ORIGINAL ARTICLE

Check for updates

The effect of judges' gender on decisions regarding intimate-partner violence

Joan Josep Vallbé¹ | Carmen Ramírez-Folch²

¹Department of Political Science, Faculty of Law, University of Barcelona, Barcelona, Spain ²European University Institute, San Domenico di Fiesole, Italy

Correspondence

Joan Josep Vallbé, Department of Political Science, Faculty of Law, University of Barcelona, Diagonal 684, 08034 Barcelona, Spain. Email: vallbe@ub.edu

Funding information "la Caixa" Banking Foundation, Grant/Award Number: LCF/PR/SR19/52540006

Abstract

This article aims at disentangling the effect of judges' gender, experience, and caseload in the assignment of restraining orders in IPV cases. Previous literature has independently looked at the effect of gender on judicial decisions and found that it becomes relevant in gender-related cases. However, we find that such effects are better understood in interaction with other contextual factors such as the experience of judges and the amount of work they face, because these determine the levels of uncertainty and information costs surrounding decisions. For our empirical analysis, we use data from on-duty pretrial court decisions on restraining orders in Spain between 2010 and 2018. We find conditional effects of gender depending on experience and workload: more experienced female judges are more likely to grant protection orders than their male counterparts when the amount of caseload is high. These findings are relevant to understand the mechanisms behind judicial inequality under civil law systems, where judges' attributes tend to be unobservable by institutional design.

INTRODUCTION

Does the sex of judges matter when deciding on Intimate-Partner Violence cases? Previous studies have found that judges' sex is relevant in cases where women's rights are at stake (Boyd, 2016; Segal, 2000). While most studies have looked at cases where female judges are still a minority in the judiciary, in this

J Empir Leg Stud. 2023;20:641-668.

⁻⁻⁻⁻⁻⁻

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.

^{© 2023} The Authors. *Journal of Empirical Legal Studies* published by Cornell Law School and Wiley Periodicals LLC.

JOURNAL OF

article, we take advantage of the high gender parity levels on the lower courts in Spain to explore how gender effects arise in IPV cases when female judges are no longer a minority. Additionally, we propose that gender effects are better understood in relation to other institutional and personal constraints such as judges' workload and their seniority. Both of these can determine how costly it is to take a good decision on a case.

Ever since the first empirical analyses of judicial politics, research in the determinants of the behavior of judges (e.g., gender, race, career interest, or ideology) has been characterized by three relevant properties. First, studies have been overwhelmingly focused on the behavior of U.S. judges, with just a few exceptions from other common-law countries (Hanretty, 2013) and just a handful of studies from civil-law countries such as Germany (Schneider, 2005) or Spain (Garoupa et al., 2013). In the common-law system, such as the U.S., judges are more likely to be appointed by a political administration. Concerning gender effects, some of the female justices were added to the bench to advance gender equality rights, thus results reporting gender effects might just be driven by the political administration's efforts on pursuing women's rights.

The second characteristic relates to the political and salient nature of the selection process of higher courts, most research has focused on the behavior of judges from Supreme, Constitutional, or Appellate courts, thus leaving the behavior of a massively larger share of lower, pretrial, or trial courts unexplored (Guthrie et al., 2007). On the one hand, in the majority of cases, the selection of lower court judges is disconnected from the political process, we can explore judicial behavior in a setting where judges have the incentive to reveal sincere behavior. On the other hand, unlike appellate courts that deal mostly with contentious cases (decisions by lower trial courts that have been appealed), lower-court routine cases are less likely to trigger identity-based or politically-based judicial behavior.

Third, most of the research exploring the effect of certain factors on judicial decisions has intended to untangle the causal effect of the variables of interest through only weak regard for the conditions required to make such causal claims (Boyd, 2016). Importantly, random treatment assignment is difficult to comply with when exploring the behavior of higher court judges, because they may have certain control over their dockets, and are usually appointed through a political process. In this context, case allocation to judges may well be endogenous to judges' own characteristics, thus raising many problems of identification. This article exploits a naturally occurring random treatment assignment, that is, a random assignment of cases to judges within districts, thus allowing us to estimate the causal effects of gender and other factors on judicial decisions. We are able to sort these three limitations in the literature by carrying out our analysis in lower courts of the Spanish judiciary, an example of a civil law system, where judges are promoted based on merit. This allows us to minimize potential endogeneity problems derived from politicized appointment mechanisms commonly found in higher courts.

In addition, our theory refines the gender-effect argument. We find that both experience and individual workload are relevant because, according to our theoretical model, they impact the levels of uncertainty and information costs surrounding judicial decisions. An essential element of theories dealing with judicial behavior lies in the discretionary margin that all judges have when making decisions (Epstein et al., 2013). This discretion to rule may become a burden when judges do not have the experience to make decisions in complex cases, when their decisions are strongly bound by tight procedural deadlines, and when they have high volumes of accumulated cases to be solved. In such situations, judges are more likely to make use of personal cues or cognitive heuristics that ease decision making. Our argument is that when dealing with intimate-partner violence (IPV) cases, these heuristics are driven by judges' gender identity, which for instance would make female judges prone to empathize with victims. Thus, we contribute to the literature by revealing the conditions needed for gender to be a determinant of judicial decisions.

To test our expectations, we use data from on-duty pretrial court decisions on restraining order petitions in Spain between 2010 and 2018. Our results indicate that gender explains only a small portion of the observed variation in judicial decisions on IPV cases. We find that gender effects are better understood as conditional, dependent on levels of experience and workload. For instance, experienced female judges are more likely to grant protection orders than their male counterparts, but workload makes protecting IPV victims much less likely for both male and female judges. Therefore, when dealing with IPV cases, the gender of judges is critical for both male and female judges when it interacts with the levels of uncertainty and information costs surrounding decisions.

The next section is devoted to briefly present the ways in which our paper dialogues with the judicial behavior literature, as well as explaining how IPV courts are managed by Spanish courts. Section 3 presents our model of judicial decisions on IPV cases, in which we outline our main argument and expectations. Section 4 presents the data and explains the identification strategy, while the main empirical results are presented in Section 5.

LITERATURE AND BACKGROUND

How judges judge

Like doctors, firefighters, or pilots, judges are expert, professional decisionmakers whose decisions depend on certain incentives and constraints (Epstein et al., 2013). Incentives include genuine love for the work, job tenure, promotion, social prestige, ideology, or the advancement of group-based interests shaped by ethnicity, class, or gender. Ideology (or nominating party) has been found relevant to predict judicial decisions when cases deal with either controversial or value-laden cases, as is typically the case of supreme courts in common-law systems (Segal & Cover, 1989; Weinshall et al., 2017) and constitutional courts of civil-law systems (Garoupa et al., 2013; Hanretty, 2012), although some examples of ideological voting have been found in lower courts of common-law (Geyh, 2016; Keith et al., 2013; Sunstein et al., 2007). In a similar way, identity-based factors are especially relevant when dealing with cases where these factors are salient, when judges tend to decide in favor of plaintiffs who match judges' own identity (Boyd, 2016; Boyd et al., 2010; Segal, 2000).

On the constraints side of the equation, judicial decisions, as any other kind of decisions, are constrained by the uncertainty surrounding the problem at hand—shaped only in part by the legal facts of the case (Dyevre, 2010)—and by the information costs of making good decisions (Kahneman et al., 1982; Newell, 1990; Simon, 1955, 1957). As most experts, judges deal only imperfectly with probabilities and with the "evidence-belief-decision" path (Sonnemans & van Dijk, 2012), but they become better professional decision-makers by identifying patterns of events from past decisions that can be used as templates to make better and more efficient decisions in the future (Gigerenzer and ABC-Research-Group, 2008; Richards, 2016; Vallbé, 2015). Therefore, as they learn from experience judges need less time to gather information. Still, quality decisions do require time, and judges must make a lot of quick decisions, typically attending several cases simultaneously, each of which requiring quick and parallel decisions, and thus raising well-known problems of attention allocation (March & Simon, 1958) that are aggravated by the "wicked" nature of the environment provided by judicial systems (Guthrie et al., 2007). Especially in cases where fundamental rights are at stake (e.g., victims' security and defendants' freedom of movement) the law in most democratic systems compels courts to expedite procedures to prevent unnecessary harm to plaintiffs' rights. In that sense, overloaded court dockets reduce the time to make decisions (Galanter, 2004), and judges tend to resort to more intuitive solutions (Guthrie et al., 2007), which usually take the form of cognitive shortcuts or heuristics that help them make quicker decisions by reducing information costs (Bainbridge & Gulati, 2002; Gigerenzer & Engel, 2006; Vallbé et al., 2019). When these decisions must be made during pretrial conferences, with scarce or no evidence at all, judges resort even more to intuition (Guthrie et al., 2007). Yet, the need for quicker decisions also makes judges less creative (Epstein et al., 2013) and more prone to error (Jolls & Sunstein, 2006; Kahneman, 2012), which on average produces worse judicial outcomes for plaintiffs (Engel & Weinshall, 2020).

Gender and judging

Studies seeking to identify gender effects on judicial behavior have found only partial support in the data. In general, the idea that male and female judges "speak different voices" and that these are reflected in different kinds of judicial behavior (Boyd et al., 2010) has hardly passed empirical scrutiny, except in cases

JOURNAL OF

EMPIRICAL LEGAL STUDIES

where the rights of women are clearly at stake, which female judges tend to decide distinctly from male judges (Boyd, 2016; Segal, 2000). In cases where the "women's issues" label is not clearly applicable, results are less conclusive.

Since the seminal article by Boyd et al. (2010) there is agreement that the reasons behind these unclear results are largely methodological. On the one hand, most research on the estimation of the causal effect of gender on judging is doomed by its inability to ensure random assignment of treatment, given that case, allocation among upper trial or appellate judges is almost never random. In one of the few exceptions, Knepper (2018) used random assignment of federal district court to civil cases to test whether gender plays a significant role in settlement cases (as opposed as trials, the most common types of cases exploited by the literature). To deal with the random assignment of treatment problem, Boyd et al. (2010) applied nonparametric matching techniques to ensure a proper comparison of treatment-outcome pairs and therefore to allow for unbiased causal effects to be estimated. On the other hand, the extent to which gender effects on judging are only to be found when cases involve "women's issues" faces another potential problem in the way "women's issues" are defined and measured. There are two common strategies on this (Boyd et al., 2010). First, authors may decide to test gender effects on cases that hardly have any genderrelated connotation (e.g., tax law; Schneider, 2001) and include a dummy variable for the gender of the applicant. The mechanism behind this design seems to be that female judges may decide more for female applicants than male judges. The second strategy is very similar to the former, except for the fact that the cases selected do have gender connotations (e.g., workplace discrimination; Boyd, 2016). This second strategy has received more empirical support than the former (Boyd et al., 2010).

However, there is also a theoretical problem affecting works on gender effects. As a large corpus of works that examine the behavior of judges has focused on high courts (e.g., the U.S. Supreme Court), they depict scenarios where gender parity is rare (Crowe, 1999; Farhang & Wawro, 2004). This has often caused literature to treat women as "tokens"-a minority in a field of study and professional practice dominated by men (Martin & Pyle, 2005). As a consequence, a theoretical assumption underlying this literature is that gender will produce a distinct behavior just on female judges, especially on genderrelated cases-cases that are potentially relevant for female judges who are aware of so-called "gender issues". This bias thus frames the behavior of male judges as "default" while the behavior of female judges is presented as a deviation from normality. Assuming that women are the only "group" on whom gender has an effect is the result of perceiving female judges as the minority group who will make decisions to protect their group's interests or even as an unconscious bias, which assumes that male judges are somehow gender-neutral in their decisions-that is, that they will not tend to identify with male defendants. This issue, which has been pointed out for studies exploring race effects

(Gazal-Ayal & Sulitzeanu-Kenan, 2010), has been largely overlooked referring to gender, with the exception of Ash et al. (2021). We understand gender as an identity mechanism and as such it can produce in-group bias—judges may favor those defendants or victims who match their identities. The existence of this mechanism and its consequences on inequality have been empirically tested on numerous fields such as career choice (Correll, 2001) or labor-market discrimination (Rodríguez-Menés & Rovira, 2019; Rudman et al., 2012). Very recently Ash et al. (2021) have tested it on outcomes of criminal courts in common-law India, where the presence of female judges is very limited. Yet, its effects in lower courts with a more balanced gender distribution and under a civil-law system is unexplored.

This study contributes to the literature focusing on lower pretrial courts in Spain, where sex composition has been balanced for the last two decades. In addition, it does so by looking at the behavior of judges under a typical civil-law system, whose strategic behavior has been largely overlooked. The institutional framework of civil-law legal systems tends to limit political influence to the top positions of the judiciary, which projects an aura of professionalism and impartiality on lower court judges: "judges should be the mouth of the law and passively apply it as produced by the legislature" (Guarnieri, 2003, 2010; Hilbink, 2007).

Background: Decisions on gender violence in Spain

The Spanish judicial system consists of three distinct levels: 2600 pretrial courts (courts of first instance and of inquiry), 1000 trial courts, which try cases, 16 regional courts of appeals, and the Supreme Court, which has the last word on appeals. Moreover, the Judicial Council (Consejo General del Poder Judicial, CGPJ) is an agency with powers to appoint, promote and discipline the members of the judiciary. 60% of members of CGPJ are directly appointed by the Spanish parliament (i.e., government), and 40% are elected by judges themselves among peers. In the last 15 years, the CGPJ has been overwhelmingly conservative. This double hierarchical structure induces judges at different levels "to consider how their decisions will be received by other actors within the system" (Zorn & Bowie, 2010), whether these actors are other courts or politicians.

At the bottom of the system, there are lower courts, which include both pretrial and trial courts. In 2004, the Spanish government created courts specialized in intimate-partner violence (IPV) at the lower-court level. These have exclusive competence to process all incoming cases of this type and to carry out pretrial preliminary investigations of these cases. These courts were established only in more populated areas, while in rural areas (the vast majority of judicial districts in Spain) these duties were assigned to one of the ordinary lower courts already in place. Appointments to these courts are made on the same grounds as in regular pretrial lower courts—they are decided by the CGPJ based on merits and seniority. One of the most effective measures to protect victims in pretrial

JOURNAL OF

EMPIRICAL LEGAL STUDIES

procedures of IPV are restraining orders. These orders can be taken when there is strong evidence of the commission of crimes or misdemeanors of gender violence and in cases where there is an objective situation of risk to the victim. When the order is granted by a judge, the court also activates social assistance and protection measures so that protection is effective at all levels. Given the liberty-limitation nature of this measure (aggressors have limits to their liberty to move), its effect is immediate though limited in time (usually 30 days), after which a judge must revise the case and decide again. Its immediate but short application makes appeal rather improbable and ineffective.

Restraining orders may be taken by lower-court judges in three different scenarios. First, in areas where specialized IPV courts have been established, restraining orders will be taken by their specialized judges. Second, in areas where specialized courts are not in place, judges from the courts that have been assigned the exclusive competence in IPV cases will make these decisions. Finally, in both previous scenarios, if an IPV case is brought to court out of ordinary court hours, a judge on duty will process the incoming case and can therefore decide on restraining orders. Given the typical urgency surrounding cases coming to court off regular hours, it is expected that restraining orders adopted by on-duty courts refer to critical cases. This explains why in Spain on-duty courts consistently approve restraining orders at higher rates than ordinary courts, as shown in Figure 1.



FIGURE 1 Rate of approval of restraining orders from Spanish ordinary and on-duty courts, 2010–2018.

JOURNAL OF EMPIRICAL LEGAL STUDIES

Thus, when deciding on IPV cases, the incentives of Spanish lower court judges are shaped by two different lines. On the one hand, the law in Spain defines intimate-partner violence (also called "gender-based violence") as "violence directed against women for the mere fact of being women", which makes courts distinguish IPV from general domestic violence cases. Therefore, in IPV cases the stakes are well defined on gender lines-applicants are always women who plead their right to security and integrity, and defendants are always men who plead their right to move freely-and thus gender might work as an identity-based mechanism that influence judges' decisions. On the other hand, the fact that a restraining order is granted by court or not has an impact beyond the immediate interests of both the plaintiff and the defendant. When a court grants a restraining order it assumes (per art. 544 of the Spain's Criminal Code) the actual existence of an IPV victim and her objective situation of risk, and thus the restraining order must be understood as just the first step of an ensuing criminal procedure. In this context, judicial decisions are constrained by the uncertainty and information costs provided by the environment within which their decisions are made, which is the result of the interaction between the information they have, their experience as decision-makers, and the workload they are facing.

MODELING JUDICIAL DECISIONS OVER IPV CASES

Theoretical model

The research question behind this article is how and in what way judges' sex affects lower-court judges when making decisions on IPV cases, and to what extent uncertainty and information costs surrounding such decisions moderate that effect. In this section, we present our model of judge's decisions regarding IPV cases, incorporating the factors discussed above. To model the relationship between deciding on IPV cases and these factors, we build on the judicial utility model provided by Epstein et al. (2013) and Guthrie et al. (2007), the cognitive framework for decision making of Simon (1997) and Gigerenzer and Engel (2006), and the judicial nominations model developed by Cameron et al. (2013).

The decision of a judge to grant a restraining order will depend on a number of factors associated with both expected benefits and costs. We can conceptualize costs distinguishing between the uncertainty function (E)—uncertainty surrounding the definition of the situation posed by a particular case—, and the information costs (T) associated with the time needed to search, measure, and weigh the attributes of the situation. We include both in the cost component of a standard utility function:

JOURNAL OF EMPIRICAL LEGAL STUDIE

$$U_i = B - (E + T)$$

where *B* is the expected benefit of taking the right decision and (E + T) the cost factors. According to the model, judge *i* will decide to grant a restraining order for IPV victim *v* if *i*'s utility in protecting *v* (u_{pv}) is greater than not granting the restraining order U_{npv} , that is whenever the value of the expected benefits is greater than the sum of costs involved: B > (E + T).

With this in mind, we conceptualize a multiplicative factor δ that captures the effect of certain factors on the whole cost function. Conceptually, δ measures how *difficult* is for the judge to accurately assess the components of the elements of the cost function. For instance, experience makes judges gradually more able to connect new situations to experienced patterns of events, which will decrease the amount of uncertainty of the situation and the time needed to assess certain facts of the case. Another example is judges' gender identity, which according to our model works as a mechanism that can make judges orient their decision in favor or against certain parties in the case. In a sense, δ allows us to formalize why different judges may decide differently on identical cases. To δ to have a multiplicative effect on the costs it is formalized as a factor of discount:

$$U_i = B - \left(\frac{1}{1-\delta}\right) (E+T)$$

Note that conceptually the value of δ cannot be negative, and empirically it should be standardized to lie between 0 and 1. Let us illustrate the discount effect of δ in two opposite scenarios. Let us imagine two judges facing the same IPV case, one with 20 years on the bench and the other an inexperienced judge. A victim's report and the facts of the case provided by the police are presented in a way that establishes the costs of deciding on the situation. For instance, there are no marks of physical violence on the victim's body, the defendant denies the charges, and the police provides evidence of previous reports against the defendant for drug abuse. In a very simplified way, these elements provide a certain level of uncertainty surrounding the problem and certain costs of gathering further information on the case (e.g., whether the defendant has a police record for battery or abuse). In our model, a judge with 20 years on the bench and an inexperienced judge will confront these costs at different levels of difficulty. As an expert, the senior judge will contrast this situation with past events and rapidly assess the level of uncertainty surrounding the problem, and new information coming in regarding the defendant will be weighted in a rather efficient way. This implies that, regardless of the objective costs of the decision situation, δ for the senior judge will be very small, very close to zero, so she will incur no extra costs for making the decision than (E+T). On the contrary, the first-time judge will have a hard time deciding whether evidence of physical

violence is a necessary condition for granting a restraining order. In the extreme case of a judge facing her first IPV case, the value of δ would be close to 1, and then the judge would incur in infinitely high extra costs for making a decision. If $\delta = 1$, then $U_i = B - (\frac{1}{0}E + T)$, and therefore, $U_i = -\infty$.

When would δ be too high for a judge to grant a restraining order? Following up the equation, we can obtain a critical point for δ (δ^*) where the difficulty of assessing the costs will make the costs higher than the expected benefits. According to this, for judge *i* to grant the restraining order, the utility function must satisfy $U_i > 0$, which can also be expressed like the following equation:

$$(U_i =) B - \left(\frac{1}{1-\delta}(E+T)\right) > 0$$

Using some algebra, we can solve the inequality for δ obtaining its critical value (δ^*):

$$\delta^* \leq 1 - \left(\frac{(E+T)}{B}\right)$$

which means that the critical δ (δ^*) must be smaller than one minus the ratio between costs (E + T) and expected benefits (B). This is logical since the equation implies that if the costs of making a particular decision (E + T) were higher than the expected benefits *regardless* of the value of δ , the value of δ^* would be negative, which is impossible. Therefore, according to the model judges with a high δ will only grant restraining orders if regular costs (E + T) are tiny and smaller than benefits.

Our first expectation mimics what has been found about gender effects on judging in other jurisdictions (Boyd, 2016; Segal, 2000). In particular, given that in IPV cases victims are always women (and aggressors are always men), we assume that all else equal average δ will be smaller for female judges due to the sex identity mechanism at work because female judges should empathize more with female victims (who match their group identity) than male judges, who might show higher levels of empathy toward male defendants (i.e., aggressors). Therefore, we should observe (H1) that female judges should grant restraining orders for IPV victims at higher rates than male judges, all else being equal.

For career judges in civil-law systems, lower courts are the starting point in their careers and the scenario in which they acquire professional expertise. As they become experts they learn from past decisions identifying patterns of events from similar cases (March & Simon, 1958; Simon, 1997), thus decreasing average uncertainty with experience. Past evidence shows that at early stages of their career (i.e., when uncertainty is high), judges feel less secure about being creative with their decisions, especially when they are on duty (Vallbé, 2015), and thus their decisions tend to be more bound by rules and grounded on legalistic terms

IOURNAL OF

MPIRICAL LEGAL STUDIES

(Epstein et al., 2013), because they are less prone to the discretionary use of their attitudes and values to make decisions (Martin & Pyle, 2005). This, in turn, has been found to translate into worse judicial outcomes for plaintiffs (Engel & Weinshall, 2020). This mechanism can be explained by status quo bias, which under certain conditions makes decisionmakers opt for status quo choices at higher rates than expected by rational choice models (Samuelson & Zeckhauser, 1988). During the recent COVID crisis administrative courts have offered many examples of such bias when reviewing governmental (in)action (Cafaggi & Iamiceli, 2021). In the case of IPV cases in Spain, the law establishes that restraining orders are extraordinary measures only to be taken when judges perceive an objective situation of risk for victims. According to our model, uncertainty makes judges less sure about the merits of the case, which should trigger status quo bias. If we assume that both the rate of real and false accusations, and the amount of noise surrounding the facts of cases are randomly distributed across judges, we should expect that the rate of approval of restraining orders should be lower among less experienced judges (because of larger δ) and grow with professional experience. Therefore, more experienced judges should grant restraining orders for IPV victims at higher rates than less experienced judges, all else equal (H2).

In addition, courts are usually packed with cases and short in staff, and judges find themselves in the need of attending multiple cases at a time (Guthrie et al., 2007), requiring quick and parallel decisions, and constantly facing bottlenecks of attention (March & Simon, 1958; Simon, 1997). In appellate courts, judges facing growing dockets have been found to decide cases in less timeconsuming ways that can affect opinion writing (Hanretty, 2020) and even the actual legal outcome of the case (Huang, 2011; Lavie, 2016). In lower courts that are in charge of investigating cases (such as IPV courts), democratic legal procedures typically compel them to minimize delays when deciding on preliminary measures that affect citizens' fundamental rights (as is the case of restraining orders in IPV cases), which adds in even more time pressure. In this context, judges facing large workloads might clear off their dockets dismissing restraining orders at higher rates, which minimizes the costs of starting new criminal procedures, which should be done for each granted restraining order. These decisions need not be conscious, but they are triggered by cognitive shortcuts and biases that allow them to make quicker decisions (Gigerenzer & Engel, 2006), minimize information costs (Vallbé et al., 2019), and blur the connection between the case and later potential events in the case (Guthrie et al., 2007). Again, quicker judicial decisions tend to come at a high price for plaintiffs, who face worse outcomes when judges are less creative (Engel & Weinshall, 2020). In this sense, while experience helps judges to become better decision-makers, individual workload counteracts this effect by increasing the value of δ , seriously harming the quality of their decisions giving them incentives to cut off cases. Therefore, judges facing higher workloads should grant restraining orders for IPV victims at lower rates than judges in courts with lower caseload, all else equal (H3).

1740146, 2023.3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jels.12361 by Radkube (Labiva Inc.), Wiley Online Library on (2904-2025). See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for ulses of use; OA anicels are governed by the applicable Centure Commons License

However, because judges' gender has a different impact on the value of their δ , we should observe that male and female judges react differently to the effects of experience and caseload. On the one hand, experience should work positively on both: if experience reduces their uncertainty, all else equal both male and female should grant more restraining orders. But, if female judges on average decide more in favor of female victims and uncertainty and information costs produce unequal outcomes for plaintiffs (in this case, worse outcomes for IPV victims), female and male judges might react differently to the factors that affect their identity-based preferences. For instance, more experienced judges have seen more cases on gender violence, and that might have made them more or less conscious about their gender role in the legal procedure. For that reason, female judges should react more positively to the effects of experience (which favor victims) while male judges should react more negatively to the effects of workload (which do not favor victims; H4).

Empirical model

The empirical approach to this model requires considering the structure of our data. Our data set contains decisions over all restraining orders from IPV cases decided by Spanish on-duty pretrial courts between 2010 and 2018 (next section will provide more details on that). Judicial decisions at this court level are hierarchical (or multilevel). In general, judges represent the individual, lowest level in our data, and each decision on an IPV case takes place at this level. Above this level, we have two groups. On the one hand, each judge works in a court district and each district has a court on duty (CoD). Because districts can gather more than one judge at a time, we find that one CoD gathers decisions made by different judges. On the other hand, given the time span of our data, years constitute a second group. However, the identification of the group levels can differ according to our empirical strategy. In a first set of models, we want to measure the average effect of our main covariates across on-duty courts and years, and there we combine CoD and Year into a single one CoD-Year group. In a second set of equations, we want to test the effect of experience and caseload (and their interaction with judges' sex) net of any individual characteristics that judges may present, and in these models, we estimate our coefficients taking into account the Judge-Year level. In all cases, we measure judges' sex, level of experience, and individual caseload at the individual judge level.

According to this multilevel structure, we assume that observations within each CoD-Year (Judge-Year) are not independent. Following the notation by Gelman and Hill (2007), we formalize our model as a multilevel, varying-intercept logistic regression model:

$$Pr(y_i=1) = \text{logit}^{-1}(X_i\beta + \alpha_{j[i]})$$

JOURNAL OF

for i = 1, ..., n, where X is the matrix of individual-level predictors and $j_{[i]}$ indexes the CoD-Year (Judge-Year) where and when judge *i* makes a decision. The second part of the model, which underscores its multilevel nature, is the regression of the CoD-Year (Judge-Year) coefficients:

$$\alpha_j \sim N(U_j \gamma, \sigma_\alpha^2)$$

for j = 1, ..., J CoD-Year (Judge-Year) combinations, where U_i is the matrix of CoD-Year(Judge-Year)-level predictors, γ is the vector of coefficients for the CoD-Year(Judge-Year)-level regression, and σ_{α} is the standard deviation of the unexplained group-level errors.

Providing a compromise between no pooling (fitting models without accounting for the multilevel structure of the data, which would overestimate individual-level effects) and complete pooling (fitting one separate model for each group), multilevel models allow for partial pooling of coefficients and model variation between groups, which provide efficient estimates especially when groups are unequal in size (Gelman & Hill, 2007).

In our model to test H1, we will assume a different intercept for each of the 2214 CoD-Year observations, and a nonvarying effect for sex across all groups. Our second model (H2) will add two individual-level variables (judge workload and experience) within the varying-intercept model. After that, third and fourth models will shift to varying intercepts for Judge-Year to further test the effect of experience and its interaction with sex (H2), and to test the effect of workload and its interaction with sex (H3). Finally, a last model will deal with the moder-ating effect of workload on experience and sex (H4).

DATA SOURCE AND IDENTIFICATION STRATEGY

In order to fit our models, we use data on decisions made by lower pretrial court judges in Spain (2010–2018). The choice of court level (the lowest possible) and country (of civil-law tradition) is a relevant contribution to the existing literature, but it also faces a number of challenges. In what follows, we explain our sources for data, and how this article deals with the existing challenges.

Data set

Because open access to public judicial data is still in its infancy in Spain, our data set has been built by matching data from three separate sources. On the one hand, the official database of the Spanish Judicial Council (Consejo General del Poder Judicial, CGPJ) provides aggregated statistics on the number of

1740146, 2023.3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jels.12361 by Radkube (Labiva Inc.), Wiley Online Library on (2904-2025). See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for ulses of use; OA anicels are governed by the applicable Centure Commons License

JOURNAL OF

restraining order petitions granted and denied each year by each court in Spain—including ordinary and on-duty courts. On the other hand, every 2 years, the CGPJ publishes the list of all members of the Spanish judiciary, which gives us the court in which each judge works and the exact number of years, months, and days passed since the judge entered the judiciary. Judges' sex was deduced from their first names. In order to fill in the information gaps in between years, we scrapped the State Official Register (*Boletin Oficial del Estado*) and identified all promotions and court changes occurred on those years. After joining the three data sets, we expanded the rows of the resulting data set so that each row represented one judge-court-year-decision observation. The outcome variable (decision) takes value 1 if a judge grants a restraining order and 0 if he/she denies it.

Due to the restrictive policy of the Spanish government regarding access to public judicial data, we do not have information about the facts of each IPV case. This is potentially challenging for our analysis because certain facts of cases (e.g., the extent of physical violence used against a victim) can seriously influence judges' decisions. However, the random rotation system to assign cases to judges (more information below) allows that judges are assigned cases regardless of the facts, thus minimizing the risk of selection bias.

In total, we have data consisting of 44,335 restraining order petitions filed across 1063 different courts on duty. Petitions for restraining orders were heard by 1528 unique on-duty judges, of which 63% were female and 37% male. Overall, 72% of restraining order petitions were granted and 28% were denied.

Accounts of judicial politics in common-law systems tend to use some measure of judges' ideology or policy preferences in the analysis of their decisions. The very nature of civil-law legal systems notwithstanding, the Spanish legislation is very restrictive regarding the involvement of judges and prosecutors in politics. Spain's Constitution (article 127) explicitly bans that the judiciary join trade unions or political parties. Nevertheless, there are several so-called judicial unions that coordinate their members' interests in matters ranging from salary reform to nominating candidates to the upper appeal courts. Although judicial unions claim to be independent of political parties, they lean toward the two main Spanish political parties—progressive PSOE or conservative PP. However, because membership in these unions is not public, the measurement of the ideology of lower court judges is all but impossible. Moreover, around 43% of judges are unaffiliated. For this reason, we cannot measure the ideology of lower-court judges.

Table 1 shows basic summary statistics of the main variables once the data set is organized at the judge-year-decision level. Female judges decided on 60% of the petitions (they represent 63% of judges in the data). This distribution of cases supports the idea that the distribution of male and female judges across judicial districts is very close to random. Each on-duty judge heard on average 13 petitions per year, although variation is quite high. In addition, the average

EMPIRICAL LEGAL STUDIE

Statistic	N	Mean	Standard deviation	Min	Max
Restraining order granted	44,335	0.72	0.45	0	1
Female judge	44,335	0.60	0.49	0	1
Caseload	44,335	13.28	8.75	1	54
Experience (years)	38,948	12.23	7.70	0	36

TABLE 1 Summary statistics of the main covariates of the IPV decisions dataset.

level of experience among our judges is around 12 years, with a standard deviation of almost 8 years.

Identification strategy

We exploit the random distribution of case files among on-duty judges within court districts to estimate the effect of judges' sex, uncertainty, and information costs on granting restraining orders in IPV cases. Typically taken as a preliminary action within a criminal procedure against an alleged IPV abuser, a restraining order commands an (alleged or confirmed) abuser to stay away from the victim establishing a minimum distance parameter. In Spain, restraining orders may be taken by ordinary or on-duty lower-court judges. If petitions are filed with court within office hours, they will be taken by judges who have the exclusive competence to take IPV cases. If an IPV case is brought to court during off hours, a judge on duty will process the incoming case and will have to decide on the restraining order petition. Because judges working in the same district are on duty according to a random rotation system, we focus on IPV cases heard by on-duty courts, since the assignment of IPV cases to on-duty judges is as good as random, which we need to estimate causal effects (Angrist & Pischke, 2014).

A major advantage of this strategy is that it helps us addressing potential selection bias. If a judge is known to grant only a small share of restraining orders, lawyers could advice victims to defer their petitions until a more favorable judge can hear it. In addition to the random allocation produced by the rotation system, this scenario is unlikely in our dataset for at least two more reasons. First and foremost, when victims go to the court on duty to file a petition they usually do that as an urgent call—literally because there *has been* an aggression—which makes strategic behavior rather unlikely. Second, in Spain, victims usually do not file their restraining order petitions directly in court premises, but initial reports of aggression take place in police stations, and it is the police that inform victims of their right to get a restraining order from a judge. The petition is actually prepared within police premises and only later brought to court.

655

	Male		Female			
	Mean	Standard deviation	Mean	Standard deviation	Difference in means	р
Caseload	14.22	8.86	12.65	8.62	-1.57	0.00
Experience	14.05	8.66	10.98	6.69	-3.07	0.00

TABLE 2	Balance table of	covariates on	gender treatment
	Dalance table of	covariates on	gender treatment.

MPIRICAL LEGAL STUDIES

IOURNAL OF

In a related project, we had access to a sample of 127 files of restraining order petitions within the Spanish region of Catalonia. From those data, only 4% of victims had filed their petitions directly to court, while 95% had done it through the police. Once in the police station, 48% of victims had requested legal advice to go to court (87% of which through free legal aid). Still, 52% of victims who went to court after going to the police to file their petition, did so without a lawyer. A large number of victims only get a lawyer once the petition has been processed, so that further criminal charges can be brought against the aggressor. Therefore, we assume that sex treatment is as good as randomly assigned to petitions within judicial districts, and that the threat of reverse causality is very small.

Table 2 shows a balance table of our main covariates comparing our treatment groups (petitions heard by female and male judges). The results show that differences in experience and the number of petitions heard by male and female judges, though small, are significant. Due to their later incorporation into the bench, female judges tend to be slightly less experienced than male judges. This would also explain why male judges tend to hear a little more petitions than female judges. Due to the dynamics of the selection and promotion system of the Spanish judiciary, more experienced judges have higher probabilities to serve in courts located in larger cities. Given than female judges are on average in earlier stages of their judicial careers, they tend to be overrepresented in rural judicial districts.

MAIN EMPIRICAL RESULTS

In this section, we will first discuss the unconditional means model—including only varying intercepts and no predictors—to compare variability within and between subjects (Roback & Legler, 2021). After that, we will present the result of the rest of multilevel models.

With the unconditional one-intercept model we obtain an estimation of the average probability of a judge approving a restraining order across all groups (CoD-Year). We transform the intercept in the logit scale (1.344) to the probability scale (0.793), and then use its standard error to get a ± 2 standard error

interval, which is (0.782, 0.804). The point estimate differs slightly from the unpooled average rate of approval we find in the historic data (0.72) because in the multilevel context, α is the weighted average of the unpooled mean and the mean over all CoD-Year categories (Gelman & Hill, 2007). This explains why the intercepts corresponding to CoD-Year observations with fewer observations have larger standard errors than groups with a higher number of observations ones. The same holds for Judge-Year observations.

Focusing now on the results including our main covariates, the first column of Table 3 shows a first model including judges' sex as the sole predictor. According to the results, when restraining orders are requested to female judges, the probability of these being granted is larger compared to male judges. This first model predicts that an average female judge is 1.05 times more likely to grant a restraining order to an IPV victim than a male judge. In terms of probability, female judges will grant a restraining order with a 0.8 probability, while male judges will do so with an average probability of 0.78. It should be taken into account that models cannot include information at case level, which introduces noise into our estimates. In other words, even without controlling for the facts of the case, judges' sex seems to make a difference when resolving IPV cases. Furthermore, the effect of sex increases when in subsequent models we introduce judges' workload and level of experience, as we shall see in the next model.

The second model (second column of Table 3) compares decisions of male and female judges controlling for levels of workload and career experience. Although the original logit coefficients can be seen in the second column of Table 3, we present the results of this model translated into odds ratios in Figure 2 for the sake of readability and clarity. First and foremost, keeping workload and experience constant at average levels, in this new model female judges are 1.08 times more likely to grant restraining orders than male judges. Moreover, the coefficient plot shows that across CoD-Year observations, on average increasing workload works negatively on the odds of granting restraining orders, whereas experience works in the opposite direction.

Regarding the effect of workload, the model predicts that one additional case in judges' dockets will reduce the odds of granting restraining orders by 15%, although as seen in Figure 3, the effect is marginally decreasing—the negative effects are stronger for initial additions to judges' workload and will decrease as workload gets larger.

As expected by our argument, as judges get more experience, they tend to grant restraining orders more frequently, which is consistent with our model of declining uncertainty with experience. In particular, each additional year of experience on average increases the odds by 1.8%. With models having CoD-Year as group level, this positive effect of judges' experience could be caused by a generational trend across the Spanish judiciary. In particular, the positive effect of experience might be caused because younger generations of judges on average decide less in favor of IPV victims.

1740146, 2023.3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jels.12361 by Radkube (Labiva Inc.), Wiley Online Library on (2904-2025). See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for ulses of use; OA anicels are governed by the applicable Centure Commons License

	Dependent variable Decision in favor of vi	ctim			
	(1)	(2)	(3)	(4)	(5)
Female	0.056** (0.027)	$0.081^{***}(0.030)$	$-0.139\ (0.104)$	0.068 (0.134)	-0.436* (0.251)
Caseload		-0.237^{***} (0.027)		-0.258^{***} (0.049)	-0.525^{***} (0.094)
Experience (years)		$0.018^{***} (0.003)$	$0.011^{**}(0.005)$		-0.030^{**} (0.015)
Female * Experience			0.013*(0.008)		0.070*** (0.021)
Female * Caseload				-0.072 (0.063)	0.088 (0.121)
Caseload * Experience					0.022*** (0.006)
Female * Caseload * Experience					-0.021^{**} (0.009)
Constant	$1.309^{***} (0.037)$	$1.581^{***}(0.069)$	$1.359^{***} (0.083)$	1.993^{***} (0.109)	2.327*** (0.199)
Group level	CoD-Year	CoD-Year	Judge-Year	Judge-Year	Judge-Year
Observations	44,335	38,948	38,948	44,335	38,948
Log Likelihood	-24,138.89	-20,995.92	-20,312.90	-23,335.32	-20,248.76
Akaike Inf. Crit.	48,283.78	42,001.83	40,635.80	46,680.64	40,515.51
Bayesian Inf. Crit.	48,309.88	42,044.68	40,678.65	46,724.14	40,592.64

Results of multilevel logistic regression models of judicial pretrial restraining orders in gender violence cases. TABLE 3

Note: *p < 0.1; **p < 0.05; ***p < 0.01.

JOURNAL OF EMPIRICAL

LEGAL

STUDIE

JOURNAL OF EMPIRICAL LEGAL STUDIE





FIGURE 2 Coefficient plot of the model including judges' sex, level of experience, and workload, without interactions.



FIGURE 3 Predicted values of the probability of granting a restraining order for an IPV victim at different levels of individual workload.

Our next models take up the Judge-Year level as the group to sort out this potential problem and to capture all the effects net of any particular characteristics judges may have that could make them lean more or less in favor of IPV



FIGURE 4 Estimates of the multilevel logistic regression model of decisions on IPV cases on sex and level of experience. Coefficients are expressed in odds ratio scale.

victims or aggressors. The third column of Table 3 gives further support to our second expectation (H2) once coefficients are estimated at the Judge-Year level, and the same holds for the effect of caseload (fourth column of Table 3), which gives support to our third hypothesis. Because we have observations from several years for each judge, what these models suggest is that judges' decisions change as they get more experienced and accumulated cases.

Let us now turn to the conditional effect of sex on both experience and workload. Figure 4 shows the interaction of experience with sex. First, the coefficient for the sex constitutive term captures the compared odds of granting restraining orders between male and female judges that have no professional experience at all. The error of the coefficient, though, is too large for the standard significance levels to hold. The constitutive term for experience yields a positive coefficient, capturing the effect of experience among male judges, while the coefficient of the interaction term captures that effect for female judges. Despite the significance of both the interaction and the experience terms, Figure 5 shows that both male and female judges present positive slopes for years of experience, although the slope for female judges is slightly larger, which suggests that on average the effect of experience is stronger among female judges, who by default are already more likely to grant restraining orders.

Turning now on the effect of individual workload, our model predicted an opposite result, that is, that judges' workload should work negatively on their odds of granting restraining orders, especially for male judges because it would reinforce male judges' tendency to grant fewer restraining orders.

JOURNAL OF

JOURNAL OF EMPIRICAL LEGAL STUDIE



FIGURE 5 Marginal effect of years of experience on granting restraining orders for IPV victims, conditional on sex.



FIGURE 6 Estimates of the multilevel logistic regression model of decisions on IPV cases on sex and workload. Coefficients are expressed in odds ratio scale.

Figure 6 shows the coefficients of the model including an interaction between sex and caseload. The constitutive term for sex shows that at zero levels of workload, there are no differences among male and female judges in granting

661

17401461, 2023, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jels.12561 by Readcube (Labria Inc.), Wiley Online Library on [2904/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/etms-and-conditions) on Wiley Online Library or notes of use; O A articles are governed by the applicable Creative Commons Lienses



FIGURE 7 Estimates of the multilevel logistic regression model of decisions on IPV cases on sex, experience, and workload. Coefficients are expressed in odds ratio scale.

restraining orders. Given that there is no judge with zero workload in the data, this result is of little use. However, the coefficient of the interaction term suggests that caseload does not have a significant effect on female judges, while the term for caseload indicates that it does for male judges and in a negative way.

To further test this hypothesis, our last model fits the interaction between sex, experience, and caseload altogether, which are presented in Figure 7. Our objective here is to explore the weight these three factors have on judicial decisions of both male and female judges. Again, the constitutive terms for both sex and experience are of little use, as they capture effects when caseload is zero. The constitutive term for caseload is negative as it was in the previous model, indicating the negative effect of workload on inexperienced male judges, thus suggesting that when males judges have high uncertainty (due to little experience), workload reinforces their likelihood to grant less restraining orders.

In contrast, the interaction between females and caseload indicates that among inexperienced female judges, workload has no significant effect, which would point to the expected result of our model. However, the last two interactions of the model add some more information to the results. On the one hand, the interaction between caseload and experience captures the effect of these two factors on male judges, and it is positive, suggesting that experience moderates the negative effects of workload on male judges. In other words, the results suggest that at least in male judges experience can moderate the negative effect of

662

JOURNAL OF

EMPIRICAL LEGAL STUDIE

workload. In contrast, on the other hand, the last interaction suggests that among female judges the presence of caseload might counteract the positive effect of experience.

To clarify the consequences of these results, Figure 8 plots the marginal effects of experience between male and female judges at different levels of caseload. The left panel of the figure indicates the difference in slopes between those male and female judges that have workloads below one standard deviation from the mean. In these situations, when work pressure is low, experience gives incentives to female judges to grant more restraining orders, while it does not seem strong enough to increase the chances of males judges to favor victims at higher rates. Once male judges get higher levels of workload, though, experience does increase their odds to grant more restraining orders, and differences between male and female judges dissipate.

This can be further observed when we plot the predicted probabilities of male and female judges granting restraining orders at different levels of experience and workload, as in Figure 8. Notice that at lower workload levels, experience makes female judges significantly more likely to grant restraining orders than male judges. However, when workload increases, the role of experience is stronger for both male and female judges, which gives support to the idea that when information costs increase due to workload judges resort more to intuition to make their decisions (Guthrie et al., 2007), and professional experience provides them with tools to do so rather efficiently.



FIGURE 8 Marginal effects of sex on levels of experience at different levels of workload.

CONCLUSIONS

JOURNAL OF

EMPIRICAL LEGAL STUDIES

In this article, we have explored factors behind judicial inequality in pretrial decisions on restraining orders as part of IPV cases in Spain. In particular, we have analyzed to what extent gender as an in-group identity mechanism operates together with uncertainty and information costs associated with decisions around cases where women's rights are clearly at stake.

To test the effect of these factors, we have used data from on-duty pretrial court decisions on restraining orders in Spain between 2010 and 2018. First of all, the paper conceptualizes gender as an in-group mechanism that should work on both male and female judges, unlike a large part of the literature that tends to observe the behavior of female judges as deviations from normality.

The paper also provides a model of judicial decision that unpacks the cost component of the judicial utility framework provided by Epstein et al. (2013). In particular, we have provided a model that captures how difficult is for judges to assess the cost function of a decision. Once the model has been set up, we have used data on intimate-partner cases to test our model.

Our empirical analysis of judicial behavior yields four main findings. First, gender explains a portion of the observed variation in judicial decisions on IPV cases, especially when controlling for levels of experience and workload. Our second finding is that, as judges gain experience and therefore reduce their average levels of uncertainty, they tend to approve restraining orders at higher rates, which is consistent with our model. Our third major result is that individual workload, our measure of information costs, is a major determinant of judicial outcomes, to the point of significantly decreasing the probability of judges granting restraining orders to protect victims. This finding is consistent with were vious literature on common-law courts, where high caseload is associated with worse judicial outcomes for plaintiffs. The argument is that when the information costs of making decisions increase, decision-makers stick harder to the rules.

Related to this, finally, we have also observed that the effects of both caseload and professional experience are conditional on judges' gender under some circumstances. Results show that professional experience has positive effects among both male and female judges, except at low levels of workload, in which case it has no effect on male judges. This suggests two main implications. First, when workload is low, experience reinforces the role of gender as an in-group mechanism, as it makes male judges less likely to grant restraining orders even at different levels of experience. However, on the other hand, when judges face higher volume of workload, experience is a strong predictor of judicial decisions as it positively affects the changes of both male and female judges to grant restraining orders. This last result suggests that when information costs are high (workload), judicial expertise provides judges with a useful tool to reduce the uncertainty of decisions. Major sources of judicial inequality for IPV victims are both endogenous and exogenous. On the one hand, judges' gender becomes relevant when dealing with cases where fundamental rights are at stake, and female judges are systematically more ready to protect victims than male judges. This effect gets stronger as judges become expert decision-makers through experience. On the other hand, an exogenous factor, such as workload, can affect the conditions under which judges must decide, to the point of systematically reducing the chances of IPV victims of getting their restraining orders approved.

Beyond our specific results on IPV cases, we believe the paper makes an important contribution to judicial politics. Lower pretrial judges in civil-law countries typically are appointed on purely technical grounds according to their performance in competitive exams based on legal knowledge, after which they become civil servants with life tenure. This feature of civil-law systems helps depicting judges as anonymous decision makers with completely unobservable attributes, a picture that is even sharper when we observe behavior down the judicial ladder. However, we show that even in civil-law systems, lower court judges' behavior can be shaped by identity-based mechanisms to the point of producing unequal decisions.

Since unequal court decisions on IPV cases have critical consequences on the everyday lives of millions of women in Spain, they may have an impact on the overall legitimacy and impartiality of the justice system.

ACKNOWLEDGMENTS

The project leading to these results GENEQUAL has received funding from "la Caixa" Banking Foundation under the project code LCF/PR/SR19/52540006. The data that support the findings of this study are available at https://www.pepvallbe.cat/publications.

DATA AVAILABILITY STATEMENT

Data necessary to replicate the results of this article are available upon request from the corresponding author.

REFERENCES

- Angrist, J. D., & Pischke, J.-S. (2014). *Mastering 'metrics': The path from cause to effect*. Princeton University Press.
- Ash, E., Asher, S., Bhowmick, A., Chen, D., Devi, T., Goessmann, C., Novosad, P., & Siddiqi, B. (2021). Measuring gender and religious bias in the Indian judiciary. Working Paper. Center for Law & Economics Working Paper Series.
- Bainbridge, S. M., & Gulati, G. M. (2002). How do judges maximize? (the same way everybody else doesboundedly). Rules of thumb in securities fraud opinions. *The Emory Law Journal*, 51, 83–151.
- Boyd, C. L. (2016). Representation on the courts? The effects of trial judges' sex and race. *Political Research Quarterly*, 69(4), 788–799.
- Boyd, C. L., Epstein, L., & Martin, A. D. (2010). Untangling the causal effects of sex on judging. *American Journal of Political Science*, 54(2), 389–411.

- JOURNAL OF EMPIRICAL LEGAL STUDIES
- Cafaggi, F., & Iamiceli, P. (2021). Uncertainty, administrative decision-making and judicial review: The courts' perspectives. *European Journal of Risk Regulation*, 12, 792–824.
- Cameron, C. M., Kastellec, J. P., & Park, J.-K. (2013). Voting for justices: Change and continuity in confirmation voting 1937–2010. *The Journal of Politics*, 75(2), 283–299.
- Correll, S. J. (2001). Gender and the career choice process: The role of biased self-assessments. *American Journal of Sociology*, 106(6), 1691–1730.
- Crowe, N. E. (1999). The effects of judges' sex and race on judicial decision making on the U.S. Courts of Appeals, 1981–1996. PhD thesis, University of Chicago, Department of Political Science.
- Dyevre, A. (2010). Unifying the field of comparative judicial politics: Towards a general theory of judicial behaviour. *European Political Science Review*, 2(2), 297–327.
- Engel, C., & Weinshall, K. (2020). Manna from heaven for judges: Judges' reaction to a quasirandom reduction in caseload. *Journal of Empirical Legal Studies*, 17(4), 722–751.
- Epstein, L., Landes, W. M., & Posner, R. A. (2013). *The behavior of federal judges: A theoretical and empirical study of rational choice*. Harvard University Press.
- Farhang, S., & Wawro, G. (2004). Institutional dynamics on the U.S. Court of Appeals: Minority representation under panel decision making. *The Journal of Law, Economics, and Organization*, 20(2), 299–330.
- Galanter, M. (2004). The vanishing trial: An examination of trials and related matters in federal and state courts. *Journal of Empirical Legal Studies*, 1(3), 459–570.
- Garoupa, N., Gómez-Pomar, F., & Grembi, V. (2013). Judging under political pressure: An empirical analysis of constitutional review voting in the Spanish constitutional court. *The Journal of Law, Economics, and Organization*, 29(3), 513–534.
- Gazal-Ayal, O., & Sulitzeanu-Kenan, R. (2010). Let my people go: Ethnic In-group Bias in judicial decisions—Evidence from a randomized natural experiment. *Journal of Empirical Legal Studies*, 7(3), 403–428.
- Gelman, A., & Hill, J. (2007). Data analysis using regression and multilevel/hierarchical models. Cambridge University Press.
- Geyh, C. G. (2016). *Courting peril: The political transformation of the American judiciary*. Oxford University Press.
- Gigerenzer, G., & ABC-Research-Group (Eds.). (2008). Simple heuristics that make us smart. Oxford University Press.
- Gigerenzer, G., & Engel, E. (Eds.). (2006). Heuristics and the law. The MIT Press.
- Guarnieri, C. (2003). Courts as an instrument of horizontal accountability: The case of Latin Europe. In J. M. Maravall & A. Przeworski (Eds.), *Democracy and the rule of law* (pp. 223– 241). Cambridge University Press.
- Guarnieri, C. (2010). Judicial Independence in authoritarian regimes: Lessons from continental Europe. In R. Peerenboom (Ed.), *Judicial Independence in China: Lessons for global rule of law* promotion (pp. 234–246). Cambridge University Press.
- Guthrie, C., Rachlinski, J. J., & Wistrich, A. J. (2007). Blinking on the bench: How judges decide cases. Cornell Law Review, 93, 1–43.
- Hanretty, C. (2012). Dissent in Iberia: The ideal points of justices on the Spanish and Portuguese constitutional tribunals. *European Journal of Political Research*, 51(5), 671–692.
- Hanretty, C. (2013). The decisions and ideal points of British law lords. British Journal of Political Science, 43(3), 703–716.
- Hanretty, C. (2020). A court of specialists: Judicial behavior on the UK supreme court. Oxford University Press.
- Hilbink, L. (2007). Judges beyond politics in democracy and dictatorship: Lessons from Chile. Cambridge University Press.
- Huang, B. I. (2011). Lightened scrutiny. Harvard Law Review, 124(5), 1109-1152.
- Jolls, C., & Sunstein, C. R. (2006). The law of implicit bias. California Law Review, 94(4), 969-996.

666

Kahneman, D. (2012). Thinking, fast and slow. Penguin Books.

- Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). Judgment under uncertainty: Heuristics and biases. Cambridge University Press.
- Keith, L. C., Holmes, J. S., & Miller, B. P. (2013). Explaining the divergence in asylum grant rates among immigration judges: An attitudinal and cognitive approach. *Law & Policy*, 35(4), 261–289.
- Knepper, M. (2018). When the shadow is the substance: Judge gender and the outcomes of workplace sex discrimination cases. *Journal of Labor Economics*, *36*(3), 623–664.
- Lavie, S. (2016). Appellate courts and caseload pressure. Stanford Law & Policy Review, 27, 55-99.
- March, J. G., & Simon, H. A. (1958). Organizations. John Wiley.
- Martin, E., & Pyle, B. (2005). State high courts and divorce: The impact of judicial gender. University of Toledo Law Review, 36(4), 923–943.
- Newell, A. (1990). Unified theories of cognition. Harvard University Press.
- Richards, D. (2016). When judges have a hunch: Intuition and experience in judicial decision-making. Archives for Philosophy of Law and Social Philosophy, 102(2), 245–260.
- Roback, P., & Legler, J. (2021). Beyond multiple linear regression: Applied generalized linear models and multilevel models in r. CRC Press.
- Rodríguez-Menés, J., & Rovira, M. (2019). Assessing discrimination in correspondence studies. Sociological Methods & Research, 50(4), 1584–1622.
- Rudman, L. A., Moss-Racusin, C. A., Phelan, J. E., & Nauts, S. (2012). Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology*, 48(1), 165–179.
- Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. Journal of Risk and Uncertainty, 1, 7–59.
- Schneider, D. M. (2001). Assessing and predicting who wins federal tax trial decisions. Wake Forest Law Review, 37, 473–529.
- Schneider, M. R. (2005). Judicial career incentives and court performance: An empirical study of the German labour courts of appeal. *European Journal of Law and Economics*, 20(2), 127–144.
- Segal, J. A. (2000). Representative decision making on the federal bench: Clinton's district court appointees. *Political Research Quarterly*, 53(1), 137–150.
- Segal, J. A., & Cover, A. D. (1989). Ideological values and the votes of US supreme court justices. *American Political Science Review*, 83(2), 557–565.
- Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1), 99–118.
- Simon, H. A. (1957). Rational choice and the structure of the environment. In Models of man, social and rational (pp. 261–273). John Wiley & Sons.
- Simon, H. A. (1997). Administrative behavior: A study of decision-making processes in administrative organizations (4th ed.). The Free Press.
- Sonnemans, J., & van Dijk, F. (2012). Errors in judicial decisions: Experimental results. Journal of Law, Economics, and Organization, 28(4), 687–716.
- Sunstein, C. R., Schkade, D., Ellman, L. M., & Sawicki, A. (2007). Are judges political? An empirical analysis of the Federal Judiciary. Brookings Institution Press.
- Vallbé, J.-J. (2015). Frameworks for modeling cognition and decisions in institutional environments. A data-driven approach. Law, governance and technology (Vol. 21). Springer.
- Vallbé, J.-J., González-Beilfuss, M., & Kalir, B. (2019). Across the sloping meadow floor: An empirical analysis of pre-removal detention of deportable non-citizens. *Law and Society Review*, 53(3), 740–763.
- Weinshall, K., Sommer, U., & Ritov, Y. (2017). Ideological influences on governance and regulation: The comparative case of supreme courts. *Regulation & Governance*, 12, 334–352.
- Zorn, C., & Bowie, J. B. (2010). Ideological influences on decision making in the Federal Judicial Hierarchy: An empirical assessment. *The Journal of Politics*, 72(4), 1212–1221.

JOURNAL OF EMPIRICAL LEGAL STUDIES

How to cite this article: Vallbé, J. J., & Ramírez-Folch, C. (2023). The effect of judges' gender on decisions regarding intimate-partner violence. *Journal of Empirical Legal Studies*, 20(3), 641–668. <u>https://doi.org/10.1111/jels.12361</u>