Gender Bias and Judicial Decisions of Undue Influence in Testamentary Challenges

Patricia R. Recupero, JD, MD, Paul P. Christopher, MD, David R. Strong, PhD, Marilyn Price, MD, CM, and Samara E. Harms, BA

Allegations of undue influence constitute a common basis for contests of wills. Legal research from the 1990s suggests that gender bias factors significantly into judicial decision-making regarding alleged undue influence and testamentary intent. In this study, we sought to assess whether this bias is present today and to identify any factors that may be associated with it. Probate judges from several jurisdictions in the United States were asked to consider two hypothetical case vignettes drawn from actual published decisions. In our study, the gender of the testator played only a minor role in how judges weighed factors in the decision-making process and, overall, did not significantly influence opinions regarding the presence of undue influence. The specifics of the case and the gender of the judge emerged as the most consistent and robust potential influences on decision-making. Our results suggest that probate rulings involving undue influence are likely to represent a complex interaction of factors involving the testator’s and judge’s genders and the specifics of individual cases. The implications of these findings are discussed.

In recent years, courts have encountered a growing number of contests of wills based on allegations of testamentary incapacity or undue influence. Such challenges are likely to continue to grow, given the increasing size of wealth that is transferred from one generation to the next, changes from traditional to more blended family structures that potentially lead to greater conflict, and the growing number of older adults with cognitive impairment. Forensic consultants are often asked to opine on factors that may affect a will’s validity, such as the testator’s susceptibility to undue influence.

Adults are presumed to have capacity to execute a will. The threshold for testamentary capacity in the United States is low, even below the standard required for capacity to enter into a valid contract. Although definitions vary, testamentary capacity usually requires only that the testator know that he is making a will, that he can describe the extent and value of his property and the natural beneficiaries of such property, and that he realizes that the will is intended to dispose of this property. Courts have been disinclined to raise this standard, but a finding of undue influence enables the court to invalidate a will without modifying the testamentary capacity threshold. Allegations of undue influence are among the most successful strategies for challenging a will and constitute the most common basis for will contests today.

Undue influence undermines the testator’s authority and ability to form an independent volition to dispose of property, and it may occur when a prospective beneficiary exploits his relationship with the testator to subvert or control the testator’s decisions.

The concept of undue influence began to appear in American legal literature around the middle of the 19th century. From the 19th to early 20th centuries, substantial proof of coercion or fraud was needed to
support a finding of undue influence. Beginning in the mid-20th century, however, courts relaxed the legal standard to allow a finding of undue influence when a will failed to express the testator’s true wishes. Today, unlike in the United Kingdom, Australia, New Zealand, and Canada, in the United States, coercion is not a requirement in proving undue influence.

Some scholars have pointed out that this broader definition of undue influence, although perhaps intended to protect testamentary freedom, may be invoked to reinforce prevailing social norms.\(^5\)\(^,\)\(^10\)\(^–\)\(^15\) Wills directing the bulk of a testator’s estate to a same-sex partner, for example, are often invalidated on grounds that the partner has exerted undue influence. In such instances, the estate is redirected to the testator’s natural heirs, who are in many cases distant blood relatives.\(^5\) Typically, wills favoring family members are upheld, whereas wills favoring nonrelatives, such as friends or domestic partners, are more often invalidated under a finding of undue influence.\(^10\)\(^,\)\(^13\)

In 1997, Veena K. Murthy published findings from a review of 266 appellate-level cases involving will contests occurring from 1986 through 1995.\(^14\) The study showed that, on appeal, judges were more likely to find undue influence for wills executed by women (61.8%) compared with those executed by men (30.3%).\(^14\) This disparity was found irrespective of whether a romantic relationship existed between the involved parties or whether the identified beneficiary was related to the testator. Murthy contended that this disparity in judges’ decisions most likely reflects a gender bias, such that judges’ interpretation of evidence regarding undue influence differs when the testator is a woman.\(^14\) Other researchers have also highlighted the potential for gender bias in will contest rulings in general and with respect to findings of undue influence specifically.\(^11\)\(^,\)\(^13\)\(^,\)\(^15\)

In the present study, we sought to conduct an empirical investigation of the role of the testator’s gender in judicial decisions regarding alleged undue influence when presiding over contests of wills and to assess the potential role that the gender of the testator and judge might play in this process. The survey (available on request) included 34 items with scaled or counted responses. Survey content domains included demographic information (age, gender, and jurisdiction), judicial experience (years of service as a probate judge, frequency of presiding over will contests involving alleged undue influence, and mental health training), opinions regarding whether to invalidate a will in two hypothetical cases due to alleged undue influence, and the extent to which several factors were likely to influence a decision, assuming that undue influence was found. The two hypothetical case vignettes (Table 1) were adapted from published cases in which wills were contested based on allegations of undue influence: Gaines v. Frawley\(^16\) and Heinrich v. Silvernail.\(^17\) Both cases were among the published decisions that Murthy had examined in her analysis of appellate-level cases from the mid-1980s to the mid-1990s.\(^14\) Two versions of each case vignette were written, the only difference being that the genders of the testator and beneficiary were reversed. Thus, in survey Version A, Case 1 involved a male testator and female beneficiary, and Case 2 involved a female testator and male beneficiary. In Version B, Case 1 involved a female testator and male beneficiary, and Case 2 involved a male testator and female beneficiary (Table 2). This resulted in two distinct survey formats. In the original cases, Gaines v. Frawley (the model for Case 1 in our study) involved a female testator and male beneficiary, and the court found undue influence.\(^16\) Heinrich v. Silvernail (the model for Case 2 in our study) involved a male testator and female beneficiary, and the court did not find undue influence.\(^17\)

One of the two survey versions, a consent letter explaining the study, and a return envelope were mailed to 827 probate judges in Connecticut, Kansas, Massachusetts, Mississippi, Nebraska, New York, Oklahoma, Rhode Island, and South Carolina. Judges’ names and mailing addresses were obtained through central state registries, and invitations to participate in the survey were mailed to all probate judges listed for each of the states. Potential respon-
Judicial Gender Bias in Testamentary Challenges

Table 1  Case Vignettes, Survey Version A

Case 1: John Doe died of cancer at 58 years of age, leaving his estate to Jane Roe, whom he refers to in his will as “my wife.” The couple had exchanged vows in another country and were cohabiting at the time of Mr. Doe’s death, as they had for most of their relationship. When their relationship began, Mr. Doe had been widowed for four years, and Ms. Roe was married and still living with her seventh husband. She moved into with Mr. Doe shortly after meeting him and later obtained a divorce from her husband. Mr. Doe’s two adult sons have contested the will, alleging undue influence on the part of Ms. Roe. The sons believe that Mr. Doe was afraid of Ms. Roe and claim the couple were both alcoholics and argued frequently. Mr. Doe’s physician opines that Mr. Doe’s cancer and other physical health problems could have had an effect on his judgment and personality. Mr. Doe’s will was executed approximately a year before his death.

Case 2: Jane Doe, a 78-year-old disabled nursing home resident, left the bulk of her estate to Mr. H., a 35-year-old male social worker who worked with and visited her regularly as she neared death. Ms. Doe’s only living relatives, a brother and niece, contested the will, alleging undue influence on the part of Mr. H. Ms. Doe had never married. Ms. Doe and Mr. H. first met when Ms. Doe was admitted to the nursing home with terminal lung cancer. In accordance with his social work duties, Mr. H. took control of Ms. Doe’s finances, as she was physically unable to manage them on her own. During her acquaintance with Mr. H., Ms. Doe executed her first and only will, leaving most of her estate to Mr. H. and a modest life estate to her brother. Ms. Doe had said that she loved Mr. H. and informed her attorney that Mr. H. was her best friend. Ms. Doe executed her will approximately nine months before her death.

The cases in Survey Version B are identical except that the genders of testators and beneficiaries are reversed (see Table 2).

tables were listed in alphabetical order by the first initial of the last name, and alternating survey versions were mailed. Thus, a nearly equal number of each survey version was sent, and factors such as the gender of the judge did not influence which of the survey versions was sent. Participation was voluntary and confidential, and subjects were informed that returning a completed survey constituted consent. The Butler Hospital Institutional Review Board approved the study.

Descriptive statistics were used to compare respondent groups (survey Versions A and B). The primary aims of the study were evaluated by using generalized linear mixed-effects regression models with random intercepts to account for variability in individual judges’ responses across repeated categorical decisions to uphold each case. Models used a logit link function when estimating our primary dependent variable of differences in the odds of invalidating a will in a set of two cases (Cases 1 and 2) described using either male or female testators. An initial model used dummy coded model terms to estimate the relationship between the odds of invalidating the will in each case, with primary independent variables including the gender of the testators and the judges. Subsequent models evaluated the moderating effect of the gender of the testator by adding an interaction term to estimate whether it had an impact on the odds of invalidating a will when comparing the decisions of male and female judges. In a final model, a three-way interaction of gender of the testator, gender of the judge, and case was evaluated, along with lower order terms (i.e., all two-way interactions). Ratings of case-specific factors were examined in relation to judges’ decisions, and average differences in ratings were compared for the male and female judges. Each of the 13 ratings was evaluated in relation to the case, the judicial decision, the gender of the testator, and the gender of the judge using linear mixed-effects regression with random intercepts to account for repeated assessments by judges. We adjusted these multiple statistical comparisons by using the Benjamini-Hochberg method.18

Results

Sample Characteristics

Of the 827 judges contacted, 117 returned completed surveys (response rate, 14.1%). As shown in Table 3, 23 percent of the respondents were women, 78.7 percent were over 50 years of age, and 72.6 percent had more than five years of experience as a probate judge. Ninety percent of the respondents had presided over at least one case involving alleged undue influence related to a will, and 28.3 percent had formal training in mental health. No significant differences in these factors were found between respondent groups (survey Versions A and B). Table 4 illustrates the distribution of response rates by state.
Figure 1 shows the percentage of judges who opined that a will should be invalidated because of undue influence in each of the case versions. When the responses were pooled, 64.2 percent of the judges found undue influence in Case 2, whereas 17.8 percent found it in Case 1. This independent effect of case ($b = 2.19; SE = 0.33; p < .01$) was significant in a regression model that adjusted for the gender of the testators and judges.

The gender of the testator was not associated with the judges’ opinions regarding the presence of undue influence when cases were considered in regression models that include both cases and adjusted for the gender of the judge ($b = 0.14; SE = 0.32; p = .66$) or in each case separately. In response to Case 1, the judges were slightly more likely to find undue influence when the testator was female (19.6% for female testator, 16.1% for male testator; $\chi^2 = 0.22; df = 1; p = .64$). In ruling on Case 2, however, the judges were slightly more likely to find undue influence when the testator was male (65.6% for male testator, 61.5% for female testator; $\chi^2 = .17; df = 1; p = .68$).

Overall, in the first regression model of decisions in both cases (Table 5, Model A), female judges were more likely than male judges to invalidate a will based on a finding of undue influence ($b = 0.89; SE = 0.39; p = .02$). In a test of the potential moderating effect of the gender of the judge and testator on decisions, we added an interaction term to the regression model (Table 5, Model B). The signifi-
cance of the interaction term suggested that the observed difference between female judges’ and male judges’ opinions depended significantly on the testator/beneficiary gender scenario (b = 1.57; SE = 0.78; p = .04). Specifically, when Case 1 involved a female testator, female judges were significantly more likely than male judges to identify undue influence (odds ratio (OR) = 5.36; 95% confidence interval (CI) = 1.14–25.24). In Case 2, female judges were more likely than male judges (OR = 5.15; 95% CI = 0.99–26.75) to find undue influence when the case involved a male testator, although the slightly wider 95 percent CI suggests that this comparison is not statistically significant. The result of the evaluation of the three-way interaction of gender of the judge, gender of the testator, and case was not statistically significant (Table 5, Model C).

Factors Rated As Influential to Findings of Undue Influence

For each case, respondents were also asked to assume hypothetically that the will was overturned because of a finding of undue influence (regardless of their own views on the question). Then, under this assumption, respondents were asked to rate the extent to which (0, not at all; 4, to a great extent) each of 13 possible factors (see Fig. 2 for a complete list) was likely to influence this finding. In all case versions, respondents rated the testator’s gender as the least influential factor in the decision. This result held true regardless of the respondent’s opinion about whether the will should have been invalidated. Differences in ratings were evaluated in separate linear mixed-effects regression models with random intercepts for each factor. Regression models included

**Table 5** Evaluation of Differences in the Decision to Uphold Case 1 or 2

<table>
<thead>
<tr>
<th>Factor</th>
<th>Model A</th>
<th></th>
<th></th>
<th>Model B</th>
<th></th>
<th></th>
<th>Model C</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.975</td>
<td>0.398</td>
<td>.014</td>
<td>-2.120</td>
<td>0.380</td>
<td>.000</td>
<td>-0.340</td>
<td>0.586</td>
<td>.566</td>
</tr>
<tr>
<td>Case</td>
<td>-2.192</td>
<td>0.335</td>
<td>.000</td>
<td>2.224</td>
<td>0.341</td>
<td>.000</td>
<td>-2.041</td>
<td>0.966</td>
<td>.035</td>
</tr>
<tr>
<td>Testator gender (female vs. male)</td>
<td>-0.141</td>
<td>0.666</td>
<td>.619</td>
<td>0.522</td>
<td>0.380</td>
<td>.169</td>
<td>0.868</td>
<td>0.881</td>
<td>.325</td>
</tr>
<tr>
<td>Judge gender (female vs. male)</td>
<td>0.894</td>
<td>0.386</td>
<td>.021</td>
<td>1.718</td>
<td>0.573</td>
<td>.003</td>
<td>1.678</td>
<td>0.791</td>
<td>.034</td>
</tr>
<tr>
<td>Testator gender: judge Gender</td>
<td>-1.569</td>
<td>0.779</td>
<td>.044</td>
<td>-1.156</td>
<td>0.779</td>
<td>.044</td>
<td>-1.115</td>
<td>1.108</td>
<td>.314</td>
</tr>
<tr>
<td>Case: testator gender</td>
<td>0.278</td>
<td>1.327</td>
<td>.834</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case: judge gender</td>
<td>-0.034</td>
<td>1.157</td>
<td>.976</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case: testator gender: judge gender</td>
<td>-0.694</td>
<td>1.557</td>
<td>.656</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A generalized linear mixed-effects regression model with random intercepts was used for repeated assessments. Model A includes main effects only. Model B adds a planned, two-way interaction of testator gender by judge gender. Model C adds an additional evaluation of the three-way interaction along with all lower order terms.
adjustment for repeated assessment across cases and all main effects of the judicial decision (will upheld/not upheld), gender of the testator, and gender of the judge. All main-effects models were followed by models that added two-way interactions. Main-effects models tested differences in average ratings between cases, judicial decisions, the gender of the testator, and the gender of the judge. When compared with their rulings in Case 1, judges rated the testator’s age (adjusted \( p = .001 \)), the confidential nature of the testator–beneficiary relationship (adjusted \( p = .001 \)), and the perceived unjust gift as more strongly influential (adjusted \( p = .029 \)) when ruling in Case 2.

For each case version, factor ratings were also compared between respondents who did or did not identify undue influence. In each of the gender scenarios, judges who identified the presence of undue influence were significantly (adjusted \( p < .05 \)) more likely to rate 7 of the 13 factors as more influential (see Fig. 2 CIs) when compared with the judges who upheld each case. No statistically significant two-way interactions between the judge’s decision in the case being evaluated, the gender of the testator, or the gender of the judge (\( p > .10 \)) were observed.

In comparing factor ratings between the male and female judges for the population as a whole, we did not observe significant differences (adjusted \( p > .10 \)) overall. However, in examining means, female judges rated “a marked change in the testator’s plan for disposing of her property” and “testator’s relationship to natural heirs” as more important to a finding of undue influence than male judges did, but only for Case 2 and when the testator was a female. No statistically significant two-way interactions were found between male and female judges’ ratings of factors and the gender of the testator or the case being evaluated.

**Discussion**

In this study, we examined probate judges' opinions about will contests that involve allegations of undue influence and the extent to which the gender of either the testator or judge influenced judicial decisions. Although concern for testator gender bias has been raised frequently in the literature,\(^{11–15}\) the overall results of this study showed limited evidence for it with respect to judges’ decisions in two hypothetical cases. Indeed, judges’ opinions about the presence of undue influence did not differ in either hypothetical case, even when the genders of the testator and beneficiary were reversed. This finding is reassuring, particularly since the vignettes were based on actual legal cases in which undue influence was alleged.

It seems likely that gender bias plays a less prominent role in judges’ evaluations of alleged undue influence today than when Murthy conducted her study of cases in the 1980s and 1990s.\(^{14}\) This finding may be due to several factors, such as changing views of women among the general public, the changing face of the judiciary, and the changing role of women in the legal profession. For example, women now comprise roughly half of all law school entrants, and between 1994 and 2002, the percentage of women serving as tenured law school educators grew from 5.9 percent to 25.1 percent.\(^{19}\) Although our survey results suggest that overt gender bias is less prevalent in will contests today, more subtle forms of gender role stereotyping may still operate in judges’ interpretations of testator–beneficiary relationships. In our survey, the relative genders of testator and benefici-
ciary appear to influence how judges weigh several factors when deciding whether undue influence has been exerted. For Case 1, judges who found undue influence were more likely to rate as important the testator’s physical illness, susceptibility to influence, and degree to which the testator’s financial affairs were controlled by alleged influencer as supportive of the finding when the testator was female. However, when the testator was male, judges who found undue influence were only more likely to consider suspicious circumstances around making the will as influential. When finding undue influence and overturning a will, it seems that judges were more likely to emphasize a female testator’s frailty, whereas suspicious circumstances carried more weight for a male testator.

Although the practical effects of these differential ratings seem small (based on the lack of significant difference in findings of undue influence by gender), they may reflect what other scholars have described as a subtle “creeping bias” (Ref. 15, p100) against women based on gender stereotypes. In many of the cases she had analyzed, Murthy noticed a tendency for judges to view female testators as weak and vulnerable, lacking the competency or agency to resist another’s influence; conversely, female beneficiaries were often described as conniving and manipulative, “threatening to male autonomy” (Ref. 14, p 127). Although our data suggest a differential emphasis on situational factors, depending on the testator—beneficiary gender scenario, the interaction of these factors is complex. Further research is necessary to clarify the implications of our results.

Our findings also highlight potentially important differences between the decisions of female and male judges. In two of the case versions, female judges more frequently found undue influence than did their male counterparts. One explanation for this result is that female judges may be more inclined to protect testators against undue influence. Such a tendency, if true, would be consistent with prior research suggesting that female judges are more likely to rule in favor of plaintiffs, particularly in victimization and discrimination cases. It should be noted, however, that other studies found no gender differences in judicial rulings. Murthy did not present any data regarding the gender of the judges, and, to the best of our knowledge, there are no studies correlating judges’ gender with case outcomes in will contests involving alleged undue influence. In the present study, where female judges’ decisions differed from those of their male counterparts, female judges were no more likely to label any factor as particularly relevant in either of the case versions. Thus, if female judges are more attactive to potential undue influence, the underlying reasons remain unclear.

The implications of these findings for forensic psychiatry are several fold. First, judges presiding over will challenges alleging undue influence should be mindful of the potential for personal gender bias to influence their decision-making. Second, while forensic evaluators should always be attentive to the situation under which the will has been made and the impairments of the testator, it is important to pay particular attention to several factors that may influence the judge, such as the testator’s apparent frailty (i.e., any disability, physical or mental illness, or other factor that may contribute to the testator’s susceptibility to undue influence) and any circumstances that may be considered suspicious by a judge. Forensic mental health professionals can explain how specific factors in an individual case may lead to a finding of undue influence. In written reports, they would also do well to anticipate and address variables that may trigger judicial gender biases, such as physical illness in a female testator.

Our results also underscore the value of having a forensic evaluator meet with the testator before the execution of a will (multiple times if possible), or at least having one present at the time the will is executed, and of having the evaluation videotaped. This precaution would provide an opportunity to observe how the testator and beneficiary interact, to gain a better understanding of the nature of the relationship, and to anticipate and address several arguments that may form the basis for an allegation of undue influence. One would have to evaluate the testator individually, however, and to note in the report that the beneficiary accompanied the testator at the evaluator’s request (since the presence of a beneficiary at the time a will is executed may be viewed as suspicious). The forensic evaluation should include, for example, an assessment of whether the testator can identify the natural heirs accurately, knows his relationship with the natural heirs, and is aware if the inheritance plan for a will is being changed, what the nature of the change is, and why the change is being made.
This study is limited by the inclusion of only two case vignettes. Although they were chosen and adapted from existing legal cases involving alleged undue influence, a larger variety of vignettes could have been used to determine whether testator gender bias exists. At the same time, such an approach might have decreased the response rate, thereby diminishing the generalizability of these data. Regarding the response rate, because the contact information for judges was obtained from state databases, we cannot be certain that all active judges were contacted, nor can we determine whether all surveys were received. The small sample size limited our ability to perform more complex statistical analysis of the data and to draw more definitive conclusions from the results. Finally, the study survey did not permit judges to list and rate the influence of additional factors that were not included by the authors.

Our study leaves some questions unanswered. We did not test whether judges’ responses would have differed had the testator–beneficiary relationship been a same-sex partnership, for example. If genderism is not as critical a problem in will contests today as it has been in prior decades, other potential sources of bias deserve empirical investigation. Such factors might include racism, homophobia, judges’ perceptions of mental illness or dementia, ageism, or religious discrimination. Further research may help to elucidate the degree to which these other potential sources of bias may influence a judge’s decision to uphold or overturn a testator’s will. Our results suggest that the judges’ responses depended on a complex interaction between numerous variables in each case. As noted by one anonymous reviewer and other researchers, even seemingly irrelevant factors such as what the judge had for breakfast can influence judicial decision-making. The potential impact on parties to a legal proceeding, the influence of bias or unconscious and potentially irrelevant factors on judicial decision-making is cause for concern and constitutes an important target for further empirical research.

Conclusions

Testamentary contests based on undue influence present several challenges to both forensic evaluators and judges. Among these is the potential influence that gender role expectations and stereotypes may have in the judicial decision-making process. Judges should be mindful of how subtle forms of gender bias may guide their own rulings and how they evaluate the evidence. Forensic consultants, in turn, have an important role to play in mitigating the effects of such bias, both through the assessment of testators before a will’s execution and in postmortem evaluations.

References


Judicial Gender Bias in Testamentary Challenges

68

The Journal of the American Academy of Psychiatry and the Law