impact on ICE, *Brito* would have increased the observed rate at which bond was granted in Boston. However, if the decision also affected ICE, then the decision would have driven the rate down while simultaneously shrinking the Boston Immigration Court’s caseload. To address the second challenge, we use our model to structure a synthetic control study to estimate the causal effect of *Brito* on bond grant rates and immigration court caseload. Our study’s design and interpretation of its implications are informed by a legal analysis complemented by in-depth interviews of immigration attorneys across the United States. This sample of immigration attorneys includes Boston attorneys with whom we directly discussed the case and its aftermath.

Although we find suggestive evidence that bond rates and caseloads declined as a consequence of *Brito*, we ultimately cannot reject the possibility that *Brito* did not affect IJ or ICE behavior at all. On the other hand, the evidence is strongly inconsistent with a sensible but ultimately naive prediction that by shifting the burden of proof to the government and making it harder to prove dangerousness, *Brito* should have increased the rate of granting bond in Boston. These findings speak most directly to the *Brito* decision itself, but they also have implications for theoretical and empirical scholarship on the burden of proof (Sherwyn and Heise, 2010; Finley and Karnes, 2008; Cheng, 2012; Wexler, 1999; Holper, 2016). Our study is also relevant for the literature on decision-making in immigration courts (Blasingame et al., 2023; Hausman, 2016; Kim and Semet, 2019; Ryo, 2016; Eagly and Shafer, 2015; Durham, 2005), in which scholars have revealed significant ideological and political patterns of IJ behavior. Finally, we believe that our findings speak to more general scholarship on the challenges courts face in controlling executive agencies (e.g. Spriggs, 1996, 1997; Ferejohn and Shipan, 1990; Staton and Vanberg, 2008).

We divide the remainder of our paper as follows. We first summarize immigration detention and the *Brito* decision. We then review general perspectives on the ability of courts to control agencies. We next offer a conceptual summary of the burden of proof and develop a theoretical account of the immigration bond process. We then test our theoretical predictions using a synthetic control study of *Brito* that

reconstructs case data from the Executive Office for Immigration Review made available through FOIA requests. Our final section considers the implications of our findings for future scholarship.

2 Immigration Detention

When noncitizens are charged with violating the country’s immigration laws, they enter removal proceedings in immigration court. The Immigration and Nationality Act (“INA”) authorizes the Department of Homeland Security (DHS) to detain noncitizens arrested for immigration violations and awaiting determinations as to whether they will be ordered removed. After noncitizens are arrested, an ICE officer determines whether they should remain in custody or be released at any time during their removal proceedings. Noncitizens may request a review of ICE’s decision by an IJ at a bond hearing. INA Section 236(a) requires the noncitizens to demonstrate that they are not a danger to the community and that they are likely to appear for future removal proceedings “to the officer’s satisfaction.” Based on this statute, the Board of Immigration Appeals (“BIA”) has held that noncitizens bear the burden of proving they should be released from custody.

Brito v. Barr was a class action lawsuit brought before the United States District Court of Massachusetts on behalf of two classes of noncitizen immigrants, each of whom is held under the discretionary detention regime pending the final outcome of their removal cases. The two classes challenged immigration regulation 8 C.F.R. §236.1(c)(8) as a violation of the 5th Amendment’s Due Process Clause, the Administrative Procedure Act (“APA”), and the INA, which requires immigrant detainees seeking release from detention to “demonstrate to the satisfaction of the officer that such release would not pose a danger to property or persons, and that the alien is likely to appear for any future proceeding.” On November 27, 2019, the District Court of Massachusetts granted summary judgment in favor of the two classes. It agreed with other federal court decisions, including from the Massachusetts District Court, that due process requires the government bear the burden of proof in bond hearings. It then asserted that the standard applied during bond hearings, “to the satisfaction of the immigration judge,” is standardless and produces arbitrary and inconsistent decisions. Recognizing the liberty interest at stake, the District Court ordered immigration courts holding custody hearings for detainees held in Massachusetts to place the burden of proof on the government to prove by clear and convincing evidence that the detained immigrants are dangerous or by the preponderance of the evidence that the noncitizen immigrant

(or “respondent” in an immigration proceeding) is a flight risk and “that no condition or combination of conditions will reasonably assure the alien’s future appearance and the safety of the community.” The District Court’s decision aligned immigration detention requirements in Massachusetts with the constitutional due process requirements for pretrial detention and other forms of civil detention. Its order applies to all immigration courts conducting bond hearings in Massachusetts effective December 13, 2019. We consider whether this decision was influential in fact.

3 Judicial Oversight of Agencies

Before we consider how *Brito* could have affected bond outcomes, it is useful to consider what scholars generally suggest about the ability of courts to control agencies. Judicial oversight of government agencies is an important institutional feature encouraging agencies to operate within the bounds of the law and adhere to the principles of justice and fairness (Dahl, 1957; Humphries and Songer, 1999; Rogers, 2001). By providing a check on agencies, judicial review fosters transparency, accountability, and the rule of law. Several factors enhance a court’s ability to effectively change the behavior of actors in other branches of government, including legal authority, ruling clarity, agency culture, and public interest. Legal authority enables courts to issue rulings that are legally-binding (Stobb, Miller and Kennedy, 2023). Specific and clear court orders, with unambiguous directives, make it easier to implement the required changes (Hitt, 2016). Likewise, agencies with respect for judicial processes and a culture of legal compliance should be more likely to change their behavior in response to court rulings. Public attitudes and media attention can embolden courts or pressure agencies to comply with court decisions (Clark, 2009), particularly in high-profile cases or issues of significant public interest.

Conversely, the absence of these factors complicate the ability of courts to control agencies. Courts may lack direct enforcement mechanisms to compel compliance from agencies. If the leaders of agencies do not support the enforcement of court decisions, street-level bureaucrats may feel empowered to ignore them (Ferejohn and Shipan, 1990). Agencies might resist court oversight if they have strong directives to enforce the law in particular ways (Johnson, 2014, 2015). These directives can encourage agencies to challenge or overlook judicial rulings, particularly in politically charged matters. Just as clear orders promote implementation, orders that lack it promote non-compliance (Spriggs, 1997; Staton, Gauri and

Cullell, 2015). Even judicial rulings that can be clearly expressed and understood in natural language can be inherently vague in application, undermining judicial control.

3.1 Judicial Oversight of Immigration Courts and Plausible Effects of *Brito*

How might these general understandings of judicial oversight apply to immigration courts? First and foremost, although the law over how much deference federal courts owe administrative agencies in the interpretation of federal statutes was in flux at the time of *Brito* (Gasaway and Parrish, 2016; Coglianesse and Walters, 2024), whether an order issued by a federal court to an administrative agency needed to be obeyed was not in question.² More importantly, there would have been no ambiguity in the courts' authority to interpret the Constitution or in their responsibility to remedy constitutional violations. There was no plausible legal argument at the time that the *Brito* ruling would have been simply advisory. The district court was legally empowered to command a change in the burden of proof in Boston. And the IJs were obligated to adopt the new framework.

IJs are also all lawyers and the legal career is characterized by a socialization process that promotes a commitment to legal norms (Knight and Epstein, 1996; Baum, 2009; Friedman, 2006; Gillman, 2001). Unsurprisingly, judicial scholars find empirical evidence linking judicial decision-making to legally relevant information (Bailey and Maltzman, 2008; George and Epstein, 1992; Bartels, 2009; Baldez, Epstein and Martin, 2006). Scholars of immigration court decision-making have also found that legally-relevant case facts affect outcomes in ways that the law anticipates (Ryo, 2016; Eagly, 2014; Eagly and Shafer, 2015; Hausman, 2016). It is thus arguable that IJs are embedded in a traditional American legal culture, one that generally respects rule of law values including the notion that federal court decisions interpreting the Constitution should be binding on all individuals, including government officials like IJs.

Further, in at least one regard, the decision was sufficiently clear and should have limited IJ discretion. Shifting the burden of proof from detained individuals to the government and requiring a heightened standard of review for evaluating the dangerousness of the individual seeking bond are not exotic judicial constructions. They have clear legal interpretations and are familiar to all individuals

²See *Baez-Sanchez v. Barr*, 947 F.3d. 1033 (2020) for an example of a Board of Immigration Appeals effort to simply disregard a decision of the Court of Appeals for the 7th Circuit.

trained in the common law. They are, in fact, used routinely in bond settings in the criminal law and in other forms of civil detention. Perhaps unsurprisingly, many of the attorneys we interviewed reported strong expectations that *Brito* would have an effect:³

“[E]verybody was in shock from it all. And defense lawyers thought that it was going to be the silver bullet.” (Research Interview #51)

“[W]e thought [Brito] was going to change our entire world.” (Research Interview #7)

“I was obviously really excited about it. I remember hearing about it the day before Thanksgiving in 2019. Obviously a huge win. [I was] optimistic because it had the reach that I had wanted for a long time, which is through a class as opposed to these individual petitions.” (Research Interview #69).

This perspective on judicial oversight of immigration courts suggests that *Brito* should have caused some kind of change. The most natural effect should have been to make it harder to detain people and more likely to be awarded bond by immigration courts.

Yet, we might be skeptical of *Brito*'s likely effects of several reasons. Immigration courts are administrative bodies operating under the Department of Justice (DOJ) and, more specifically, the Executive Office of Immigration Review (“EOIR”). IJs are neither federal judges nor administrative law judges, and they are not subject to the Code of Conduct for United States Judges. Instead, they are subject to the policies and regulations set out for them by the DOJ and EOIR. These regulations instruct IJs to “act as the [Attorney General]’s delegates in the cases that come before them.”⁴ The history of constitutional law around the power to regulate immigration conceives of immigration policy as a foreign relations issue under which Congress’s power to regulate is plenary and the President is afforded great deference in the implementation of the regulatory framework established by Congress. Although recent research finds scant empirical evidence supporting this exception (DeMattee, Lindsay and Ludsin, 2023), this legal rationale may embolden Attorneys General to more aggressively encourage IJs to use whatever discretion they can devine to arrive at the preferred political result.

³Emory University IRB Protocol ID #STUDY00003467.

⁴8 C.F.R. §1003.10(a)

Scholars have found strong effects of presidential administration on IJ decision-making (e.g., Kim and Semet, 2019). Not only do IJs not enjoy formal independence from the Attorney General, the BIA grants IJs enormous discretion in bond determinations; written decisions are not required. During the period of our study, the sitting president, Donald J. Trump, held strong anti-immigration preferences and actively attempted to influence immigration court judge behavior (Blasingame et al., 2023). IJs would have had strong incentives to resist the district court’s decision, and the standards’ discretion could have undermined a judicial effect.

Further, generations of scholarship on judges of multiple types and locations around the world have provided compelling evidence that judicial decision-making can be understood in ideological and partisan terms (Segal and Spaeth, 2002; Lauderdale and Clark, 2012; Ríos-Figueroa, 2007; Desposato, Ingram and Lannes Jr, 2015). From this perspective, legal rules are flexible enough and difficult enough to audit to allow a judge wide discretion in their application. Although the standards that *Brito* set have known meanings in the law, they are inherently vague. Whereas what it means to evaluate evidence against the clear and convincing standard is comprehensible and readily distinguished from other known standards, e.g., “beyond a reasonable doubt,” whether a judge has implemented the required standard validly is not observable; it must be inferred. The threshold degree of confidence in the evidence necessary for a particular finding is certainly latent and best understood as a mental construct developed and used in the minds of particular judges. Judges can apply the same standard in different ways simply because their thresholds for persuasion differ. Although the law implies an ordering of thresholds among standards it does not specify a particular threshold in precise terms; and, even if it did, it would be impossible to ensure that all judicial minds apply it in the same way. Consistent with this perspective, scholars have also found that shifts in the burden of proof in the context of the U.S. Tax Court did not impact decision-making in intended ways (Finley and Karnes, 2008).

Our research team also spoke with several immigration attorneys who expressed skepticism regarding *Brito*. They identified flexible interpretations, discretion, and barriers to reviewing an IJ’s decisions as tactics the government would use to make it appear it complied with *Brito* (RI #10, #51, and #52). These attorneys also expected the government to fight the implementation on other fronts, including

foot-dragging its enforcement, complying with only parts of the decision, and appealing the decision to the First Circuit.⁵

In summary, while there are some general features of US immigration law that suggest that federal courts ought to be able to control immigration courts, there are also features of US immigration law that suggest that federal courts will not be able to do so effectively, at least through the kind of decision issued in *Brito*. Critically, even if *Brito* impacted the Boston Immigration Court, we have no clear account of how it could have done so. As we discuss in the next section, if the decision affected outcomes, it would have done so in surprising ways.

4 A Theoretical Account of Immigration Bond Decisions

In the Appendix, we include a game theoretic model of immigration bond hearings that fully characterizes our argument. We focus our argument on the equilibrium that characterizes well basic features of the immigration bond process.⁶ Here we summarize its key elements in four steps. We provide a conceptual account of the burden of proof as it is understood legally. We then characterize the bond decision-making process, drawing on past work that has provided a statistical interpretation of the burden of proof. We describe how *Brito* would have changed IJ decision-making if it only affected IJ behavior. Finally, we consider how *Brito* could have changed outcomes by influencing not only IJs but the decisions of ICE when making initial bond decisions. Our argument reveals that *Brito* would have increased the rate of granting bond is if it only influenced IJ decision-making. If instead, it influenced ICE as well, the decision would have decreased the Boston immigration court’s caseload and lowered the observed rate of bond.

⁵In June 2022, the U.S. Government filed a petition to have *Brito’s* burden ruling reconsidered by the full (“en banc”) group of First Circuit judges.

⁶We do not consider for the purposes of empirical implications equilibria in which the immigration courts have no docket, in which ICE grants bond to absolutely no person, or in which immigration courts never grant bond when asked. We characterize the conditions under which those equilibria exist, but for our purposes they are not particularly helpful in understanding well-known empirical patterns.

4.1 The Burden of Proof

The burden of proof has two components: the burden of production and the burden of persuasion. The burden of production imposes an obligation to offer sufficient evidence to show that a factual claim warrants a decision by the trier of fact. The burden of persuasion is the responsibility to persuade the trier of fact that a factual claim should be accepted. These concepts derive from the idea that decisions to accept or reject a factual assertion in court carry risks for the parties. Erroneously accepting that a defendant has committed a crime unjustly impinges upon the defendant's liberty rights. The Supreme Court describes the allocation of the burden of proof as a decision over which litigant should carry the greater risk of a wrongful decision against her.⁷ The litigant with the most at stake will suffer the most if the process reaches a wrongful decision. For this reason, the litigant with less at stake receives the burden of proof. Carrying the burden of proof effectively puts a thumb on the scale of evidence, weighting it against the litigant that carries it (Wexler, 1999, p. 75).

Assuming that the party satisfies the burden of production, the party still carries the burden of persuading the trier of fact. Standards of proof are criteria that guide decisions about the party's persuasiveness. In a sense, the standard of proof determines how much benefit of the doubt to give the litigant entitled to it. This assessment again requires a balancing of each litigant's interests. The weightier the interests of the litigant with the benefit of the doubt relative to the litigant with the burden of proof, the greater the benefit of the doubt offered. The Supreme Court recognizes three different standards of proof. *Preponderance of the evidence*, which requires the litigant with the burden of proof to prove her facts are more likely than not, tolerates the greatest amount of risk of a wrongful decision and provides the least benefit of the doubt.⁸ *Clear and convincing evidence* tolerates a lower risk of error against the party with the most to lose and provides a greater benefit of the doubt, requiring the party with the burden of proof to present evidence that makes it "highly probable" that her factual

⁷Addington v. Texas, 441 U.S. 418, 423 (U.S. Sup. Ct. 1979)

⁸The Supreme Court describes this standard as dividing "the risk of error in roughly equal fashion" between litigants and particularly appropriate where "mere loss of money" is at stake. Ibid, 424.

contentions are true.⁹ The *beyond a reasonable doubt* standard tolerates almost no risk of an error harming the party with the most to lose and, therefore, provides the greatest benefit of the doubt. It is reserved for criminal cases where “the interests of the defendant are of such magnitude that historically and without any explicit constitutional requirement they have been protected by standards of proof designed to exclude as nearly as possible the likelihood of an erroneous judgment.”¹⁰

4.2 The Decision to Grant or Deny Bond

Our theoretical account adopts a “case-space” approach to modeling judicial decision-making (Kornhauser, 1992; Lax, 2011). IJs evaluate facts that reflect an individual’s true but unobserved dangerousness and flight risk. The problem IJs confront is that the evidence they observe imperfectly reveals the truth and thus they must draw inferences. Legal standards are thresholds that IJs use to evaluate how confident they need to be about a factual assertion’s truth (e.g., “The respondent is dangerous.”) in order to deny bond. We assume that IJs are motivated simply to arrive at the “correct” disposition of bond cases, i.e., to grant (deny) bond to those who (do not) deserve it. This simple dispositional utility function nevertheless captures key features of immigration court decision-making including potential pressures that IJs confront from their superiors (Cameron and Kornhauser, 2024).¹¹

We focus on two steps of the bond process following the detention of an individual who is subject to removal proceedings. The first step involves ICE evaluating the facts of the case and making an initial determination about bond. Should ICE deny bond,¹² we will assume that the decision is reviewed by

⁹The Supreme Court considers this standard is appropriate where the litigant with the benefit of the doubt has a “particularly important interest” at stake. *See* *Addington v. Texas*, 441 U.S. 424 (U.S. Sup. Ct. 1979) and *Colorado v. New Mexico*, 467 U.S. 310, 316 (1984) defining clear and convincing evidence.

¹⁰*Ibid*, 423.

¹¹Clearly, we might imagine that the law has nothing to do with bond decisions. In such an approach, we might believe that *Brito* would have had no effect on outcomes. Our research design accounts for this possibility.

¹²We set aside the issue of setting an amount of bond.

an IJ.¹³ The **burden of production** is modeled as the cost that ICE faces if it must argue for a denial of bond before an IJ. We argue that it is more costly to make such an argument when ICE carries the burden of production.¹⁴ The second step is IJ review. The IJ observes evidence presented in a bond hearing and asks whether it is sufficiently likely that an individual is either dangerous or a flight risk (and thus does not warrant bond). Sufficiency is defined by the relevant standard of proof, which emerges endogenously.

We adopt a familiar statistical interpretation of standards of proof (e.g. Cheng, 2012; Hay and Spier, 1997). We will suppose that an IJ approaches the record with a prior belief about the truth of the assertion that a respondent does not warrant bond, π , which can be interpreted as a probability that the respondent is either a flight risk or dangerous. We assume that the evidence observed by the IJ can be understood as a noisy signal about whether the respondent warrants bond. We suppose that IJs observe either “strong” or “weak” evidence that the respondent is either dangerous or a flight risk. We assume further that IJs will see strong evidence in cases where individuals truly warrant denial; however, the evidence may appear strong even if denying bond is the wrong decision. That is, some factual records point toward detention even if the person is neither truly a flight risk nor dangerous. We let q , the “coarseness” of the signal, represent the probability that the evidence is strong even though the individual truly warrants bond.

Finally, we assume that ICE has better information about the true state of affairs than does the IJ.¹⁵ ICE is motivated to detain all individuals, yet the costs of releasing a truly dangerous individual are higher than those associated with releasing an individual who is not. This reflects a “blame avoidance” logic among public officials derived from beliefs that negative outcomes will be weighed more significantly

¹³We address this assumption in the Appendix. The consequence of modeling the individual’s choice to seek review is not material as long as ICE takes account of changing incentives to seek review when it makes an initial determination on bond, which we could do at no material benefit to the argument.

¹⁴We also assume that the costs of making a case to deny bond increase if the case facts generally point to an individual who is not dangerous than if they point to an individual who is dangerous.

¹⁵Our model assumes that ICE observes the individual’s true dangerousness/flight risk perfectly.

by principals and the public generally than positive outcomes (e.g., Hood, 2007).¹⁶ The upshot of this is that ICE always detains individuals whom it believes warrant a denial, and yet ICE sometimes detains individuals it believes to be truly deserving of bond. ICE detains individuals who deserve it with a probability denoted λ .

4.2.1 IJ Decision-making

Detaining a respondent who does not warrant bond or granting bond to a respondent who does are correct decisions. In contrast, granting bond to a respondent who does not warrant bond or detaining a respondent who does reflect two types errors in decision-making. We assume that incorrect decisions are costly. We will say that the cost of the first type of error is $\beta > 0$ and that the cost of the second type of error is $\alpha > 0$. Correct decisions are costless.¹⁷

After observing evidence, the IJ's belief about whether the respondent warrants bond can be understood as a Bayesian posterior. Specifically, IJ's posterior belief is the probability that the assertion is true, given the evidence that she has observed: $Pr(\text{Respondent does not warrant bond}|\text{evidence}) \equiv \hat{\pi}$. A standard of proof is then a threshold indicating the value of the posterior belief above which the judge should conclude that the assertion is true. If we continue to assume that IJs are trying to make correct decisions, the expected cost of granting bond is $-\hat{\pi}\beta$ and that the expected cost of denying bond is $(1 - \hat{\pi})(-\alpha)$. Thus, the IJ should grant bond if and only if $\hat{\pi} < \bar{\pi}$, or if and only if

¹⁶The most obvious negative outcome would be to release a truly dangerous person who commits a violent crime upon release.

¹⁷In a hypothesis testing framework commonly applied to liberty questions before courts, we might say that the null hypothesis is that the respondent is neither dangerous nor a flight risk, and thus should be granted bond. We do not describe the setting in this way since the burden of proof is itself in flux and thus what the null hypothesis should be is unclear. That said, if you were to take this approach than in our case, a Type I error would be to deny bond when bond should be granted, the cost of which is α . A Type II error is to grant bond to someone who does not warrant it, and the cost of this error is β .

$$\underbrace{\frac{\pi}{\pi + (1 - \pi)\lambda q}}_{\text{Pr(Does not warrant bond | strong evidence)}} < \underbrace{\frac{\alpha}{\alpha + \beta}}_{\text{Standard of proof}}. \quad (1)$$

This decision rule reflects the conceptual connection between error costs and standards of proof, which in turn influences what it means for a litigant to carry the **burden of persuasion**. In our framework, error costs fully characterize any standard of proof. For example, setting $\alpha = \beta$ yields the “preponderance of the evidence” standard. Similarly, setting $\alpha = 3\beta$ yields something resembling “clear and convincing evidence,” and setting $\alpha = 19\beta$ yields something like the “beyond a reasonable doubt” standard. An instruction to use a particular standard of proof is a statement about how judges ought to weigh the costs associated with decision errors. Requiring a stricter standard of proof in a bond hearing is to communicate that judges should assign a higher cost to detaining a peaceful person (i.e., α increased), or that they should assign lower costs to granting bond to a person who is truly dangerous (i.e., β decreased), or both. In summary, on this account, when the District Court fixed standards of proof in *Brito*, it was instructing the Boston Immigration Court about how to weigh the costs of decision errors, and thus how heavy the burden of persuasion would be for the government.

4.2.2 Empirical Implications of this Argument

The empirical implications of the argument depend on whether we believe that *Brito* would have affected both ICE and IJs or only IJs. If only IJs responded to *Brito*, then the setting is decision-theoretic. The implications of *Brito* for this setting only requires inspection of Expression 1 for a fixed $\lambda > 0$. In this case, bond rates would increase after the burden of proof shifted.¹⁸ To see how, consider the left-hand side of the inequality in (1), which represents an IJ’s belief about the respondent after observing strong evidence. These beliefs would not have been changed by the decision since λ (under the control of ICE) would not have changed. However, the right-hand side of (1) would have shifted upward due to an increase in α , a decrease in β or both. All else equal, granting bond would have become more likely.

¹⁸We assume that the best interpretation of the standard “to the satisfaction of the IJ” is preponderance of the evidence.

But what happens if ICE also responded to *Brito*? This requires consideration of how behavior would have changed in a game theoretic equilibrium (see Appendix, pp. 41-44). In the equilibrium we consider, ICE denies bond to peaceful people with probability λ and the IJ grants bond with probability p . The former probability reflect the optimal denial rate for ICE given its expectations about the IJ's bond granting behavior; the latter probability reflects the optimal bond grant rate given the IJ's information, beliefs, and expectations about ICE's bond denial behavior. The decision had two effects. First, *Brito* increased the burden of production by making it harder for ICE to win cases with weak facts (i.e., increasing ICE's costs of denying bond to a peaceful person). When the burden of production rested with the respondent, detained individuals were required to produce evidence showing that they are neither flight risks nor dangerous. ICE would not have had to put on much of a case. When the burden shifted, ICE was required to produce evidence showing dangerousness or flight risk. If the facts were weak, this would clearly have been more difficult than simply responding to whatever respondents could have pulled together when they held the burden, remembering that they largely did so without the assistance of counsel, from within a prison, and often in a foreign language. The direct impact of this change in equilibrium is a reduction in the rate that IJs grant bond (p^* in the Appendix). IJs internalize ICE's changed incentives. In equilibrium, ICE continues to deny bond to some peaceful individuals, and that requires the bond grant rate to decline.

The second effect of *Brito* runs through the costs of decision errors (by increasing α , decreasing β , or both). These changes decrease the rate at which ICE denies bond to peaceful individuals. As IJs reduce their grant rates, ICE must reduce the rate at which denies bond unjustifiably. Like IJs, ICE prosecutors would have internalized the IJ's changed incentives related to the burden of persuasion.¹⁹ A complementary effect of *Brito* is that this decrease in the rate of unjustifiable denials would have resulted in stronger IJ beliefs in the dangerousness of respondents conditional on observing strong evidence. And that in turn would have rationalized the IJ's lower bond grant rates.

This argument leads to three empirical implications.

¹⁹As we show in the Appendix, this is possible only if the signal linking a respondent's true dangerousness or flight risk and the evidence observed is sufficiently coarse, i.e., the evidence cannot be overwhelmingly discriminating.

Independent ICE Decisions (Bond Rates): If ICE decisions about setting bond were not affected by the *Brito* decision, then *Brito* would have increased the rate at which Massachusetts IJs grant bond relative to immigration court judges not subject to *Brito*.

Dependent ICE Decisions (Caseload): If ICE decisions about setting bond were affected by the *Brito* decision, then *Brito* would have decreased the caseload of the Massachusetts Immigration Court relative to immigration courts not subject to *Brito*.

Dependent ICE Decisions (Bond Rates): If ICE decisions about setting bond were affected by the *Brito* decision, then *Brito* would have decreased the rate at which Massachusetts IJs grant bond relative to immigration court judges not subject to *Brito*.

5 An Empirical Study of *Brito v. Barr*

Brito offers an opportunity to evaluate the impact of changing the burden of proof in immigration custody decisions. The question is to which immigration judges or immigration courts should the judges in Boston be compared. We use the synthetic control method (Abadie and Gardeazabal, 2003; Abadie, Diamond and Hainmueller, 2010, 2015, 2020) to estimate *Brito*'s effect on immigration courts and judges.²⁰ We thus construct a counterfactual, “synthetic” Boston Immigration Court (which was not exposed to *Brito*) by combining the experiences of other U.S. immigration courts such that the trends in bond rates and caseloads approximate the trends for the Boston Immigration Court prior to *Brito*.

5.1 Data and Sample

We analyze publicly available data from the EOIR.²¹ Specifically, we use weekly hearing base city (HBC) panel data from March 19th, 2019, through March 13th, 2020. An HBC is the administrative designation for the immigration court having jurisdiction over the assigned hearing locations. The EOIR data contain over 70 unique HBC codes and over 400 unique hearing locations. Our sample contains 38 pre-treatment periods because the synthetic control method requires a perfectly balanced panel,

²⁰We use Boston and Massachusetts interchangeably because Boston is the only HBC in Massachusetts.

²¹Retrieved January 4, 2021, from <https://www.justice.gov/eoir/foia-library-0>.

and starting earlier forces us to drop HBCs from the donor pool. We truncate the sample the week President Trump declared COVID-19 a National Emergency on Friday, March 13th, 2020. We do this because COVID did not impact all jurisdictions simultaneously, and not all jurisdictions had identical responses to the pandemic. The 13 post-treatment periods are sufficient to evaluate the decision’s effect on bond hearing decisions; the ending ensures our outcome variables are unaffected by COVID or a new administration enacting policy changes.

To augment missing information about IJs, our research team scraped data from investiture announcements and coded relevant information for over 550 IJ’s, including their professional background and the president who appointed them. The result is a rectangular dataset where each row represents a unique bond hearing decision by an IJ. Thus, each row provides information on an IJ’s decision in one of potentially many hearings concerning a detained individual’s claim. We collapse 215,762 bond hearing decisions into an HBC-week panel. This process transforms our variables from observed values into rates or averages.²²

The synthetic must not include courts exposed to similar rules as *Brito* imposed on Boston. Our team contacted immigration lawyers within the jurisdiction of each US immigration court to determine the operational burdens of production and persuasion. At the time of the study, no other jurisdiction had yet experienced a change similar to Boston’s. A subsequent legal review found two other federal courts enacted legal rules identical to *Brito*, but those changes occurred two and seven months after our study period. Archival evidence from government documents confirms this and guides IJs to conduct bond hearings in the same manner in December 2020 as November 2017 (Office of the Chief Immigration Judge, 2017, 2020*a,b*). We are confident that Massachusetts is the only jurisdiction where the burden of proof was shifted as required by the district court.²³

5.1.1 Outcome Variables

Our first outcome, *caseload*, is a simple count of the number of bond hearing decisions made in the jurisdiction in a particular period. Our second outcome, *rate granted bond*, is the percentage of bond

²²Bond hearing decisions provide a concrete example: *rate granted bond*_{*i,t*} represents the percentage of decisions in which IJs in HBC *i* during week *t* granted bonded release.

²³We remove from the donor pool jurisdictions making fewer than four decisions per week.

hearings in which IJs granted bond.²⁴ The data required extensive cleaning and we kept all observations with bond decision dates before 2019 and after and after President Trump declared COVID-19 a National Emergency on March 13th, 2020, leaving an initial dataset containing 201,469 observations. We removed fewer than 5% of observations in which data are missing or the EOIR code for decision type indicates that the decision is outside our scope, undefined, or illogical (see Appendix for further details).²⁵

Table 1 summarizes the number of decisions made by Democratic- and Republican-appointed IJs throughout the period: across all 24 HBCs, IJs made 71,983 decisions and granted bond 41% of the time. The table also suggests that across the immigration court system, there were no meaningful differences across judges appointed during either Democratic or Republican administrations.²⁶ Pre-*Brito*, IJs decided 1,916 Massachusetts bond hearings, with Democratic appointees granting bonded release in 26% of their 243 decisions. Republican appointees granted bond at a higher rate on seven times as many decisions (1,673).

By the time the District Court of Massachusetts announced its *Brito* opinion on November 27th, 2019, IJs appointed by Democratic presidents were active in Boston immigration courts but were tasked with responsibilities other than making bond hearing decisions.²⁷ The lack of decisions by IJs appointed by Democratic presidents post-*Brito* is unavoidable but reflects the reality of the docket at the time.

We use predictor variables in the pre-intervention period to identify the weight each unit in the donor pool contributes to the synthetic Boston. *Respondent Criminal Charge* is the percentage of decisions in which DHS alleges the respondent violated a criminal charge. *Respondent Number of Appearances* is the average number of times, between 1 and 2, a respondent appears before an IJ for a decision. Three

²⁴Appendix Table 3 shows our coding protocol for this outcome.

²⁵Applying the synthetic control method retains only 24 hearing base cities and observations between March 19, 2019, and March 13, 2020. Removing these base cities and dates in the data-clearing process leaves the 71,983 observations shown in Table 1.

²⁶Across all 24 HBCs, both Democratic and Republican IJs granted bond approximately 41% of the time.

²⁷See Appendix Table 4 for information on these 11 IJs. IJs appointed by Democratic presidents made 244 of 2,385 (10.3%) decisions in the Boston immigration court between 03/13/13 and 03/13/20, and all of those decisions occurred on or before *Brito* went into effect on 12/13/19. The rate granted bond by IJs in the Boston court was 31.0% (n=1,954) before *Brito's* effective date and 31.6% (n=431) after.

	Pre- <i>Brito</i> ^a	<i>Brito</i> ^b	Post- <i>Brito</i> ^c	Total
Treated Jurisdiction (Boston)				
—Democratic IJs	243 (25.5%)	-	1 (0.0%)	244 (25.4%)
—Republican IJs	1,673 (31.6%)	81 (30.9%)	444 (31.5%)	2,198 (31.6%)
	1,916 (30.8%)	81 (30.9%)	445 (31.5%)	2,442 (31.0%)
Donor Pool (x23 HBCs)				
—Democratic IJs	16,348 (41.7%)	992 (43.5%)	4,731 (39.8%)	22,071 (41.4%)
—Republican IJs	35,757 (41.7%)	2,006 (41.8%)	9,707 (39.3%)	47,470 (41.2%)
	52,105 (41.7%)	2,998 (42.4%)	14,438 (39.5%)	69,541 (41.3%)
All Jurisdictions (x24 HBCs)				
—Democratic IJs	16,591 (41.5%)	992 (43.5%)	4,732 (39.8%)	22,315 (41.2%)
—Republican IJs	37,430 (41.2%)	2,087 (41.4%)	10,151 (39.0%)	49,668 (40.8%)
	54,021 (41.3%)	3,079 (42.1%)	14,883 (39.2%)	71,983 (40.9%)

^a 03/19/19 thru 11/26/19. *Brito* opinion published 11/27/19.

^b 11/27/19 thru 12/13/19. *Brito* ordered effective by 12/13/19.

^c 12/14/19 thru 03/13/20. COVID-19 declared a National Emergency 03/13/20.

Table 1: *Number of decisions from IJs appointed by different presidents. Values in parentheses indicate percentage of decisions granting bonded release. Reported for decisions in 24 Hearing Base Cities 03/19/19 through 03/13/20.*

variables account for differences between judges. *IJ Trump Appointee* and *IJ Obama Appointee* measure the percentage of decisions in a jurisdiction made by IJs appointed under these administrations. *IJ Years of Experience* is a weighted average of IJs’ years of experience as an IJ. The proportion of decisions made by a particular IJ corresponds to the weight the IJ’s experience lends to the calculation. Three socioeconomic controls account for differences between jurisdictions: annual state GDP in millions of dollars transformed into equally-weighted monthly values (logged), state population in thousands of persons (logged), and government employees per capita.²⁸ All analyses utilize outcome variables in prior periods as predictors to increase the synthetic’s precision.

5.2 Synthetic Control Validation

The validity of the synthetic control approach depends on its ability to approximate pre-*Brito* trends in synthetic Boston. We do not replicate the trends perfectly, but the synthetic Boston is preferred over available alternatives, including the donor pool average or any individual HBC.

²⁸Source: Federal Reserve Bank of St. Louis.

Figure 1a plots the weekly, 38-period progression of caseload in Boston compared to the broader donor pool and synthetic. Between March and August 2019, the synthetic outperforms the donor pool average in replicating Boston’s caseload. For the remainder of the pre-treatment period, the donor pool average and the synthetic track Boston’s volatile caseload equally well. Table 2 provides another test. It compares the predictor balances in Boston, the synthetic, and the donor pool average to assess the synthetic’s relative accuracy. The average of HBCs in the donor pool does not provide a suitable control group for Boston. Instead, the synthetic Boston more accurately reproduces the values of 10 of the 12 predictor variables prior to *Brito*; keeping these two variables maximizes the synthetic’s precision. A combination of seven HBCs in Arizona, Louisiana, Nebraska, New Jersey, and New York reproduce Boston’s caseload with the most accuracy (see Appendix Table 5 for precise weightings).

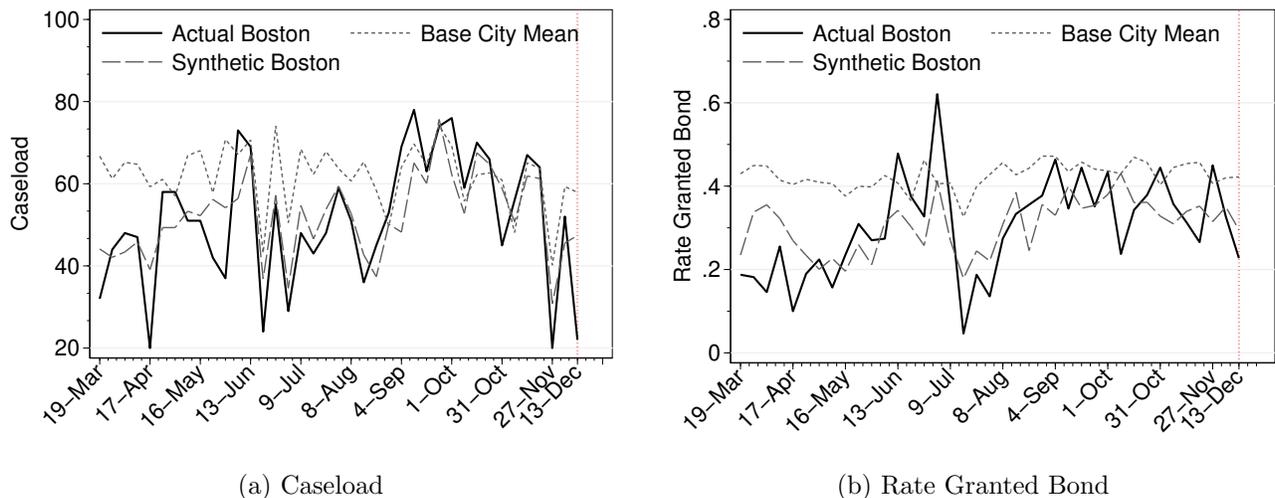


Figure 1: *Weekly Progression of Caseload and Rate Granted Bond in Boston, the Donor Pool, and Synthetic Boston.*

Figure 1b compares the rate granted bond in Boston, the broader donor pool, and the synthetic control. The rate granted bond in Boston generally ranged between 20% to 45%. Meanwhile, the base city mean was regularly above 40%. The synthetic mirrors Boston’s week-to-week movements closely, and Table 2 shows that the average of HBCs unaffected by *Brito* does not provide a suitable control group for Boston. The synthetic Boston more accurately reproduces the values of 10 of the 12 predictor variables prior to the decision; keeping the underperforming variables maximizes the sythetic’s precision. A weighted combination of HBCs in Adelanto, CA (3%); Jena, LA (49%); and Tacoma, WA (48%) reproduce Boston’s rate granted bond with the most accuracy.

Predictor	Caseload			Rate Granted Bond		
	Boston			Boston		
	Actual	Synth ^a	Pool ^b	Actual	Synth ^a	Pool ^b
IJ Trump Appointee	0.86	0.46	0.51	0.86	0.63	0.51
IJ Obama Appointee	0.13	0.15	0.26	0.13	0.18	0.26
IJ Years of Experience	1.97	9.17	6.59	1.97	4.94	6.59
Respondent Criminal Charge	0.05	0.05	0.05	0.05	0.05	0.05
Respondent Number of Appearances	1.42	1.34	1.29	1.42	1.33	1.29
State GDP (logged)	13.25	12.98	13.33	13.25	12.91	13.33
State Population (logged)	15.75	15.80	16.13	15.75	15.65	16.13
State Government Employees/Capita	0.07	0.07	0.07	0.07	0.07	0.07
Caseload (<i>week</i> = 8)	51.00	43.00	66.87			
Caseload (<i>week</i> = 22)	36.00	45.29	65.30			
Caseload (<i>week</i> = 24)	53.00	53.01	49.70			
Caseload (<i>week</i> = 37)	20.00	30.94	40.17			
Rate Granted Bond (<i>week</i> = 13)				0.48	0.34	0.41
Rate Granted Bond (<i>week</i> = 17)				0.35	0.27	0.41
Rate Granted Bond (<i>week</i> = 20)				0.14	0.22	0.43
Rate Granted Bond (<i>week</i> = 38)				0.33	0.35	0.42

^a Weighted combination of donor units selected to approximate response variables in Boston in terms of the outcome predictors. All predictors averaged over pretreatment period.

^b Average for jurisdictions in the donor pool excluding Boston, MA.

Note: Pretreatment (weeks 1-38), treatment (week 39), posttreatment (weeks 40-52).

Table 2: *Predictor Means. Variables averaged over pre-intervention period but excluding the intervention period. Predictors selected to minimize MSPE in the pre-treatment period.*

5.3 Results

If the decision affected both ICE and IJs, we expect to observe that the difference between the actual Boston’s and synthetic Boston’s outcomes should be negative and significant in the post-treatment periods for each outcome. If the decision only affected IJs, we would expect to observe a positive difference between actual and synthetic Boston for the grant rate only.

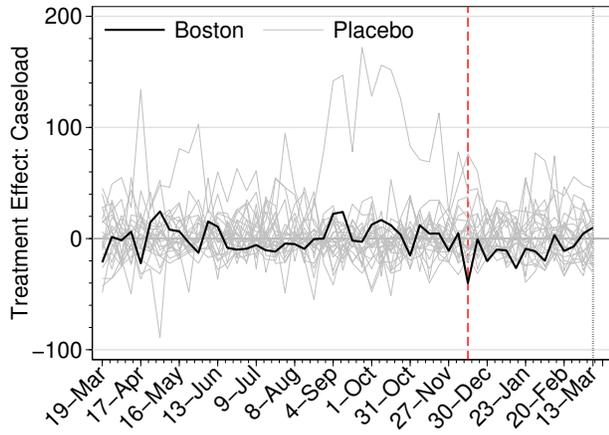
The thick lines in Figure 2 plot these differences. Figure 2a indicates a negative difference in caseload between Boston and synthetic Boston after *Brito*, at least in weeks of December, January, and February. The difference in rate granted bond (Figure 2b) is positive in December but is largely negative by January and remains so with two exceptions. Taken together, these initial findings appear to confirm our theoretical predictions on how *Brito* affected IJ decision-making or prosecutorial discretion, assuming that the decision impacted both ICE and the IJs.

5.3.1 Inference About *Brito*'s Effect on Boston IJs

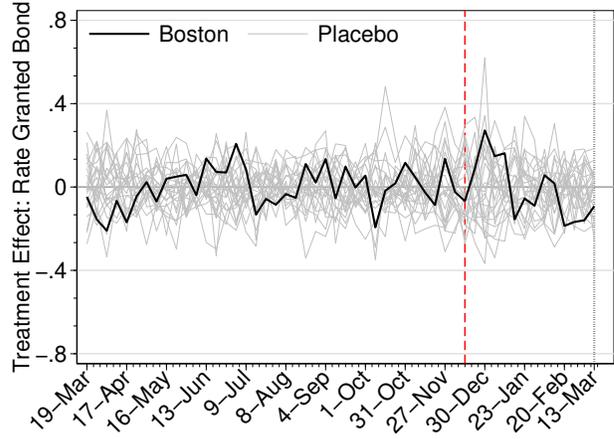
To evaluate *Brito*'s statistical significance, we create a reference distribution through the use of placebo tests. As in Abadie, Diamond and Hainmueller (2010), we conduct placebo studies by applying the synthetic control method to 23 donor pool units we know *Brito* did not affect (see also Abadie and Gardeazabal, 2003). That is, we assume that each of these 23 units was subject to *Brito*, and we carry out the same synthetic control analysis. This results in a distribution of placebo effects to which we can compare the actual effects from Boston. If the Boston effects are extreme relative to the placebo distribution, we can conclude that they are unlikely to have been caused by chance alone. If, on the other hand, the Boston effects are not particularly extreme relative to the placebo distribution, then we should be more skeptical about the possibility that *Brito* affected either IJ decision-making or prosecutorial discretion.

In each placebo iteration, we reassign in our data the *Brito* intervention to one of the donor pool units, replacing that unit with Boston in the donor pool. We then compute the estimated effect associated with each placebo run. This iterative procedure provides us with a distribution of estimated gaps for the states where no intervention took place. Figures 2a and 2b display placebo study results. Thinner lines represent the difference associated with each placebo test. Each line measures the difference in the outcomes between each HBC in the donor pool and its corresponding synthetic. As Figures 2a and 2b make apparent, the estimated difference for Boston after the decision's December 13 effective date is strong and negative compared to the distribution of other HBCs in the donor pool. The figures also indicate that caseloads and rates granted bond exhibit a noisy data generating process during the 38-week pretreatment period, which cannot be well reproduced for some HBCs using an amalgam of caseloads and rates granted bond in other HBCs.

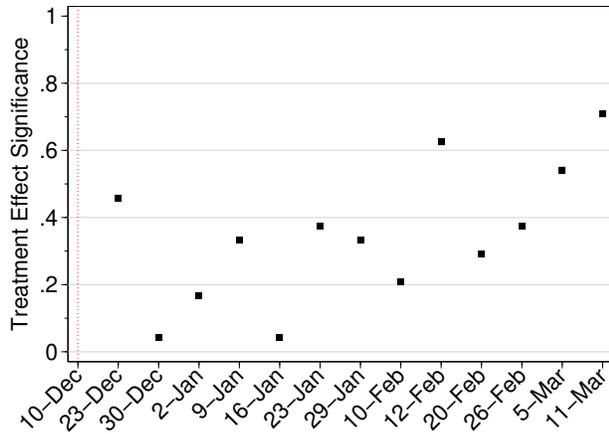
Figures 2a and 2b also show the distribution of causal effects for our outcomes for Boston and the donor pool. We next inspect the distribution of weekly treatment effects in the post-treatment period and test the sharp null hypothesis that the treatment had no impact on any unit. We reject this null by estimating Fisher-exact P-values corresponding to the proportion of treatment effects that are as strong or stronger than the negative effects we predict. Because each figure includes 23 control units, the probability of estimating a difference of the magnitude of the one observed in Boston under a random permutation of the *Brito* intervention in our data is 0.043, a level typically used in conventional tests



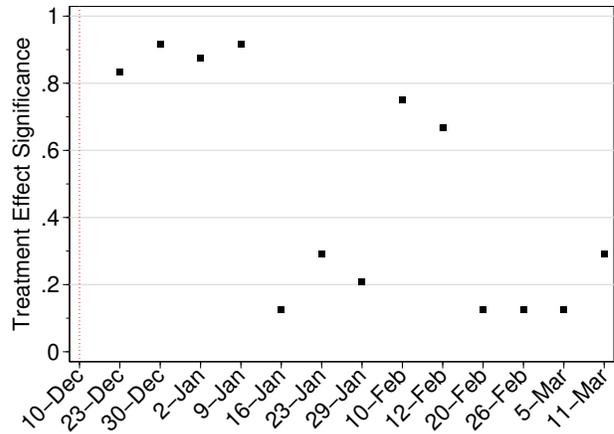
(a) Caseload differences



(b) Rate Granted Bond differences



(c) Caseload treatment significance



(d) Rate Granted Bond treatment significance

Figure 2: Subfigures (a) and (b) show outcome differences between actual and synthetic units for Boston (thick line) and placebos (thin lines). Subfigures (c) and (d) show randomization inference of treatment effect on outcomes using Fisher-exact p -values.

of statistical significance.²⁹ The test is one-sided because we predict the decision negatively affected caseloads and the rate granted bond.

Brito's effect on caseload is significant in weeks two ($p = 0.042$) and five ($p = 0.042$) after the decision (Figure 2c). Fisher-exact P-value for rate granted bond show the effect is borderline significant in weeks five ($p = 0.125$), ten ($p = 0.125$), eleven ($p = 0.125$), and twelve ($p = 0.125$) after the enforcement date (Figure 2d). These effects are neither strong enough nor consistent enough to support our argument that *Brito* reshaped IJ decision-making or prosecutorial discretion.

²⁹The minimum attainable Fisher-exact P-value that can be achieved is $1/(\text{number of jurisdictions})$.

One final way to assess *Brito's* treatment effect relative to the placebo runs is to look at the distribution of the post-/pre-treatment mean squared prediction error (MSPE) ratios. The pre-MSPE is the squared average deviation between the Boston series and the synthetic control before *Brito*, and the post-MSPE is the squared average deviation between the Boston series and the synthetic control after the decision. Inflated post-/pre-treatment MSPE ratios identify synthetics that are less accurate post-treatment than pre-treatment, which they should be if the treatment had an effect. Where a ratio falls in the distribution of all ratios is the probability of obtaining a similar MSPE ratio if the treatment were randomly assigned. We infer that the treatment effect is significant if its prediction error ratio is an outlier in the distribution.

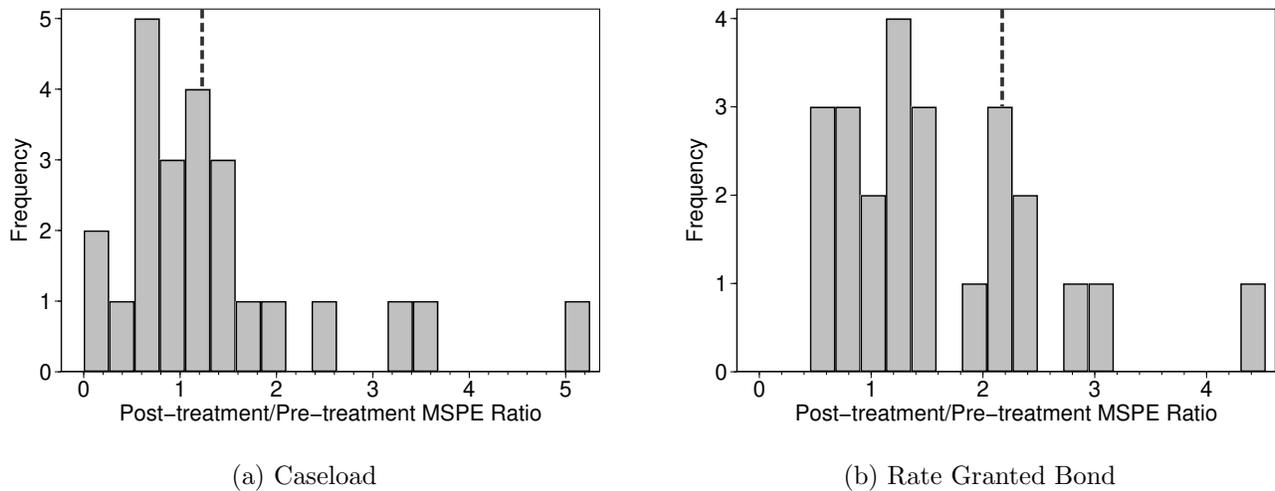


Figure 3: *Post-treatment/pre-treatment MSPE Ratios*. Dashed line represents *Boston's* value within the 24-unit sample.

Figure 3 shows the distribution of ratios for all HBCs for caseload and rate granted bond. Boston's ratio is 1.2 for caseload, ranking it 10th among 24 jurisdictions (Figure 3a). Boston's caseload ratio location means that if one were to assign the intervention randomly, the probability of obtaining a ratio as large as Boston's is 0.42 (10/24 = 0.417). Figure 3b communicates similar information for rate granted bond. Boston's ratio is 2.2, ranking it 7th in size. This means that if the intervention were assigned randomly, the probability of obtaining a ratio as large as Boston's is 0.29 (7/24 = 0.292). Boston's pedestrian MSPE ratios are evidence against the assertion that *Brito* changed bond hearings. Taken together, these placebo studies put *Brito's* treatment effect into perspective: a negative but not necessarily significant nor permanent effect on caseload and rate granted bond.

5.3.2 Interpreting Findings with Expert Interviews

Our research team used qualitative methods to explore the *Brito* decision, providing insights that increased our confidence in our relatively weak findings. As one attorney shared: “I’ve definitely come home from bond hearings before where I’ve said *Brito* was just a dream. It was so good when it happened, but it didn’t mean anything. It didn’t have impact” (Research Interview #69). Elite interviews and document analysis to identify six factors that stopped *Brito* from achieving its expected effects.

First, poor implementation undermines *Brito*’s effect on immigration courts and judges. The National Association of Immigration Judges (NAIJ) suggests IJs cannot keep up with rulings and cannot implement directives as required. It argues that the EOIR does not give IJs the time to keep current on legal development and does not provide IJs adequate training (National Association of Immigration Judges, 2019, 3). Immigration attorneys argue ignorance of the decision is no excuse. Interview participants report EOIR and ICE had full knowledge of *Brito* and could easily comply with the clearly defined standards and burdens it set forth (RI #69, 71). These courtroom professionals explained that no institutions effectively hold IJs accountable to *Brito*’s directives.

Another factor limiting *Brito*’s effect on immigration courts and judges is an IJ’s ability to ignore or shirk *Brito*’s directives. Numerous immigration attorneys reported that IJs and ICE personnel ignore *Brito* and maintain the status quo wherever possible (RI #7, 10, 51, 52). An attorney who once had high hopes for the decision now sees it as impotent: “Judges are very careful to at least make it look as if they’re following Brito...‘I’m going to recite these magic words. And I’m going to say that there is clear and convincing evidence of dangerousness, and that there is no reasonable alternative to detention that can protect the community or ensure the client’s safety.’...Once you have a judge who said all the magic words, even if the evidence doesn’t meet the clear and convincing standard, the district court will probably say no jurisdiction to review that” (RI #69). A common theme from these interviews is that IJs are aware of *Brito* but are unwilling to incorporate its directives.

Third, *Brito* does not change the reality that immigrants are often unrepresented, non-English speakers forced to square off against an ICE attorney who has litigated hundreds or thousands of cases (National Association of Immigration Judges, 2020, 5). Several immigration attorneys explained that despite the due process *Brito* promises, detained immigrants who lack competent representation are

unable to assert their rights during a fast-tempo bond hearing (RI #7, 23, 52). We expect representation to affect individual outcomes positively (see also Ryo and Peacock, 2021), but our results are unchanged when using representation as a predictor in our synthetic control.

Contrasting directives from political appointees in the executive is the fourth factor why the ruling did not affect immigration courts and judges. The NAIJ has made multiple public statements warning lawmakers that IJs are highly susceptible to political influence. The NAIJ argues that the DOJ considers IJs not as independent judges but as mere attorneys employed “to enforce the political will of the then current administration” and creates profound conflicts of interest concerning decisional independence and due process (National Association of Immigration Judges, 2021, 1).

Fifth, the administration’s priorities guidance and professional incentives erect perverse incentives for IJs. An administration’s priorities guidance directs immigration enforcement activities, and IJs comment that these directives are “drastic pendulum swings” that inhibit consistent policy between administrations (National Association of Immigration Judges, 2020, 15). The Trump administration directed enforcement that detained individuals with minor offenses (RI #7, 10, 64, 69).³⁰ These directives inflated the case backlog from 500,000 to over one-million between 2017 to 2019 (National Association of Immigration Judges, 2020, 2). At the cost of \$99 per day to detain a noncitizen immigrant (RI #10), the administration used the backlog to justify policies that demanded IJs make more decisions at faster rates. In October 2018, the DOJ implemented a performance management system that subjected IJs to arbitrary quotas and deadlines. Immigration attorneys argue the system is at odds with due process (RI #64), and IJs assert the performance metrics empower the EOIR “to dismiss judges who fail to follow their policy preferences under the pretext of inadequate performance” (National Association of Immigration Judges, 2020, 7). Unchanged caseloads can be an artifact of a large backlog coupled with IJs responding to incentives to keep their judgeships. Meanwhile, high caseloads provide ICE attorneys opportunities for advancement because building trial experience is how individuals advance within the DHS (RI #51, 69), which is why ICE trial attorneys “still try to put on their case even when their cases are weak” (RI #7).

³⁰By contrast, the Biden administration’s directive (09/30/2021) narrowed enforcement and prioritized those individuals that threaten national security, public safety, and border security.

Finally, immigration attorneys identify faulty enforcement mechanisms as the final reason *Brito* did not affect immigration courts and judges. Perhaps the greatest challenge preventing IJs from complying with *Brito* is the inability to review their decisions as long as the IJ completes the analysis. Respondents clarified that 8 US Code 1226 Section E, denies federal courts the jurisdiction to review a discretionary detention decision, making litigating the standard’s application to bond hearing a significant challenge (RI #69). IJs appear to take full advantage of these rules. Respondents concurred that *Brito* changed bond hearings procedures and shifted the burden of production and persuasion to the government. Yet, attorneys observe that IJs have not adjusted their decisions accordingly and offer several explanations for why. An IJ’s personality profoundly affects their decisions, and “having a precise legal standard [does not] take that out of the equation” (RI #23).

Moreover, reflecting the blame avoidance logic we drew on, IJs often overestimate immigrants’ risks to communities (RI #10, 69, 71). Two experienced immigration attorneys explained it this way:

“The judge is always [asking,] ‘What am I going to see in the newspaper the next day? What am I going to see on TV? Am I going to see that this person with a drunk driving arrest [from] 10 years ago going to be in five more [accidents and] kill somebody?’” (RI #10)

“[Brito] didn’t change the underlying fact that whose burden it is isn’t going to change the results. [If] you have a long enough criminal record[, which] doesn’t technically bar you from bond, it makes the judge feel that you’re not safe.” (RI #71).

6 Conclusion

We study whether *Brito v. Barr* affected the outcomes of immigration bond decisions in the Boston Immigration Court. The District Court’s decision created an opportunity to evaluate the causal effect of assigning the burden of proof to the government in this process, as immigration detention requirements remained unchanged for the rest of the country. We reviewed existing theoretical literature on the ability of courts to control executive agencies, concluding that there are both reasons to be optimistic and skeptical about whether *Brito* would have an effect. Assuming that the decision would be impactful, we then developed a theoretical model of the burden of proof as applied to the bond process to identify how *Brito* would have worked. This model produces three empirical implications: (1) If *Brito* impacted

IJs only, then it would have increased bond rates in Boston, (2) if it impacted both IJs and ICE, then it would have lowered bond rates in Boston and (3) reduced the caseload of the Boston Immigration Court. Using a synthetic control approach we found suggestive evidence that is consistent with our theoretical argument in the case where both ICE and IJs were affected by the decision. That said, our findings are not strong, and the fairest interpretation of them is that *Brito* had no effect on ICE or IJ behavior.

Scholars have raised concerns about procedural defects in the immigration law space, highlighting problems with the lack of a right to affordable counsel, remote adjudication procedures, and the inability of the federal courts to review several important yet discretionary decisions of IJs (Eagly, 2014; Eagly and Shafer, 2015; Hausman, 2016). Placing the burden of proof on the detained is another element of due process in immigration law that has garnered scholarly and public attention.³¹ Most relevant to our study are Brady (2022) and Holper (2016), who make legal arguments for shifting the burden of proof to the government in detention cases. We believe that Brady’s and Hoper’s normative cases for doing so are strong, but our results do not suggest that shifting the burden will result in meaningful change. Empirically, our findings reflect most closely those of Finley and Karnes (2008), who found that shifting the burden of proof from the taxpayer to the Tax Court in fiscal disputes did not meaningfully change outcomes.

6.1 Implications

We focus on a single decision of a federal court that oversees a single immigration court. We ought to be careful drawing conclusions from this case to other federal courts, other immigration courts, and other decisions. Some speculation may be useful nevertheless to highlight implications and themes for future research. We summarize likely explanations for our findings, drawing on our interviews, our theoretical analysis, and related ideas about judicial-bureaucratic interactions. We then speculate about possible interventions that could address what seems like the most likely (if uncertain) way to describe what happened in Boston.

³¹See for example https://immigrantjustice.org/sites/default/files/content-type/research-item/documents/2023-06/NIJC-Policy-Brief_ICE-Bond-Reform_May-2023.pdf

6.2 Accounting for our Findings

One simple explanation is that the Boston IJs were already applying the rules laid out in *Brito*. Holper (2016) claims that although the detained hold the burden of production in bond hearings, the government routinely reveals evidence of flight risk and dangerousness. Yet, Holper also argues that the poor allocation of the burden of persuasion was the key problem, and there is no reason to believe that Boston was special – that their IJs were placing the burden of persuasion on the government. Three other ideas suggest that this explanation is implausible: the class-action lawsuit would have been unnecessary if this were the case; the Office of the Chief Immigration Judge (2017, 2020*a,b*) would not produce Immigration Court Practice Manuals instructing IJs to apply rules that are different than those the District Court prescribed; and, the Boston IJs who were working on bond cases in this period were overwhelmingly Trump appointees. It is simply hard to believe that the Boston judges were already applying a standard as strict as clean and convincing evidence, much less that they were assigning the burden of persuasion under this standard to the government before the *Brito* decision.

Our interviews with attorneys suggest other explanations: (1) IJs were unaware of the implications of *Brito*, (2) detainees are rarely represented by counsel and thus there are typically no professionals to help them take advantage of an improved legal rule, and (3) strong pressures from above combined with the flexibility of the rule resulted in simple non-compliance. We view the first idea as certainly possible but unlikely. *Brito* was a major case for the Boston Immigration Court, and we know that IJs applied it at least formally. If the problem was merely about knowledge, we would have seen the results emerge consistently later in our study window. That did not happen. The second idea is also plausible and the lack of access to free counsel creates multiple, considerable problems for the detained. Yet, while the general lack of counsel would have presented a problem, it is hard to square this argument with the experiences of the Boston immigration attorneys we interviewed. They reported their own disappointment about *Brito's* effects, beliefs that they developed through direct observation while representing clients. These are the most likely individuals to report changes in outcomes if lack of counsel was the problem. They did not.

This leaves simple non-compliance. Scholars have found that IJs respond powerfully to incentives created by Attorneys General (e.g. Blasingame et al., 2023). It is certainly possible that these pressures, combined with the inability to clearly audit the use of a standard of proof beyond verifying the so-called

“magic words,” would have resulted in non-compliance. Our own theoretical model does capture this idea in two ways. If pressures from above are strong, then it is entirely possible that ICE’s perceived cost of releasing any detained person would have been high. If this were true, then even a strong increase in the cost associated with the burden of production would not have had a meaningful effect on the IJ’s equilibrium bond grant rate (See Equation 5 on p. 45). The rate would have decreased as we predicted, but for a very large cost of releasing detained people (v_0), the decrease could have been negligible in practice. Similarly, the decision’s effects on an IJ’s perceived error costs could have been negligible if the costs of errors strongly track what the Attorney General wishes them to be. Again, we would still predict a decrease in the rate of ICE denying bond to peaceful individuals (Equation 4 on page 45), but the decrease could have been too small to detect. Indeed, it could have been so small that we would commonly refer to the outcome as simple non-compliance.

If it is in fact true that the Boston Immigration Court simply failed to fairly implement a federal court ruling, it would not have been the first act of simple non-compliance by the immigration bureaucracy during this period. As discussed in Judge Easterbrook’s opinion in *Baez-Sanchez v. Barr* (2020), at roughly the same time as the District Court for Massachusetts issued its decision, the Board of Immigration Appeals (BIA) simply ignored a ruling of the 7th Circuit remanding a case with instructions for further analysis. Judge Easterbrook writes, “What happened next beggars belief. The Board of Immigration Appeals wrote, on the basis of a footnote in a letter the Attorney General issued after our opinion, that our decision is incorrect. Instead of addressing the issues we specified, the Board repeated a theme of its prior decision that the Secretary has the sole power to issue U visas and therefore should have the sole power to decide whether to waive inadmissibility.”

6.3 Addressing Non-Compliance

We conclude by considering the implications of our study if it is in fact true that what happened in Boston was that the District Court’s decision was basically ignored. One possible source of non-compliance follows from standard bureaucratic challenges: IJs and ICE officials face massive caseloads (see discussion in Blasingame et al., 2023). It is also entirely plausible to imagine that they are motivated strongly by blame avoidance. Large caseloads will make it difficult to evaluate claims carefully, and blame avoidance in combination with large caseloads might reasonably create resistance to change. If this is the source, hiring additional IJs would not only address the well-known immigration court case

backlog³² but also promote better decision-making. Hence, a larger cohort of IJs could result in easier compliance with a decision such as *Brito*. Key to this implication, of course, is the assumption that the problem to solve is freeing up IJs who are otherwise motivated to properly apply the law.

A related idea is that the error-correction process in a legal hierarchy also might be expected to address non-compliance, though we are skeptical about this solution. Hausman (2016) persuasively shows that the BIA is not effective at correcting IJ errors that are adverse to the interests of the detained. The third tier of Attorney General review does not seem to correct this bias. A simple explanation is that the members of the quasi-judicial hierarchy in immigration law simply do not share the federal courts' views of how cases should come out, and so the benefits of tiered review among a team are not realized. There are, if you will, at least two distinct teams working in this area of the law. If internal bureaucratic pressures on IJs and members of the BIA within the Department of Justice is the actual problem to solve, relying on tiered review, while useful for solving many other problems, is unlikely to be helpful here.

Hausman points to the lack of legal representation for the detained as a likely explanation for the poor error-correction he identifies, so we might conclude that error-correction could be improved by granting the detained the right to a government-provided lawyer. This surely must be considered a reasonable possibility, and it would likely address many other challenges in the immigration system. Yet, it too runs into conceptual challenges. If the problem to solve is that both IJs and members of the BIA are subject to DOJ pressure, it is not clear why having a clever lawyer who can help you appeal an unreasonable IJ decision will result in a better outcome before a BIA that is similarly, perhaps even more greatly, pressured.

Two final possibilities involve creating more robust oversight procedures for U.S. immigration courts or directly attempting to cut the link between the Attorney General and IJs. Greater involvement of federal courts in the review of bond decisions offers a natural solution. The INA limits judicial review of Attorney General decisions regarding detention, writing that: "The Attorney General's discretionary judgment regarding the application of this section shall not be subject to review. No court may set aside any action or decision by the Attorney General under this section regarding the detention or

³²See "Immigration Court Backlog Tops 3 Million; Each Judge Assigned 4,500 cases." TRAC Immigration. December 18, 2023. <https://trac.syr.edu/reports/734/>

release of any alien or the grant, revocation, or denial of bond or parole” (8 U.S.C. 1226(e)). Creating opportunities for federal courts to correct IJ errors might better break the link between Attorney General pressure and IJ decision-making. That said, it would come at a significant cost to the federal courts, by potentially massively increasing their own caseload.

An alternative to increased federal court involvement would be to restructure the immigration judiciary to grant IJs independence akin to that of administrative law judges (ALJs). ALJs, unlike IJs, enjoy a degree of independence from executive influence, which is crucial for fair and unbiased legal adjudication. By affording similar independence to IJs, the system could mitigate the current conflicts of interest that arise from their status as AG delegates. Such a reform might better align with the democratic ideal of a judiciary that exercises discretion based on legal standards and principles, rather than on executive policies or political agendas.

References

- Abadie, Alberto, Alexis Diamond and Jens Hainmueller. 2010. “Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program.” *Journal of the American statistical Association* 105(490):493–505.
- Abadie, Alberto, Alexis Diamond and Jens Hainmueller. 2015. “Comparative politics and the synthetic control method.” *American Journal of Political Science* 59(2):495–510.
- Abadie, Alberto, Alexis Diamond and Jens Hainmueller. 2020. “Synth: Stata module to implement synthetic control methods for comparative case studies.”.
- Abadie, Alberto and Javier Gardeazabal. 2003. “The economic costs of conflict: A case study of the Basque Country.” *American economic review* 93(1):113–132.
- Bailey, Michael A and Forrest Maltzman. 2008. “Does legal doctrine matter? Unpacking law and policy preferences on the US Supreme Court.” *American Political Science Review* pp. 369–384.
- Baldez, Lisa, Lee Epstein and Andrew D Martin. 2006. “Does the US Constitution Need an Equal Rights Amendment?” *The Journal of Legal Studies* 35(1):243–283.
- Bartels, Brandon L. 2009. “The constraining capacity of legal doctrine on the US Supreme Court.” *American Political Science Review* pp. 474–495.
- Baum, Lawrence. 2009. *Judges and their audiences: A perspective on judicial behavior*. Princeton University Press.
- Blasingame, Elise N, Christina L Boyd, Roberto F Carlos and Joseph T Ornstein. 2023. “How the Trump Administration’s Quota Policy Transformed Immigration Judging.” *American Political Science Review* pp. 1–16.
- Brady, Colin. 2022. “Freedom in the Balance: Procedural Due Process Rights and the Burden of Proof in Detention Hearings in Immigration Removal Proceedings.” *Wm. & Mary Bill Rts. J.* 31:1241.
- Cameron, Charles M and Lewis A Kornhauser. 2024. “What Do Judges Want? How to Model Judicial Preferences.” *Asian Journal of Law and Economics* 15(2):167–208.

- Cheng, Edward K. 2012. "Reconceptualizing the burden of proof." *Yale LJ* 122:1254.
- Clark, Tom S. 2009. "The separation of powers, court curbing, and judicial legitimacy." *American Journal of Political Science* 53(4):971–989.
- Coglianesi, Cary and Daniel E Walters. 2024. "The Great Unsettling: Administrative Governance After Loper Bright."
- Dahl, Robert A. 1957. "Decision-making in a democracy: The Supreme Court as a national policy-maker." *Journal of Public Law* 6:279.
- DeMattee, Anthony J, Matthew J Lindsay and Hallie Ludsin. 2023. "An Unreasonable Presumption: The National Security/Foreign Affairs Nexus in Immigration Law." *Brook. L. Rev.* 88(3):747–796.
- Desposato, Scott W, Matthew C Ingram and Osmar P Lannes Jr. 2015. "Power, composition, and decision making: the behavioral consequences of institutional reform on Brazil's Supremo Tribunal Federal." *The journal of law, economics, and organization* 31(3):534–567.
- Durham, Dory Mitros. 2005. "The once and future judge: The rise and fall (and rise) of independence in US immigration courts." *Notre Dame L. Rev.* 81:655.
- Eagly, Ingrid V. 2014. "Remote adjudication in immigration." *Nw. UL REv.* 109:933.
- Eagly, Ingrid V and Steven Shafer. 2015. "A national study of access to counsel in immigration court." *U. Pa. L. Rev.* 164:1.
- Ferejohn, John and Charles Shipan. 1990. "Congressional Influence on Bureaucracy." *Journal of Law, Economics, & Organization* 6:1–20.
- Finley, Janene R and Allan Karnes. 2008. "An Empirical Study of the Change in the Burden of Proof in the United States Tax Court." *Pitt. Tax Rev.* 6:61.
- Friedman, Barry. 2006. "Taking law seriously." *Perspectives on Politics* pp. 261–276.
- Gasaway, Robert R and Ashley C Parrish. 2016. "Administrative Law in Flux: An Opportunity for Constitutional Reassessment." *Geo. Mason L. Rev.* 24:361.

- George, Tracey E and Lee Epstein. 1992. "On the nature of Supreme Court decision making." *The American Political Science Review* pp. 323–337.
- Gillman, Howard. 2001. "What's law got to do with it? Judicial behavioralists test the "legal model" of judicial decision making." *Law & Social Inquiry* 26(2):465–504.
- Hausman, David. 2016. "The failure of immigration appeals." *University of Pennsylvania Law Review* pp. 1177–1238.
- Hay, Bruce L and Kathryn E Spier. 1997. "Burdens of proof in civil litigation: An economic perspective." *The Journal of Legal Studies* 26(2):413–431.
- Hitt, Matthew P. 2016. "Measuring precedent in a judicial hierarchy." *Law & Society Review* 50(1):57–81.
- Holper, Mary. 2016. "The Beast of Burden in Immigration Bond Hearings." *Case W. Res. L. Rev.* 67:75.
- Hood, Christopher. 2007. "What happens when transparency meets blame-avoidance?" *Public Management Review* 9(2):191–210.
- Humphries, Martha Anne and Donald R Songer. 1999. "Law and politics in judicial oversight of federal administrative agencies." *The Journal of Politics* 61(1):207–220.
- Johnson, Gbemende. 2014. "Judicial deference and executive control over administrative agencies." *State Politics & Policy Quarterly* 14(2):142–164.
- Johnson, Gbemende. 2015. "Executive power and judicial deference: judicial decision making on executive power challenges in the American states." *Political Research Quarterly* 68(1):128–141.
- Kim, Catherine Y and Amy Semet. 2019. "Presidential Ideology and Immigrant Detention." *Duke LJ* 69:1855.
- Knight, Jack and Lee Epstein. 1996. "The norm of stare decisis." *American Journal of Political Science* pp. 1018–1035.
- Kornhauser, Lewis A. 1992. "Modeling collegial courts. II. Legal doctrine." *JL Econ. & Org.* 8:441.

- Lauderdale, Benjamin E and Tom S Clark. 2012. “The Supreme Court’s many median justices.” *American Political Science Review* pp. 847–866.
- Lax, Jeffrey R. 2011. “The new judicial politics of legal doctrine.” *Annual Review of Political Science* 14(1):131–157.
- National Association of Immigration Judges. 2019. “The Immigration Court In Crisis and in Need of Reform”. Technical report National Association of Immigration Judges.
URL: https://www.naij-usa.org/images/uploads/publications/Immigration_Court_in_Crisis_and_in_Need_of_Reform.pdf
- National Association of Immigration Judges. 2020. “Statement of Judge A. Ashley Tabaddor, President National Association of Immigration Judges January 29, 2020, before the United States House of Representatives Committee on the Judiciary Subcommittee on Immigration and Citizenship Hearing on ‘The State of Judicial Independence and Due Process in U.S. Immigration Courts.’”. Technical report National Association of Immigration Judges.
URL: <https://www.naij-usa.org/images/uploads/publications/2020.01.29.00.pdf>
- National Association of Immigration Judges. 2021. “Overview of EOIR Discipline and Performance Protocols”. Technical report National Association of Immigration Judges.
URL: https://www.naij-usa.org/images/uploads/publications/Overview_of_EOIR_Discipline_and_Performance_Protocols_3-13-21_Combined_files.pdf
- Office of the Chief Immigration Judge. 2017. Immigration Court Practice Manual. Official government report U.S. Department of Justice, Executive Office for Immigration Review Falls Church, VA: . Updated November 2nd, 2017, <http://www.justice.gov/eoir>.
- Office of the Chief Immigration Judge. 2020a. Immigration Court Practice Manual. Official government report U.S. Department of Justice, Executive Office for Immigration Review Falls Church, VA: . Updated July 2nd, 2020 <http://www.justice.gov/eoir>.
- Office of the Chief Immigration Judge. 2020b. Immigration Court Practice Manual. Official government report U.S. Department of Justice, Executive Office for Immigration Review Falls Church, VA: . Updated December 31st, 2020; <http://www.justice.gov/eoir>.

- Ríos-Figueroa, Julio. 2007. "Fragmentation of Power and the Emergence of an Effective Judiciary in Mexico, 1994–2002." *Latin American Politics and Society* 49(1):31–57.
- Rogers, James R. 2001. "Information and Judicial Review: A Signaling Game of Legislative–Judicial Interaction." *American Journal of Political Science* 45(1):84–99.
- Ryo, Emily. 2016. "Detained: A study of immigration bond hearings." *Law & Society Review* 50(1):117–153.
- Ryo, Emily and Ian Peacock. 2021. "Represented but unequal: The contingent effect of legal representation in removal proceedings." *Law & Society Review* 55(4):634–656.
- Segal, Jeffrey A and Harold J Spaeth. 2002. *The Supreme Court and the attitudinal model revisited*. Cambridge University Press.
- Sherwyn, David and Michael Heise. 2010. "The Gross Beast of Burden of Proof: Experimental Evidence on How the Burden of Proof Influences Employment Discrimination Case Outcomes." *Ariz. St. LJ* 42:901.
- Spriggs, James F. 1996. "The Supreme Court and Federal Administrative Agencies: A Resource-Based Theory and Analysis of Judicial Impact." *American Journal of Political Science* 40(4):1122–1151.
- Spriggs, James F. 1997. "Explaining Federal Bureaucratic Compliance with Supreme Court Opinions." *Political Research Quarterly* 50(3):567.
- Staton, Jeffrey K. and Georg Vanberg. 2008. "The Value of Vagueness: Delegation, Defiance, and Judicial Opinions." *American Journal of Political Science* 52(3):504–519.
- Staton, Jeffrey K., Varun Gauri and Jorge Vargas Cullell. 2015. "The Costa Rican Supreme Court's Compliance Monitoring System." *Journal of Politics* 77(3).
- Stobb, Maureen, Banks Miller and Joshua Kennedy. 2023. "Who Controls the Immigration Bureaucracy? The Relative Influence of the Three Branches Over Asylum Policy Implementation." *American Politics Research* 51(2):235–246.
- Wexler, Steve. 1999. "Burden of Proof, Writ Large." *U. Brit. Colum. L. Rev.* 33:75.