

*Constitutional Sex Discrimination**

Lee Epstein, Andrew D. Martin, Lisa Baldez, & Tasina
Nitzschke Nihiser**

Abstract

Nearly thirty years have elapsed since the U.S. Supreme Court decided *Craig v. Boren*, a landmark case in the Court's constitutional sex discrimination jurisprudence. In *Craig*, the justices pronounced that they would apply neither the lowest level of scrutiny—rational basis—nor the highest level—strict scrutiny—to evaluate claims of sex

* Please send all correspondence to Lee Epstein. Email: epstein@artsci.wustl.edu; Post: Department of Political Science, Washington University, CB 1063, 1 Brookings Drive, St. Louis, MO 63130.

** Lee Epstein is the Edward Mallinckrodt Distinguished University Professor of Political Science and Professor of Law at Washington University; Andrew D. Martin is Assistant Professor of Political Science at Washington University; Lisa Baldez is Associate Professor of Government at Dartmouth College; Tasina Nitzschke Nihiser is a Ph.D. student at Washington University. We are grateful to the National Science Foundation, the Center for New Institutional Social Science, the Washington University School of Law, and the Weidenbaum Center on the Economy, Government, and Public Policy, for supporting our research. We also owe thanks to participants in Washington University's Workshop on Empirical Research in the Law ("WERL"), Susan Appleton, Jeffrey A. Segal, and Nancy Staudt for offering many useful suggestions; and to Scott Friedman, David Lewarchik, Shelby Johnston, and Peter Ryan for outstanding research assistance. We used R, SPSS, and Stata to conduct the analyses presented in this paper. The project's web site (<http://epstein.wustl.edu/research/sexdiscrimination.html>) houses a full replication archive, including a database containing all the cases and variables we used in this study, as well as the documentation necessary to reproduce our results.

discrimination. Rather, the Court invoked a standard “in between” the two, now known as intermediate or heightened scrutiny. Under this approach, the Court asks whether a law challenged on equal protection grounds is substantially related to the achievement of an important objective.

Certainly the *Craig* Court’s intermediate approach has its supporters; indeed, influential legal scholars are now advocating that courts adopt it to evaluate laws discriminating against gays and lesbians. But to many analysts, *Craig* (and its progeny) was and remains highly problematic. Among their claims is that the standard it instantiated is so “loose” and “amorphous” that it produces unpredictable results.

In this article, we seek to bring some empirical teeth to this debate by exploring patterns in sex discrimination litigation in the U.S. Supreme Court and in state courts of last resort. Our chief finding is that the critics of heightened scrutiny probably have the better case. At the very least, the *Craig* standard—while generating outcomes more favorable to parties alleging sex discrimination than did the traditional rational basis test—does, in fact, lead to far less predictable results than either rational basis or strict scrutiny. For reasons that may have little to do with the standard itself, courts are just as likely to uphold sex-based classifications as they are to eradicate them.

This finding has important implications for the future of sex discrimination litigation, as well as for the advancement of legal rights for gays and lesbians. As to the former, our results underscore the importance of elevating the standard used to adjudicate sex discrimination claims—a goal, as we demonstrate, that could be achieved in several distinct ways. As to gays and lesbians, our findings identify the possible costs and benefits associated with a litigation strategy designed to place their claims of discrimination in the intermediate scrutiny basket.

Table of Contents

I. Introduction	13
II. Equal Protection and the Supreme Court	21
<i>A. The Three-Tier Approach to Equal Protection</i>	22
<i>B. Assessments of the Three-Tier Approach</i>	25
III. An Empirical Analysis of the Equal Protection Tests	34
<i>A. Rational Basis and Strict Scrutiny</i>	36
<i>B. Heightened Scrutiny</i>	48
IV. Implications of the Analysis for Discrimination Based on Sex and Sexual Orientation	51
<i>A. Sex Discrimination Litigation</i>	53
<i>1. Equal Rights Amendment (ERA)</i>	54
<i>2. Women on the Bench</i>	58
<i>B. Discrimination Against Gays and Lesbians</i>	63
V. Conclusion	67

I. Introduction

Nearly thirty years have elapsed since the U.S. Supreme Court decided *Craig v. Boren*,¹ a landmark in its constitutional sex discrimination jurisprudence.² In *Craig*,

¹ 429 U.S. 190 (1976).

² *Id.* We follow Mary Anne Case, *Constitutional Sex Discrimination: The Law as a Quest for Perfect Proxies*, 85 CORNELL L. REV. 1447, 1447 (2000) and “use the term ‘sex discrimination’ to refer to discrimination between males and females.” The vast majority of Supreme Court inquiries into denials of equal protection on grounds of

the justices pronounced that they would apply neither the lowest level of scrutiny—rational basis—nor the highest level—strict scrutiny—to evaluate claims of sex discrimination.³ Rather, they would invoke a standard “in between” the two, now known as intermediate or heightened scrutiny. Under this approach, the Court asks whether a law challenged on equal protection grounds is substantially related to the achievement of an important

sex—our chief concern in this Article—have focused on this type of discrimination, that is, “on deprivations caused by rules that, on their face, distinguish between males and females.” *Id.*

³ Since we provide more details about these levels of scrutiny in Part II, suffice it to note here that prior to *Craig*, the Court invoked a two-tier approach to equal protection claims. Under this model, the Justices

upheld legislation so long as it was rationally related to a legitimate governmental purpose, unless that legislation impinged upon a “fundamental right” or classified on the basis of a “suspect trait” like race or ethnicity. Where a law employed a “suspect classification” or restricted the exercise of a “fundamental right,” the Court applied “strict scrutiny,” requiring that the government establish that the legislation was necessary to a compelling governmental objective and that no less restrictive alternative was available.

Peter S. Smith, *The Demise of Three-Tier Review: Has the United States Supreme Court Adopted a “Sliding Scale” Approach Toward Equal Protection Jurisprudence?*, 23 J. CONTEMP. L. 475, 477 (1997). See also LEE EPSTEIN & THOMAS G. WALKER, *CONSTITUTIONAL LAW FOR A CHANGING AMERICA: RIGHTS, LIBERTIES, AND JUSTICE* (2004); DANIEL A. FARBER ET AL., *CONSTITUTIONAL LAW* (2003); SUSAN G. MEZEY, *ELUSIVE EQUALITY* (2003); Norman T. Deutsch, Nguyen v. INS and the Application of Intermediate Scrutiny to Gender Classifications, 30 PEPP. L. REV. 185 (2003); Michael C. Dorf, *The Paths to Legal Equality*, 90 CAL. L. REV. 791 (2002); Gerald Gunther, *Foreword: In Search of Evolving Doctrine on a Changing Court: A Model for a Newer Equal Protection* 86 HARV. L. REV. 1 (1971); Risa E. Kaufman, *State ERAs in the New Era: Securing Poor Women’s Equality by Eliminating Reproductive-Based Discrimination*, 24 HARV. WOMEN’S L.J. 191 (2001).

objective.⁴

Since *Craig*, the Court has tinkered with this “in-between” standard. Most notably, in *United States v. Virginia* (the “VMI case”)⁵ it seemed to “ratchet up” *Craig*,⁶ stating that it would require an “exceedingly persuasive justification” to sustain a sex-based classification.⁷ Tinkering, as it turns out, is the operative word. Despite the speculation of some commentators,⁸ and

⁴ *Craig*, 429 U.S. at 197.

⁵ 518 U.S. 515 (1996).

⁶ Martha C. Daughtrey, *Women and the Constitution*, 75 N.Y.U.L. REV. 1, 21 (2000) (suggesting that in *VMI*, “Justice Ginsburg ratcheted up the already ‘heightened scrutiny’ another notch or two.”); see also Dorf, *supra* note 3 (claiming that “*United States v. Virginia* arguably ratcheted up the level of judicial scrutiny applicable to sex classifications from intermediate to nearly strict.”); *infra* notes 8 and 9.

⁷ *United States v. Virginia*, 518 U.S. at 531 (“Parties who seek to defend gender-based government action must demonstrate an ‘exceedingly persuasive justification’ for that action. Today’s skeptical scrutiny of official action denying rights or opportunities based on sex responds to volumes of history.”). Some trace this language back to *Mississippi Univ. for Women v. Hogan*, 458 U.S. 718, 724 (1982) and *J.E.B. v. Alabama ex rel. T.B.*, 511 U.S. 127, 136 (1994). See, e.g., Smith, *supra* note 3 at n.17 (asserting that “in recent gender discrimination challenges, the Court has applied a super-heightened scrutiny to equal protection challenges. The language for this redefinition of intermediate review derives from *Hogan* (requiring that gender-based governmental action demonstrate an ‘exceedingly persuasive justification.’)”).; see also David K. Bowers, *Cracking the Code of United States v. Virginia*, 48 DUKE L.J. 305, 320 (1998); Jason M. Skaggs, *Justifying Gender-Based Affirmative Action under United States v. Virginia’s “Exceedingly Persuasive Justification” Standard*, 86 CAL. L. REV. 1169, 1170 (1998); Laura Weinrib, *Protecting Sex: Sexual Disincentives and Sex-Based Discrimination in Nguyen v. INS*, 12 COLUM. J. GENDER & L. 222, 228 (2003).

⁸ As Heather L. Stobaugh notes:

Over the years, commentators have argued that Justice Ginsburg’s use of “skeptical scrutiny” and her heavy reliance on the “exceedingly persuasive justification” language in the majority opinion of *Virginia* introduced a

even Court members,⁹ *United States v. Virginia* has hardly generated a sea-change in the Court's approach to sex discrimination. In fact, in its most recent forays into the area the majority of the Court backed off *Virginia*,¹⁰

stricter test into cases involving gender-based classifications as opposed to the traditional level of intermediate scrutiny established in *Craig v. Boren*.

Heather L. Stobaugh, *The Aftermath of United States v. Virginia*, 55 SMUL. REV. 1755, 1755 (2002). For commentary illustrating Stobaugh's claim, see, e.g., Steven A. Delchin, *United States v. Virginia and Our Evolving "Constitution": Playing Peek-a-boo with the Standard of Scrutiny for Sex-Based Classifications*, 47 CASE W. RES. L. REV. 1121, 1134 (1997) (arguing that "VMI may be the instrument for establishing strict scrutiny"); Cass Sunstein, *Foreword: Leaving Things Undecided*, 110 HARV. L. REV. 6, 73 (1996).

⁹ For example, in his dissent in *United States v. Virginia*, 518 U.S. at 574, Justice Scalia accused the majority of "a de facto abandonment of . . . intermediate scrutiny." See also Daughtrey, *supra* note 6, at 22, who writes that "in an address to the University of Virginia School of Law shortly after the VMI decision was announced, [Justice Ginsburg said,] 'There is no practical difference between what has evolved and the ERA.'"

¹⁰ Since *VMI*, the Court has decided two cases mounting constitutional challenges to sex-based classifications, *Miller v. Albright*, 523 U.S. 420 (1998) and *Nguyen v. INS*, 533 U.S. 53 (2001). In both, it upheld the classification and invoked the specter of *Craig*, rather than *VMI*, in the process. As to *Miller*, the majority wrote, "[e]ven if, as petitioner and her amici argue, the heightened scrutiny that normally governs gender discrimination claims applied in this context, we are persuaded that the requirement imposed by §1409(a)(4) on children of unmarried male, but not female, citizens is substantially related to important governmental objectives." *Miller*, 523 U.S. at 434 n. 11 (citation omitted). Many scholars argue that this language "cannot be squared with . . . Court doctrine prohibiting sex-based classifications that are not supported by 'exceedingly persuasive justifications'." Cornelia T. L. Pillard and T. Alexander Aleinikoff, *Skeptical Scrutiny of Plenary Power*, 1998 SUP. CT. REV. 1 (1998). See also Kristin Collins, *When Fathers' Rights are Mothers' Duties*, 109 YALE L. J. 1669 (2002); Emily J. Gelhaus, *The New Lower Standard for Equal Protection Claims Concerning Gender*, 71 U. CIN. L. REV. 305. As to *Nguyen*, the Court's majority claimed that "[f]or a gender-based classification to

retreating to the familiar territory of *Craig*.¹¹ “The message of *United States v. Virginia*,” turned out to be “no different from *Craig*: gender classifications are subject to

withstand equal protection scrutiny, it must be established ‘at least that the [challenged] classification serves important governmental objectives and that the discriminatory means employed are substantially related to the achievement of those objectives.’” *Nguyen*, 533 U.S. at 60. Because the Court once again relied on the standard articulated in *Craig* rather than the standard of *VMI* and *Nguyen*, many scholars argue that this represented “a marked shift away from the Court’s gender-based equal protection analysis set forth in *Virginia*.” Stobaugh, *supra* note 8, at 1756. *See also* Dorf, *supra* note 3, at n. 93 (suggesting that while “*United States v. Virginia* arguably ratcheted up the level of judicial scrutiny . . . *Nguyen* . . . ratcheted it back down.”). Stobaugh speculates that this “retreat” from *Virginia* occurred because

five Justices united to pull in the reins on the use of the “exceedingly persuasive justification,” which they believe imposes a higher level of scrutiny on gender-based classifications under Justice Ginsburg’s *Virginia* opinion. In other words, five Justices saw this area of the Court’s jurisprudence moving in a direction they did not support—toward treating gender as a suspect class—and they used *Nguyen* to prevent that result.

Stobaugh, *supra* note 8 at 1757.

¹¹ *See* Weinrib, *supra* note 7, at 227 (stating that “in light of the landmark anti-discriminatory outcome of *United States v. Virginia* . . . the apparent retreat in *Nguyen* came to many as an unpleasant surprise.”); Stobaugh, *supra* note 8, at 1756 (noting that “[d]espite such scholarly speculation about *Virginia*’s positive impact on future gender-based equal protection claims, it now appears not enough weight was given to those Justices who had expressed a strong dislike for the ‘exceedingly persuasive justification’ language and had disapproved of Justice Ginsburg’s heavy reliance on it.”); Bowsher, *supra* note 7, at 318 (“The Court has repeatedly applied the standard expressed in *Craig* without any further changes to the message.”). Also worth noting are Bowsher’s findings, at 306-07, that “[o]f the six federal Courts of Appeals that have considered whether *United States v. Virginia* heightened the standard of scrutiny, five have concluded that it did not.”

intermediate scrutiny.”¹²

Certainly *Craig*'s intermediate approach has its supporters;¹³ influential legal scholars are now advocating that courts adopt it to evaluate laws discriminating against gays and lesbians.¹⁴ But to many analysts, *Craig* (and its

¹² Bowsher, *supra* note 7, at 320. *But see* Edward M. Gaffney, *Curious Chiasma*, 4 U. PENN. J. CONST. L. 394, 396-97 (2002) (claiming that “if the formal language of the standard used to evaluate claims of gender discrimination is not strict scrutiny, it is something very close to that sort of exacting review. It rarely meets a classification that it likes From *Hoyt v. Florida* to *United States v. Virginia*, the movement in the protection of gender equality has been from low to high, from toothless irrationality to de facto strict scrutiny.”).

¹³ *See, e.g.*, Bowsher, *supra* note 7, at 317-18 (arguing that “despite the objections of three members of the *Craig* Court, intermediate scrutiny has since proven to be very workable.”); Collin O’Connor Udell, *Signaling A New Direction in Gender Classification Scrutiny*, 29 CONN. L. REV. 521, 548 (1996) (claiming that “from a feminist perspective, the intermediate scrutiny standard was certainly preferable to the ‘mere rationality’ formulation,” but also asserting that “we can do better [than *Craig*].”); *Miss. Univ. for Women v. Hogan*, 458 U.S. 718, 724 n.9 (1982) (“when a classification expressly discriminates on the basis of gender, the analysis and level of scrutiny . . . does not vary simply because the objective appears acceptable to individual Members of the Court.”).

¹⁴ Some commentators urge the application of heightened scrutiny to laws that discriminate between homosexuals and heterosexuals, that is, these analysts “seek to garner intermediate scrutiny for gays as gays.” Kenji Yoshino, *Suspect Symbols: The Literary Argument for Heightened Scrutiny for Gays*, 96 COLUM. L. REV. 1753. *See, e.g.*, Ann Shalleck, *Revisiting Equality: Feminist Thought About Intermediate Scrutiny*, 6 AM. U. J. GENDER & LAW 31 (1997) (“As in the struggle for women’s equality, heightened constitutional scrutiny matters both symbolically and practically in the movement for gay and lesbian rights.”); Yoshino, at 1756 (“In the near future, the United States Supreme Court is likely to consider the argument that gays should receive heightened scrutiny The main purpose of this Article is to strengthen [that] argument.”) Another group of commentators suggest that discrimination against gays and lesbians is, in fact, discrimination on the basis of sex. Hence, courts should apply the same level of scrutiny to classifications based on sexual orientation as they do for

progeny) was and remains highly problematic. Among their claims is that the standard it instantiated is so “loose”¹⁵ and “amorphous”¹⁶ that it produces unpredictable results¹⁷—so much so that it was “only a partial victory in women’s rights advocates’ campaign for equality.”¹⁸

In this article we seek to bring some empirical teeth to this debate by exploring patterns in sex discrimination litigation in the U.S. Supreme Court and in state courts of last resort.¹⁹ Our chief finding is that the critics of

laws that discriminate on the basis of sex. See Andrew Koppelman, *Why Discrimination Against Lesbians and Gay Men Is Sex Discrimination*, 69 N.Y.U. L. REV. 197 (1994); Sylvia A. Law, *Homosexuality and the Social Meaning of Gender*, 1988 WIS. L. REV. 187. See also Toni M. Massaro, *Gay Rights, Thick and Thin*, 49 STAN. L. REV. 45, 81 (1996) (stating that an “equality-based argument against antigay policies is that such policies constitute gender discrimination When one is barred from entering into a marriage with a same-sex partner who satisfies all other legal criteria of marriage eligibility, then one has been denied a marriage license solely because of sex. This triggers intermediate scrutiny.”); Andrea L. Clausen, *Marriage of Same-Sex Couples in Iowa*, 6 J. GENDER, RACE, & JUST. 451, 461 (2002) (“[I]f the Iowa Supreme Court were to recognize denial of same-sex marriage as a sex-based claim, but did not want to classify sex as a suspect class, the court could use intermediate scrutiny.”); SUZANNE PHARR, *HOMOPHOBIA: A WEAPON OF SEXISM* (1988). See also *infra* Part IV. For commentary advocating the use of intermediate scrutiny in other areas, see Rosemary M. Kennedy, *The Treatment of Women Prisoners after the VMI Decision*, 6 AM. U. J. GENDER & LAW 65 (1997) and Tamra M. Boyd, *Keeping the Constitution’s Promise: An Argument for Greater Judicial Scrutiny of Federal Alienage Classifications*, 54 STAN. L. REV. 319 (2001).

¹⁵ Justice David Souter, quoted in Skaggs, *supra* note 7, at 1190.

¹⁶ Joan A. Lukey and Jeffrey A. Smagula, *Do We Still Need a Federal Equal Rights Amendment?*, 44 BOSTON BAR J. 10, 26 (2000).

¹⁷ See also, John Galotto, *Strict Scrutiny for Gender, via Croson*, 93 COLUM. L. REV. 508, 545 (1993) (writing that the “standard is chafing at the Court and [unconvincing] to most scholars.”).

¹⁸ Shalleck, *supra* note 14, at 33.

¹⁹ In Part III, we explain why we focus on state courts of last resort rather than on the lower federal bench. Suffice it to note here that

20 TENNESSEE JOURNAL OF LAW & POLICY [VOL. I, 1]

heightened scrutiny probably have the better case. At the very least, the *Craig* standard—while generating outcomes more favorable to parties alleging sex discrimination than did the traditional rational basis test—does, in fact, lead to far less predictable results than either rational basis or strict scrutiny. Courts are, for reasons that may have little to do with the standard itself, just as likely to uphold sex-based classifications as they are to eradicate them.

This finding has important implications for the future of sex discrimination litigation, as well as for the advancement of legal rights for gays and lesbians. As to the former, our results underscore the importance of elevating the standard used to adjudicate sex discrimination claims—a goal, as we demonstrate, that could be achieved in several distinct ways. As to gays and lesbians, our findings identify the possible costs and benefits associated with a litigation strategy designed to place their claims of discrimination in the intermediate scrutiny basket.²⁰

We arrive at these implications in three steps. We begin in Part II with a description of the three-tier approach federal courts use to analyze claims of discrimination under the Fourteenth Amendment's Equal Protection Clause, along with a long-standing critique of that approach—that it leads, especially in the mid-level tier, to indeterminate results. In Part III, we demonstrate that this critique has merit. In particular, we show that while the application of the lowest and highest standards does, in fact, lead to rather predictable outcomes, the “in-between” standard does not. This result leads us in Part IV to emphasize the importance of elevating sex to the highest level of scrutiny and to question efforts designed to treat discrimination against gays and lesbians as sex discrimination—at least until

because all three levels of scrutiny have been used in these courts to adjudicate claims of constitutional sex discrimination, they provide excellent laboratories for exploring patterns in this area of the law.

²⁰ See *supra* note 14 and *infra* Part IV B.

courts move away from *Craig*'s in-between test.

II. Equal Protection and the Supreme Court

To say that the Supreme Court's equal protection jurisprudence has generated its fair share of commentary over the past few years is to make a rather uncontroversial claim. In fact, in the wake of recent decisions in the areas of sex discrimination,²¹ affirmative action,²² and gay rights,²³ scholars have scrutinized virtually every aspect of the Justices' approach to classifications based on sex, race, sexual orientation, and the like.²⁴ We do not intend to review the range of commentary, which would require a book or two. Instead we have two objectives. First, since an appreciation of current debates over *Craig* and its progeny requires some knowledge of the Court's three-tier approach to equal protection, we provide a brief overview of it in Part II A. Not much more is necessary since it has been so well described elsewhere.²⁵ Second, since we explicitly seek to put the debate on firmer empirical ground, we identify in Part II B the chief claims of both proponents and opponents of the Court's jurisprudence, first with regard to the determinacy of the results yielded by the current three-tier approach, and then with particular emphasis on its application to sex discrimination.

²¹ See, e.g., *Nguyen v. INS*, 533 U.S. 53 (2001).

²² *Grutter v. Bollinger*, 539 U.S. 306 (2003); *Gratz v. Bollinger*, 539 U.S. 244 (2003).

²³ *Lawrence v. Texas*, 539 U.S. 558 (2003), *reh'g denied*, 124 S.Ct. 35 (2003).

²⁴ For a recent and particularly insightful analysis, see Robert C. Post, *Foreword: Fashioning the Legal Constitution: Culture, Courts, and Law*, 117 HARV. L. REV. 4 (2003).

²⁵ For a sampling of work, see *supra* note 3.

A. The Three-Tier Approach to Equal Protection

To analyze claims of discrimination under the Fourteenth Amendment's Equal Protection Clause,²⁶ judges, at least until *Craig*, applied one of two standards.²⁷ Under the traditional rational basis test, as Table 1 shows, courts presume the validity of whatever classification the government has made (e.g., allowing only those over the age of 18 to enter into contracts, or permitting only M.D.s to perform surgery); it is up to the party challenging the law to establish that it is irrational. Since this burden is difficult to meet, the conventional view among scholarly commentators is that rational basis leads to a predictable outcome: courts defer to the government, generally upholding its classification.²⁸

²⁶ The Equal Protection Clause of the Fourteenth Amendment is restricted to the states; the governing constitutional provision for claims of discrimination against the federal government is the Due Process Clause of the Fifth Amendment. For purposes of our discussion on sex discrimination, the two clauses are interchangeable.

²⁷ We adopt and adapt some of the discussion in this Part from Lee Epstein et al., *Do We Still Need an ERA?* at <http://epstein.wustl.edu/research/ERA.html>.

²⁸ See, e.g., EPSTEIN & WALKER, *supra* note 3; MEZEY, *supra* note 3; Barbara A. Brown et al., *The Equal Rights Amendment: A Constitutional Basis for Equal Rights for Women*, 80 YALE L. J. 871 (1971); Gunther, *supra* note 3; Kaufman, *supra* note 3. As we foreshadowed in Part I, some contemporary commentators take issue with this prediction. For their views, see Part II B.

Table 1. Equal protection tests²⁹

Test	Example of Application	Validity Standard
Rational Basis	Age discrimination	The law must be a <i>reasonable</i> measure designed to achieve a <i>legitimate</i> government purpose.
Intermediate scrutiny	Sex discrimination	The law must be <i>substantially related</i> to the achievement of an <i>important</i> objective.
Strict scrutiny	Race discrimination	The law must be the <i>least restrictive</i> means available to achieve a <i>compelling</i> state interest.

Until the 1970s, the vast majority of claims of discrimination proceeded under the rules of the traditional rational basis test—with one particularly relevant exception: race. In light of the history surrounding ratification of the Fourteenth Amendment, the Court has held that classifications based on race should be subject to a less surmountable standard, known as “strict scrutiny” (or “suspect class”).³⁰ Under this standard, judges presume that a government action is suspect or unconstitutional. Only by showing that the law is the least restrictive means available to achieve a compelling state interest can the government overcome that presumption (see Table 1). Given the difficulty of making this showing, a conventional

²⁹ EPSTEIN & WALKER, *supra* note 3, at 645. See also *supra* note 3 and *infra* note 89.

³⁰ See supporting citations *supra* note 3.

wisdom emerged: the application of strict scrutiny generally leads to outcomes just as predictable as those under rational basis—only, of course, in the opposite direction: when courts apply this stricter test, they almost always rule in favor of the party alleging discrimination. Or, as Gunther famously put it, the suspect class test is “‘strict’ in theory and fatal in fact,” whereas the traditional rational basis standard provides “minimal scrutiny in theory and virtually none in fact.”³¹

It is thus no wonder that as part of their attempt to eradicate discrimination, women’s rights groups, beginning in the late 1960s, attempted to convince courts that sex-based classifications should be subject to strict scrutiny rather than to a rational basis analysis.³² Their litigation efforts did not succeed, but neither did they wholly fail. In response to their claims, the U.S. Supreme Court in *Craig v. Boren*³³ articulated a new standard, often called intermediate or heightened scrutiny, that falls somewhere between rational basis and strict scrutiny.³⁴ Under it, the

³¹ Gunther, *supra* note 3, at 8. Justice O’Connor has taken issue with this claim, asserting strict scrutiny is not always “strict in theory, but fatal in fact.” *Grutter*, 123 S.Ct. at 2338. Some scholars agree, arguing more broadly, as we do in Part II B, that the standard three-tier approach fails to produce reliable expectations. Nonetheless, it is true that many contemporary commentators suggest that Gunther’s assertion remains generally apt. See, e.g., EPSTEIN & WALKER, *supra* note 3; MEZEY, *supra* note 3; Farber et al., *supra* note 3; Lukey, *supra* note 16. See also Part III for our attempt to empirically evaluate this debate.

³² For descriptions of these efforts, see generally Daughtrey, *supra* note 6; Karen O’Connor & Lee Epstein, *Beyond Legislative Lobbying: Women’s Rights Groups and the Supreme Court*, 67 JUDICATURE 134 (1983).

³³ *Craig*, 429 U.S. at 218.

³⁴ For a history of the litigation leading up to *Craig*, including the Court’s decisions in *Reed v. Reed*, 404 U.S. 71 (1971) and *Frontiero v. Richardson*, 411 U.S. 677 (1973), see LEE EPSTEIN & JACK KNIGHT, *THE CHOICES JUSTICES MAKE* (1998); Daughtrey, *supra* note 6; Gaffney, *supra* note 12; Weinrib, *supra* note 7.

challenged law must be substantially related to the achievement of an important government objective (see Table 1).³⁵

Many argue that application of intermediate scrutiny leads to more favorable outcomes for parties alleging sex discrimination than did the traditional standard.³⁶ At the same time, though, they suggest that the intermediate approach, as opposed to rational basis or strict scrutiny, produces far less predictable results: the Court may more often than not void sex-based classifications, but it more than occasionally upholds them.³⁷

B. Assessments of the Three-Tier Approach

As we noted at the onset of this section, no shortage of critical commentary exists on the Court's three-tier approach to equal protection. Some comes from the bench itself, such as Justice Stevens's statement that "there is only one Equal Protection Clause. It requires every State to govern impartially. It does not direct the courts to apply

³⁵ As we explained in Part I, in *United States v. Virginia*, the majority articulated a variation on this approach which would require an "exceedingly persuasive justification" to sustain a sex-based classification. See *supra* note 7. But in its two most recent cases the Court "retreated" to the standard articulated in *Craig*. See *supra* note 10.

³⁶ See, e.g., Udell, *supra* note 13 (noting that "from a feminist perspective, the intermediate scrutiny standard was certainly preferable to the 'mere rationality' formulation"). See also LESLIE F. GOLDSTEIN, CONTEMPORARY CASES IN WOMEN'S RIGHTS (1994); Case, *supra* note 2.

³⁷ See, e.g., MEZEY, *supra* note 3; Deborah L. Brake, *Sex as a Suspect Class: An Argument for Applying Strict Scrutiny to Gender Discrimination*, 6 SETON HALL CONST. L.J. 953 (1996); Erin Chlopak, *Mandatory Motherhood and Frustrated Fatherhood: The Supreme Court's Preservation of Gender Discrimination in American Citizenship*, 51 AM. U. L. REV. 967 (2002); Deutsch, *supra* note 3; Lukey, *supra* note 16.

one standard of review in some cases and a different standard in other cases.”³⁸ There are numerous critics of the heightened scrutiny standard in particular. Mary Ann Case, for example, asserts that “the components” of the intermediate standard “have rarely been the moving parts in a Supreme Court sex discrimination decision,”³⁹ whereas Jason Skaggs points out the inconsistency of “subject[ing] gender-based affirmative action programs to strict scrutiny while analyzing all other gender-based classifications under intermediate scrutiny.”⁴⁰

These scholars make interesting and useful points. Especially relevant for our project, though, are analyses centering on the reliability of the results yielded by the three-tier approach.⁴¹ The issue raised in these analyses—

³⁸ *Craig*, 429 U.S. at 211-12. Justice Stevens was responding to a potential critique of the now three-tiered, as opposed to the traditional two-tiered, approach. As he goes on to say: “Whatever criticism may be leveled at a judicial opinion implying that there are at least three such standards applies with the same force to a double standard.”

³⁹ *Case*, *supra* note 2, at 1449. Rather, she says, “to determine whether there is unconstitutional sex discrimination, one need generally ask only two questions: 1) Is the rule or practice at issue sex-respecting, that is to say, does it distinguish on its face between males and females? and 2) Does the sex-respecting rule rely on a stereotype?”

⁴⁰ *Skaggs*, *supra* note 7, at 1174-75. For other critiques, *see* Shalleck, *supra* note 14 (strict-scrutiny standard as applied to gender issues also important for gay and lesbian rights); Katharine B. Silbaugh, *Miller v. Albright: Problems of Constitutionalization in Family Law*, 79 B.U. L. REV. 1139 (1999). For more general commentary on the Court’s jurisprudential posture toward women’s rights, *see* JUDITH A. BAER, *OUR LIVES BEFORE THE LAW: CONSTRUCTING A FEMINIST JURISPRUDENCE* (1999) (suggesting that it has worked to reinforce male dominance).

⁴¹ *See* Lee Epstein & Gary King, *The Rules of Inference*, 69 U. CHI. L. REV. 1, 83 (2002):

Reliability, [in empirical research], is the extent to which it is possible to replicate a measurement, reproducing the same value (regardless of whether it

whether various legal standards produce predictable and consistent outcomes—is important for many reasons, not the least of which is the implication for the norm of *stare decisis*. If courts decline to follow precedential rules of law or do so in unpredictable ways, they risk undermining their fundamental efficacy. Members of the legal and political communities predicate future expectations on the assumption that others will follow existing rules. Should courts make radical changes or apply rules inconsistently, these communities may be left unable to adapt.⁴²

From this general logic emerge two specific critiques of the Court's equal protection jurisprudence. One, which has emerged in recent years, suggests that a three-tier approach no longer provides an adequate framework for reliable expectations about court decisions; the other, a more conventional view we previewed in Part

is the right one) on the same standard for the same subject at the same time. If any one of us stepped on the same bathroom scale 100 times in a row, and if the scale [was] working properly, it would give us the same weight 100 times in a row—even if that weight is not accurate. (In contrast, a scale that is both reliable and valid will give a reading that is both the same and accurate 100 times in a row.) In other words, in empirical research we deem a measure reliable when it produces the same results repeatedly regardless of who or what is actually doing the measuring.

Id. This, it seems to us, is very similar to the criteria many commentators use to evaluate the predictability of legal standards.

⁴² See, e.g., Lee Epstein & Jack Knight, *supra* note 34; Jack Knight & Lee Epstein, *The Norm of Stare Decisis*, 40 AM. J. POL. SCI. 1018. It is along similar lines that scholars are concerned with reliability. They suggest that unreliable measurement procedures might provide evidence that the researcher, however inadvertently, has biased a measure in favor of a personal hypothesis, which can then undermine any inferences reached therein. See Epstein & King, *supra* note 41, at 83.

IIA, asserts that while intermediate scrutiny may be relatively unpredictable, the two lowest and highest tiers are not. As a general matter, the application of rational basis leads courts to uphold classifications and the application of strict scrutiny leads courts to strike them down.

Those who believe that the standard equal protection framework no longer produces uniform results (if, in fact, it ever did) do so for various reasons. For Ashutosh Bhagwat, the problem is that the “the Court has failed to develop any coherent framework regarding how, in applying the tiers of scrutiny, courts are to assess whether the governmental interest asserted satisfies the requirements of the level of scrutiny at issue.”⁴³ In Robert Post’s view, the Court can circumvent the three-tier approach altogether by strategically avoiding (even obvious) equal protection arguments.⁴⁴ The commonalities among these and other critiques, however,⁴⁵ may be more

⁴³ Ashutosh Bhagwat, *Affirmative Action and Compelling Interests: Equal Protection Jurisprudence at the Crossroads*, 4 U. PA. J. CONST. L. 260, 270 (2002).

⁴⁴ Post, *supra* note 24, at 99-101. He specifically points to *Lawrence v. Texas*, 124 S. Ct. 441 (2003), in which passages in the majority’s opinion “sound almost entirely in equal protection,” but which the Court decided on due process grounds. *Id.* at 99. Post argues that the justices took this route to avoid determining “whether classifications based upon sexual orientation should receive elevated scrutiny or merely rational basis review.” *Id.* at 100. Such strategic “instrumentation” on the part of judges seems to occur in other areas of the law, in other forms, and for a range of reasons. See Lee Epstein et al., *Dynamic Agenda-Setting on the Supreme Court: An Empirical Assessment*, 39 HARV. J. ON LEGIS. 395 (2002); Emerson H. Tiller & Pablo T. Spiller, *Strategic Instruments: Legal Structure and Political Games in Administrative Law*, 15 J.L. ECON. & ORG. 349, 362 (1999).

⁴⁵ See, e.g., Smith, *supra* note 3, at 476 (asserting “[t]he Rehnquist Court is moving away from [a three-tiered scheme of review] toward a more flexible approach. Moreover, even under the three-tier framework, the Court balances the importance of the rights or interests at stake with the government’s justification for the discriminatory

interesting than their differences. Broadly speaking, the argument advanced in study after study of judicial decisions is that although institutions—including legal standards—are certainly important, they are not as determinative as the three-tier framework might suggest. Indeed, the extant literature typically defines an institution as a set of rules that structures an interaction,⁴⁶ not as rules that establish outcome. Furthermore, it typically views the choices judges make as a function of many other forces, including personal political preferences, jurisprudential values, personal attributes, and the external environment in which they deliberate.⁴⁷

There are multiple examples within the equal protection realm that demonstrate the inadequacy of the three-tier framework.⁴⁸ Commentators choose among a wide array of disputes to justify their position—from *Romer v. Evans*,⁴⁹ in which the Court “use[d] the heightened scrutiny mode of analysis when it claim[ed] to

legislation in selecting and defining the appropriate tier. Recent cases have made it clear that the Court covertly employs [this] approach in every equal protection challenge.”)

⁴⁶ See, e.g., EPSTEIN & KNIGHT, *supra* note 34; WALTER F. MURPHY, *THE ELEMENTS OF JUDICIAL STRATEGY* (1964).

⁴⁷ See, e.g., DANIEL R. PINELLO, *GAY RIGHTS AND AMERICAN LAW* (2003); Gregory A. Caldeira et al., *Sophisticated Voting and Gate-Keeping in the Supreme Court*, 15 J. L. ECON. & ORG. 549 (1999); Erwin Chemerinsky, *Understanding the Rehnquist Court: An Admiring Reply to Professor Merrill*, 47 ST. LOUIS U. L.J. 659 (2003); Frank B. Cross & Emerson H. Tiller, *Judicial Partisanship and Obedience to Legal Doctrine: Whistleblowing on the Federal Courts of Appeals*, 107 YALE L.J. 2155 (1998); Epstein & Knight, *supra* note 34; Tracey George & Lee Epstein, *On the Nature of Supreme Court Decision Making*, 86 AM. POL. SCI. REV. 323 (1992); Gerard S. Gryski et al., *Models of State High Court Decision Making in Sex Discrimination Cases*, 48 J. POL. 143; DANIEL R. PINELLO, *GAY RIGHTS AND AMERICAN LAW* (2003); Richard L. Revesz, *Environmental Regulation, Ideology, and the D.C. Circuit*, 83 VA. L. REV. 1717 (1997).

⁴⁸ See *supra* note 47.

⁴⁹ 517 U.S. 620 (1996).

be employing rational basis review;”⁵⁰ to *United States v. Virginia*,⁵¹ in which the Court, at least according to Justice Scalia,⁵¹ applied strict (not heightened) scrutiny to assess a sex-based classification; to *Grutter v. Bollinger*,⁵² in which the Court applied strict scrutiny but nonetheless upheld the University of Michigan Law School’s use of race in admissions decisions. In short, because “the applicable level of scrutiny remains susceptible to modification—either ratcheted up to the most demanding standard or reduced to the most permissive test”⁵³—predictability is all but lost.

Many scholars, perhaps the majority, take issue with the purported inconsistency as it pertains to the highest and lowest levels of the three-tier approach. As we foreshadowed in Part II A, they say, and have long said, that these two extreme tiers have “evolved sub silentio so that the highest level, strict scrutiny, equates to an almost automatic conclusion of unconstitutionality, and the lowest, rational basis review, leads to an equally likely result of constitutionality.”⁵⁴ These analysts are aware of *Romer* and *Grutter*. They simply suggest that these cases are the exceptions to the general rule, as espoused by Gerald Gunther: strict scrutiny is “‘strict’ in theory and fatal in fact,” while rational basis provides “minimal scrutiny in theory and virtually none in fact.”⁵⁵

While proponents of the three-tier approach differ with critics on the expectations created by the rational basis and strict scrutiny tests, there is virtually universal

⁵⁰ Smith, *supra* note 3 at 476.

⁵¹ See *supra* note 9. Justice Scalia is not alone. See *supra* note 8 (scholarly works supporting Scalia’s contention).

⁵² 539 U.S. 306 (2003).

⁵³ Melanie K. Morris, *Ruth Bader Ginsburg and Gender Equality: A Reassessment of Her Contribution*, 9 CARDOZO WOMEN’S L. J. 1, 20-21.

⁵⁴ Bhagwat, *supra* note 43, at 270.

⁵⁵ Gunther, *supra* note 3, at 8.

agreement on the inherent unpredictability of the intermediate standard. Or, as Joan Lukey puts it:

Brennan's language in *Craig* indicated that this new test was more than the rational basis standard, under which classifications were almost always upheld. Still, it fell short of the onus of the strict scrutiny standard, under which almost every classification was struck down. It seemed that the concept of 'intermediate scrutiny' constituted a malleable, rather indeterminate standard of review, providing little or no guidance for lower courts—or even for future Supreme Court cases.⁵⁶

Several members of the current Supreme Court agree. During his confirmation proceedings, Justice Souter declared that the intermediate scrutiny test “is not good, sound protection. It is too loose.”⁵⁷ Prior to joining the Court, Justice Ginsburg cautioned that “variance within the federal judiciary will persist until the High Court provides unequivocal guidance by designating sex as a suspect classification requiring the application of strict judicial scrutiny.”⁵⁸ Dissenting in *Craig*, Justice Rehnquist asked:

How is this Court to divine what objectives are important? How is it to determine whether a particular law is 'substantially' related to the achievement of such objective, rather than

⁵⁶ Lukey, *supra* note 16, at 26. See also Udell, *supra* note 13, at 547 (arguing that “the Justices can draft such divergent opinions in their application of the intermediate scrutiny standard. . .”).

⁵⁷ Skaggs, *supra* note 7, at 1190.

⁵⁸ See Morris, *supra* note 53, at 21. Ginsburg made this claim in Brief of Amici Curiae of the American Civil Liberties Union et al. at 14, *Wengler v. Druggists Mut. Ins. Co.*, 446 U.S. 142 (1980) (No. 79-381).

32 TENNESSEE JOURNAL OF LAW & POLICY [VOL. I, 1]

related in some other way to its achievement?
Both of the phrases used are so diaphanous and
elastic as to invite subjective judicial
preferences or prejudices relating to particular
types of legislation, masquerading as judgments
.....⁵⁹

Two of the Court's most recent constitutional sex discrimination decisions illustrate these concerns. In *United States v. Virginia*, the U.S. government implored the Court to apply strict scrutiny to sex-based classifications, an invitation that Justice Scalia, along with many scholars, contend the majority "effectively" accepted.⁶⁰ Nevertheless, in a scathing critique of intermediate scrutiny, Norman Deutsch writes that "[t]he shoe was on the other foot in the Court's most recent gender case, *Nguyen v. INS*."⁶¹ In the course of upholding the law at issue—one that privileges a mother over a father in citizenship proceedings—the *Nguyen* majority held that the sex-based classification achieved important government interests and passed the heightened scrutiny test. Justice O'Connor disagreed. In a vigorous dissent, she went so far as to accuse the Court of explaining and applying "heightened scrutiny. . . in a manner. . . that is a stranger to our precedents."⁶² O'Connor went on to say that:

No one should mistake the majority's analysis
for a careful application of this Court's equal

⁵⁹ *Craig*, 429 U.S. at 221 (Rehnquist, J., dissenting).

⁶⁰ See *supra* note 9. See also Delchin, *supra* note 8 (arguing that "while Justice Ginsburg never expressly referred to the Government's argument in her majority opinion, several factors support Justice Scalia's contention that the Court [adopted strict scrutiny]"); Sunstein, *supra* note 8.

⁶¹ Deutsch, *supra* note 3, at 187.

⁶² *Nguyen*, 533 U.S. at 74.

protection jurisprudence concerning sex-based classifications. Today's decision instead represents a deviation from a line of cases in which we have vigilantly applied heightened scrutiny to such classifications to determine whether a constitutional violation has occurred. I trust that the depth and vitality of these precedents will ensure that today's error remains an aberration.⁶³

In short, O'Connor "not so subtly implied that the majority had, in effect, not applied intermediate scrutiny, but rational basis review."⁶⁴

It is hardly a surprise that many scholars, regardless of what position they take over rational basis and strict scrutiny, have come to see that:

[D]ispute over the proper application of the standard of review in *Nguyen* and *Virginia* is symptomatic of the fact that intermediate scrutiny is a 'made up' rule that has had little effect on the outcome of the decisions. . . . [I]n the end, the results in the cases turn on how the Court and the individual Justices view the underlying facts and policies, rather than on the verbalization of the standard of review as intermediate scrutiny.⁶⁵

If this is true, then as long as the Court continues to invoke the "murky" *Craig* rule, it will uphold or void classifications as it sees fit; and judges on state and lower federal courts will do the same or even "concoct" their own

⁶³ *Id.* at 97.

⁶⁴ Deutsch, *supra* note 3 at 187.

⁶⁵ *Id.* at 187-88.

approaches to sex discrimination.⁶⁶ As a result, all predictability, reliability, and consistency is lost.

III. An Empirical Analysis of the Equal Protection Tests

Several commentators dispute this view of intermediate scrutiny, claiming instead that its results are just as determinant as the rational basis and strict scrutiny tests. Edward Gaffney is exemplary: “If the formal language of the standard used to evaluate claims of gender discrimination is not strict scrutiny, it is something very close to that sort of exacting review. It rarely meets a classification that it likes . . . ”⁶⁷ while this claim may be relatively anomalous, as the discussion above indicates, it is nonetheless worthwhile to assess this claim against assertions flowing from more mainstream camps that the three-tier approach reveals either very little or a great deal about the likely outcomes of equal protection suits, with the notable exception of sex-discrimination litigation.

Accordingly, in what follows we undertake this task. In Part III A, we examine empirically the degree to which the highest and lowest tiers produce reliable outcomes: i.e., decisions upholding classifications under a rational basis analysis and decisions striking down classifications under strict scrutiny. Then, in Part III B we explore the extent to which the intermediate test produces predictable results.

In conducting these investigations, we focus

⁶⁶ See, e.g., MEZEY, *supra* note 3; Brake, *supra* note 37; Roberta W. Francis, *Reconstituting the Equal Rights Amendment: Policy Implications for Sex Discrimination*, paper presented at the annual meeting of the American Political Science Association, San Francisco (2001) (on file with the authors).

⁶⁷ Gaffney, *supra* note 12, at 396-97. See also Bowsher, *supra* note 7 at 317-18 (arguing that “despite the objections of three members of the *Craig* Court, intermediate scrutiny has since proven to be very workable”).

primarily on state courts of last resort (and then on sex-discrimination litigation),⁶⁸ even though the vast majority of commentary has centered on the U.S. Supreme Court. This focus reflects the difficulty of generalizing about the High Court's adjudication of equal protection disputes from a small number of cases (e.g., since 1960, the Justices have decided only 30 sex discrimination cases implicating the Equal Protection Clause)⁶⁹ and from a lack of variation in the standards employed (e.g., with the possible exception of *United States v. Virginia*, a majority of the Court has never applied strict scrutiny to sex-based claims)⁷⁰. These difficulties evaporate when we move to state supreme courts. Since 1960, state courts of last resort have addressed constitutional questions in 416 cases involving sex-based classifications, and they have done so using all three equal protection tests: rational basis, heightened scrutiny, and strict scrutiny.⁷¹

⁶⁸ Throughout this article we use the terms "state court of last resort" and "state supreme court" interchangeably, even though some state courts of last resort are not named "supreme court."

⁶⁹ We obtained this figure from MEZEY, *supra* note 3, at 16, which lists the 29 cases decided between 1971 and 2001. The U.S. Supreme Court Database identifies one case decided before 1971, *Hoyt v. Florida*, 368 U.S. 57 (1961). U.S. Supreme Court Database, at <http://www.polisci.msu.edu/pljp/databases.html> (last accessed on January 1, 2003).

⁷⁰ In *Frontiero v. Richardson*, 411 U.S. 677 (1973), a plurality of four Justices deemed sex a suspect class. But the four could not obtain a fifth vote, which led to the "compromise" in *Craig*. See EPSTEIN & KNIGHT, *supra* note 34.

⁷¹ The figure of 416—and all other data we present in this Article—comes from a database we amassed on all constitutional sex discrimination cases resolved in state courts of last resort between 1960 and 1999. Since that database (as well as all the documentation necessary to use it and an explanation of how we collected the information) is available on our web site, suffice it to note here that we included cases in which the state justices addressed a claim of constitutional sex discrimination and invoked an equal protection test in the course of addressing it. See Table 1 and *infra* note 89; see also

The use of different levels of scrutiny by state supreme courts with respect to the same class of disputes raises a number of interesting questions. Most importantly, why do justices in one state invoke strict scrutiny, while those in another apply the intermediate standard? On a somewhat different note, our attempt to gain insight into how the federal bench (especially the U.S. Supreme Court) employs the three-tier framework by focusing on the states and on sex discrimination raises a different set of questions. We address these matters in Part IV. For now, let us turn to the results of our analyses of the various tiers encompassing the judiciary's approach to equal protection.

A. Rational Basis and Strict Scrutiny

Were we to focus our empirical investigation exclusively on the U.S. Supreme Court's use of rational basis and strict scrutiny, it is entirely possible that we could muster support for virtually all existing commentary. Consider, for example, the conventional view that strict scrutiny and rational basis lead to predictable results. Using the former, the Court strikes classification, while under the latter, it upholds them. If we eliminate affirmative action cases from consideration,⁷² this conventional expectation seems to hold. Since the 1960s, for example, it is difficult to identify a single act discriminating on the basis of race and challenged on equal protection grounds that the Justices upheld. Conversely, it is equally difficult to identify a single law involving age discrimination that the Justices struck down as a violation of equal protection.⁷³

infra note 90.

⁷² We return to these cases momentarily.

⁷³ Using the U.S. Supreme Court Database (*see supra* note 69), we identified five age discrimination cases in the employment realm: *Kimel v. Fla. Bd. of Regents*, 528 U.S. 62 (2000); *Gregory v. Ashcroft*, 501 U.S. 452 (1991); *Vance v. Bradley*, 440 U.S. 93 (1979); *Alexander*

The existence of successful challenges in the race area only and none involving age discrimination hardly seems a coincidence: the former are subject to strict scrutiny, and the latter to rational basis. This distinction in standard may explain the results we observe. The Court, in fact, suggested as much in the recent age discrimination case of *Kimel v. Florida Board of Regents*:

States may discriminate on the basis of age without offending the Fourteenth Amendment if the age classification in question is rationally related to a legitimate state interest. The rationality commanded by the Equal Protection Clause does not require States to match age distinctions and the legitimate interests they serve with razorlike precision. . . . In contrast, when a State discriminates on the basis of race or gender, we require a tighter fit between the discriminatory means and the legitimate ends they serve. . . . Under the Fourteenth Amendment, a State may rely on age as a proxy for other qualities, abilities, or characteristics that are relevant to the State's legitimate interests. The Constitution does not preclude reliance on such generalizations. That age proves to be an inaccurate proxy in any individual case is irrelevant. . . . Finally, because an age classification is presumptively rational, the individual challenging its constitutionality bears the burden of proving that the "facts on which the classification is apparently based could not reasonably be conceived to be true by

v. Fioto, 430 U.S. 634 (1977); *Massachusetts Bd. of Retirement v. Murgia*, 427 U.S. 307 (1976). In none of the five did the Court find in favor of the party alleging discrimination.

the governmental decisionmaker.”⁷⁴

If we incorporate affirmative action into our analysis, however, the picture changes dramatically, lending some support to those who proclaim the indeterminacy of even the highest level of scrutiny. Since *Regents of the University of California v. Bakke*—⁷⁵ the Court’s first major foray into the area of affirmative action—it has decided eight constitutional cases centering on preferences for minorities.⁷⁶ Despite the Court’s use of the strict scrutiny standard to examine the programs at issue, it upheld four of the eight.⁷⁷

What should we take away from this analysis, as limited as it is, of the Supreme Court’s use of rational basis and strict scrutiny? Not much, as it turns out. Too few cases exist to reach any firm conclusions. Furthermore, as we suggested earlier, once the Court uses a particular test to evaluate a particular type of classification, it generally stays the course. The Court’s repeated application of rational basis in the age context is exemplary. With the possible exception of *United States v. Virginia*, a majority of the Court has never applied strict scrutiny to a sex-based claim;⁷⁸ and in only eight cases prior to *Craig* did it invoke the rational basis standard.

In an effort to overcome these problems, we turn to

⁷⁴ *Kimel*, 528 U.S. at 83-84 (citations omitted).

⁷⁵ 438 U.S. 265 (1978).

⁷⁶ We identified these cases using the U.S. Supreme Court Database. See *supra* note 69.

⁷⁷ The four it upheld were in *Fullilove v. Klutznick*, 448 U.S. 448 (1979); *United States v. Paradise*, 480 U.S. 149 (1987); *Metro Broad., Inc. v. Fed. Communications Comm’n*, 497 U.S. 547 (1990); *Grutter v. Bollinger*, 124 S. Ct. 35 (2003). The four it struck were at issue in *Wygant v. Jackson Bd. of Educ.*, 476 U.S. 267 (1986); *Richmond v. J. A. Croson Co.*, 488 U.S. 469 (1988); *Adarand Constructors v. Pena*, 515 U.S. 200 (1995); *Gratz v. Bollinger*, 123 S. Ct. 2411 (2003).

⁷⁸ As we noted *supra* at note 70, in *Frontiero*, a plurality, not a majority, of the Justices deemed sex a suspect class.

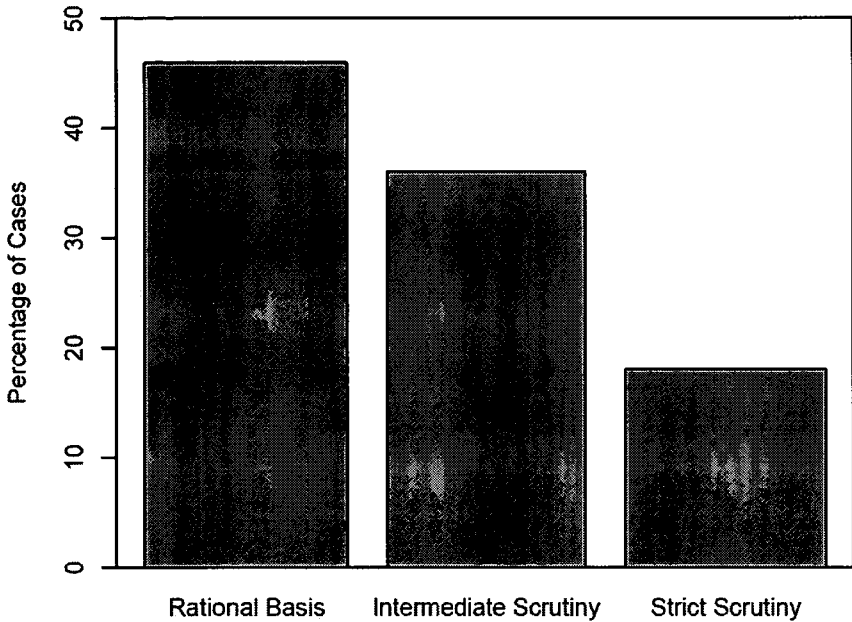
the states and, in particular, to how their courts of last resort have employed the two extreme levels of scrutiny to sex-based classifications. Since 1960, these courts have resolved 416 constitutional sex discrimination suits. As Figure 1 shows, minimal scrutiny is used in 45% (n=191), strict scrutiny in 18.0 % (n=75), and a version of the *Craig* intermediate standard in the remaining 150 (36.1%).⁷⁹ This is quite a bit of inter-state variation in its own right, and it is especially noticeable when we compare it against the U.S. Supreme Court's adjudication of sex-based discrimination cases. Between 1960 and 2002, the Justices heard 30 cases, using the intermediate standard, or a variant thereof,⁸⁰ in nearly 70%.

⁷⁹ For information on how we defined these standards, *see infra* note 89.

⁸⁰ *See supra* note 35.

Figure 1: Use of equal protection tests in constitutional sex discrimination litigation in state courts of last resort.

N=416⁸¹



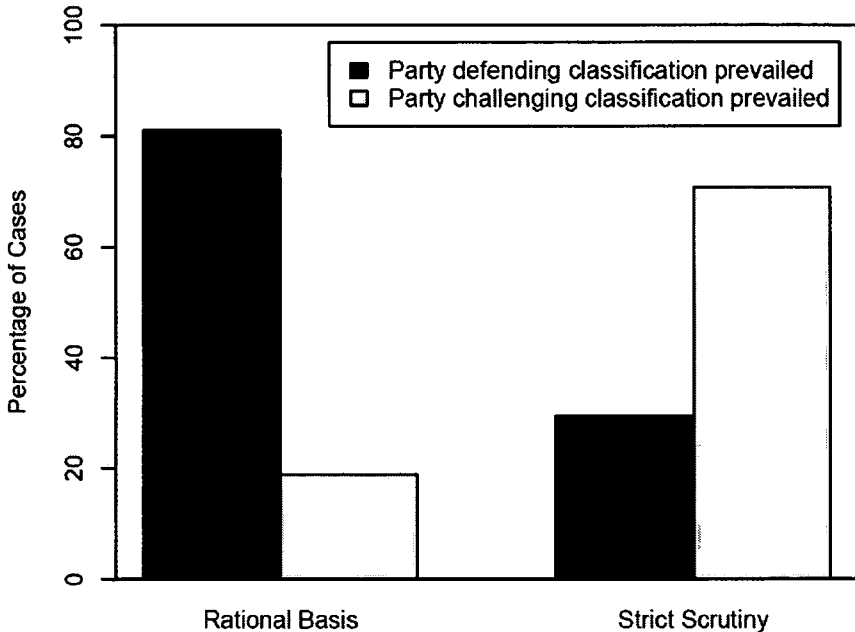
Let us consider the extent to which the upper and lower tiers produce reliable expectations about outcomes. Do courts invoking rational basis generally uphold classifications, while those employing strict scrutiny strike them down? A simple comparison, as seen in Figure 2, suggests that the answer is yes. The standard used by a court is, to a statistically significant degree,⁸² associated

⁸¹ See *supra* note 69 (information on the database we used to generate this figure); see also Table 1 and *infra* note 89 (how we defined these standards).

⁸² The association is statistically significant (Chi-Square = 62.64; $p < 0.001$).

with case outcomes. Of the 191 suits in which the court applied rational basis, the party alleging discrimination failed to prevail in 81.2% (155 of 191); when the state justices invoked strict scrutiny (22 of 75) the losing percentage was only 18.8%.⁸³

Figure 2: Prevailing party by the application of the rational basis and strict scrutiny tests in constitutional sex discrimination litigation in state courts of last resort. N=266.⁸⁴



⁸³ See *infra* note 90 (information on how we defined case outcomes).

⁸⁴ See *supra* note 69 (information on the database we used to generate this figure); see Table 1 and *infra* note 89 (how we defined these standards); see *infra* note 90 (how we treated case outcomes). See also *supra* note 80 for information on statistical significance.

From Figure 2, we can say that standards and outcomes appear to be associated in ways that many commentators would anticipate. Unfortunately, this sort of analysis does not enable us to make causal claims about that relationship; that is, from a simple comparison between standards and outcomes, we cannot claim that the standard employed caused the court to reach a particular outcome. There are many reasons for this, but most relevant here is that we have considered only the effect of the particular equal protection test on the case outcome and have failed to take into account other factors. To the extent that we have ignored various competing explanations, our simple comparison suffers from the most severe form of “omitted variable bias,” making inferences reached therein suspect.⁸⁵ For example, if we believe that politics plays a role in explaining court decisions, then failure to address the political composition of the deciding court could lead to an incorrect assessment of the true jurisprudential effect of legal standards.⁸⁶ Indeed, the impact of politics on judicial decision making could confound our results in any number of ways; for example, that only left-of-center courts invoke strict scrutiny or reach decisions in favor of the party

⁸⁵ Epstein & King, *supra* note 41; GARY KING ET. AL., *DESIGNING SOCIAL INQUIRY: SCIENTIFIC INFERENCE IN QUALITATIVE RESEARCH* (1994). Omitted variable bias occurs when a statistical comparison excludes variables that are (a) known to affect the outcome and (b) correlate with the explanatory covariate of interest.

⁸⁶ For more than six decades, political scientists and legal academics have documented the effect of the political preferences of judges on the decisions they reach. For recent examples, see JEFFREY A. SEGAL & HAROLD J. SPAETH, *THE SUPREME COURT AND THE ATTITUDINAL MODEL REVISITED* (2002); Theodore W. Ruger, et al., *The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting Supreme Court Decision-Making*, COLUM. L. REV. (forthcoming) (manuscript on file with authors). See generally Lee Epstein et al., *Childress Lecture Symposium: The Political (Science) Context of Judging*, 47 ST. LOUIS U. L.J. 783 (2003); works cited *supra* note 47.

alleging discrimination regardless of the standard employed.

In light of the large amount of literature suggesting the importance of ideology, neither of these scenarios is much of a stretch.⁸⁷ More generally, while there may be good reasons to believe that the adoption of a higher standard of law will generate outcomes more favorable to parties alleging discrimination, there are equally good reasons to question that assertion and believe that other factors come into play. At minimum, the conventional view about the determinacy of (the most extreme) tiers of the Court's equal protection framework may ask too much of institutions. While rules certainly can serve to structure choices, it seems imprudent to believe that they do all the work—especially when so many studies of judging suggest otherwise.

Accordingly, we must attend to (that is “control for”) the other factors that may affect case outcomes. Without performing statistical control, comparisons like Figure 2 are not informative about the possible causal relationships among the variables of interest. Also, we ought to account for the possibility that the choice of which test to employ may be influenced by numerous factors.⁸⁸ In

⁸⁷ See *supra* note 84.

⁸⁸ To be sure, federal courts are supposed to adhere to legal principles established by the U.S. Supreme Court, and state courts are supposed to view federal law as establishing a floor, instead of a ceiling, on civil rights and liberties. However, as the numerous studies, not to mention our own reading of state cases indicate, these norms do not always hold in this area of the law. Commentators point to federal courts that have all but ignored the current intermediate standard and have instead invoked higher or lower rules as they so desire; they also point to state courts of last resort that used a rational basis standard to adjudicate sex discrimination even after *Craig*. And, of course, there are states courts that treat sex as a suspect class. See, e.g., Brake, *supra* note 37; Branon P. Denning & John R. Vile, *Necromancing the Equal Rights Amendment*, 17 CONST. COMM. 593 (2000); Beth Gammie, *State ERAs: Problems and Possibilities*, 1989 U. ILL. L. REV. 1123 (1989);

a previous analysis of the state sex discrimination data, we took these steps, estimating a statistical model that incorporated four factors (or “variables”) to explain the equal protection test used by the court⁸⁹ and five factors, in addition to the test, to explain the outcome.⁹⁰ To be even more precise, we analyzed two equations, with two differentially measured dependent variables—the test (an ordinal variable that can take on three values: rational basis, intermediate scrutiny, and strict scrutiny⁹¹) and the outcome

Kaufman, *supra* note 3; MEZEY, *supra* note 3; Kevin Francis O’Neill, *The Road Not Taken: State Constitutions as an Alternative Source of Protection for Reproductive Rights*, 11 N.Y.L. SCH. J. HUM. RTS. 5 (1993). *See also* Part IV, in which we examine factors explaining state courts’ choice of standard.

⁸⁹ These variables are as follows: (1) the presence or absence of a state equal rights amendment, (2) the political ideology of the court, (3) the proportion of the court composed of women, (4) the existence or not of a state intermediate appellate court, and (5) whether or not the state had ratified the national ERA. *See* [<http://epstein.wustl.edu/research/sexdiscrimination.html>] (more details on these variables); *see also* Epstein et al., *supra* note 27 (the rationale behind including these variables).

⁹⁰ In addition to the equal protection test employed by the court, these variables are as follows: (1) the political ideology of the court, (2) the existence or not of a state intermediate appellate court, (3) a female as the party alleging discrimination, (4) a claim of a physical difference between men and women, and (5) the government as a defender of the sex-based classification. *See* [<http://epstein.wustl.edu/research/sexdiscrimination.html>] (more details on these variables); *see also* Epstein et al., *supra* note 27 (the rationale behind including these variables).

the data and documentation on our web site; for the rationale behind including these variables, *see* Epstein et al., *supra* note 27.

⁹¹ In the intermediate scrutiny category we also include a variation on that standard—that the government must offer an “exceedingly persuasive justification” for discriminating on the basis of sex—which some U.S. Supreme Court justices have endorsed. *See supra* note 35. In the strict scrutiny category, we also include a standard invoked occasionally by a few state courts—a standard that some observers liken to strict scrutiny, while others describe as “stricter” than strict

(a dichotomous variable in which the court either ruled for or against the party alleging discrimination⁹²)—in one model. Since no standard statistical model adequately performs this task, we developed one: a bivariate mixed response probit model, which allows for correlation across two equations and which we estimate using maximum likelihood.⁹³

scrutiny because it supposedly does not allow for sex-based classifications. To the extent that courts qualify this “stricter” standard with terms such as “absent compelling justifications” or in cases based on “actual differences,” however, the prohibition is not absolute, thereby lending credence to the view that it is akin to strict scrutiny. For our purposes, though, the key point is that it is closer, if not identical, to strict scrutiny.

⁹² To code case outcome, we rely on the approach commended by Gryski et al., *supra* note 47 and assess whether the party alleging sex discrimination won or lost the dispute. In taking this route, we are well aware of normative debates among some feminists over whether, as Goldstein, *supra* note 36, at 209, puts it, “to argue for ‘protective’ legislation for women on the grounds that without such legislation women are unfairly disadvantaged by making them play by rules that were designed with men in mind, and that are ill-adapted to women’s biology and life patterns.” While we appreciate this argument, our coding scheme remains relatively agnostic over it.

⁹³ See Epstein et al. *supra* note 27, at app. B (statistical model, along with our estimation methods). What is important here is that even though the parameter estimates resulting from these procedures admit to an interpretation akin to probit coefficients, our methodological approach is distinctive in two regards. First, it enables us to estimate parameters that, while substantively similar to those that would result from analyzing decisions over standards of law and case outcomes independently, are more efficient because we employ all the data to obtain them. Second, the approach facilitates a more exacting investigation of the dependence between the choices of standard and outcome because we are able to obtain a precise estimate of that dependence, in the form of an estimate of a correlation parameter, as a result of our ability to control for the factors that may affect both the standard and outcome in one model. Estimating this bivariate mixed response probit model leads to the results depicted in the table below—results that are quite striking: All the variables produce statistically significant coefficients in the expected direction. The estimate,

From this model, we gain a great deal of insight into the resolution of constitutional sex discrimination cases in the states. For example, we can now account for why state courts adopt different standards of law to adjudicate the same class of cases.⁹⁴ Furthermore, we can speak to the matter directly at hand: whether different equal protection tests lead to different results. The bivariate analysis presented in Figure 2 suggested that they do; and, as it turns out, this basic conclusion remains even after we control for the other relevant factors.

In particular, from our analyses, we find that the standard a court uses and the outcome it reaches are significantly correlated⁹⁵—in the direction many commentators would anticipate. Moreover, the probabilities displayed in Figure 3 reveal that the relationship is

indicating the correlation between the equal protection test used and the outcome reached, also attains statistical significance.

Parameter	MLE	Standard Error
β_1 Constant	-1.106	0.189 *
β_1 ERA	0.460	0.120 *
β_1 Judicial Ideology	0.011	0.004 *
β_1 ERA Ratification	0.456	0.143 *
β_1 Proportion Women	2.187	0.566 *
β_1 Intermediate Appellate Court	0.354	0.138 *
τ_1 Cut Point	1.179	0.084 *
β_2 Constant	-0.794	0.225 *
β_2 Judicial Ideology	0.017	0.004 *
β_2 Intermediate Appellate Court	0.376	0.153 *
β_2 Physical Difference	-0.685	0.160 *
β_2 Government Opposition	-0.725	0.130 *
β_2 Female	0.376	0.138 *
ρ Correlation	0.532	0.060 *

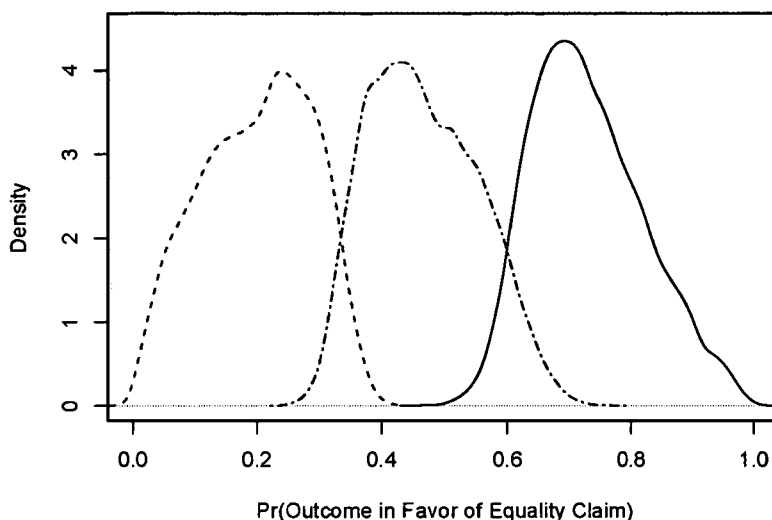
Maximum likelihood estimates and (asymptotic) standard errors for the bivariate mixed response probit model fit to the constitutional sex discrimination data. N = 416. lnL = -593.0884. * denotes statistical significance (? = 0.05.) See *supra* note 87 and *supra* note 88 (information on the variables used in this model).

⁹⁴ We describe these results later in Part IV, in which we explore the implications of our study.

⁹⁵ See the ρ coefficient in the table presented in note 91 *supra*.

substantively meaningful as well. Consider the two curves at the extreme right (representing rational basis) and the extreme left (strict scrutiny) of the figure, and notice the monotonic increase in the odds, such that when courts assess sex classifications via a rational basis test—the lowest level of scrutiny—the likelihood of finding in favor of the equality claim is just .20. That probability increases to a rather large .73 when they invoke strict scrutiny. In other words, and inline with much extant commentary, application of the lowest and highest standards leads to rather predictable outcomes—though in opposing directions: claims of sex discrimination will, on average, fail under a rational basis standard, and in all likelihood prevail under strict scrutiny.

Figure 3: Kernel density estimates of probabilities of an outcome favoring the litigant alleging sex discrimination given the rational basis standard (the left-most dashed line), the intermediate standard (the middle dashed line), and the strict scrutiny standard (the solid line).⁹⁶



B. Heightened Scrutiny

What our analysis thus far reveals is that state court adjudication of sex discrimination cases fits conventional views about the predictability of results yielded by rational basis and strict scrutiny approaches to equal protection. But what of the in-between tier—a tier that many, if not most, scholars suggest can lead to unexpected results? Are the outcomes as unpredictable as so many commentators

⁹⁶ These estimates account for all parameter uncertainty and were constructed from the simulation outlined in Epstein et al. *supra* note 27, at app. B. All covariates are held at their sample means. See *supra* note 69 (information on the database we used to generate this figure); Table 1 and *supra* note 89 (how we defined these standards); *supra* note 90 (how we treated case outcomes).

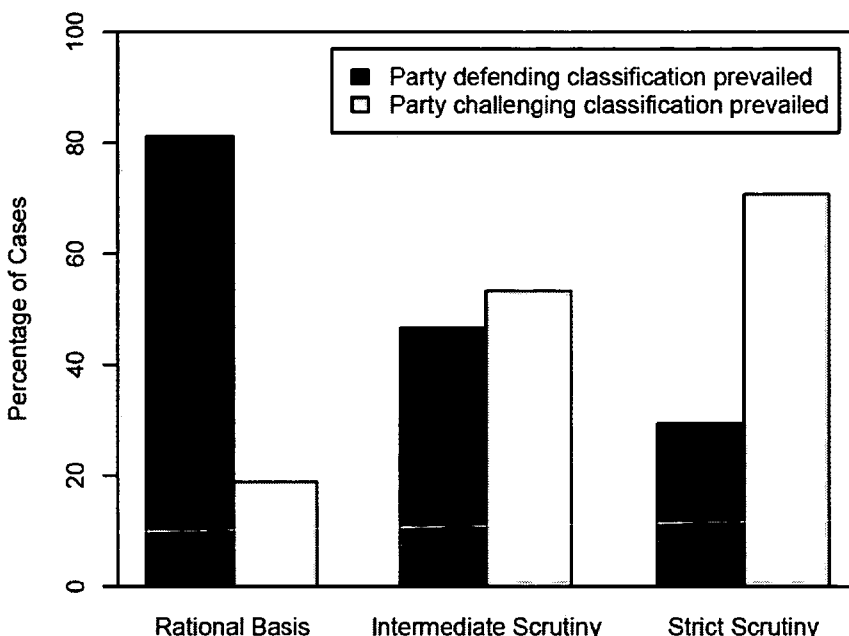
assert?

The middle curve displayed in Figure 3 (which represents heightened scrutiny) begins to provide an answer, and it is in the affirmative: When courts apply the intermediate standard, the probability that a litigant alleging discrimination will prevail is 47%. Just as many scholars would expect, under mid-level scrutiny litigants claiming sex discrimination are nearly as likely to win as they are to lose. This is in contrast to the relatively predictable outcomes generated by rational basis (under which a litigant faces only 20% likelihood of winning) and strict scrutiny (with a 73% probability of success).

Further analyses of the data do little to change the basic conclusion about the predictability—or, more pointedly, lack thereof—of the intermediate standard. For example, if we focus exclusively on the 150 cases in which state courts invoked this approach, we find once again, as Figure 4 shows, that parties alleging sex discrimination lost nearly as often as they prevailed (46.7% versus 53.3%). Moreover, the outcomes in the 150 cases are themselves somewhat difficult to predict. From analyses designed to take into account the multitude of factors that may affect court decisions, we were able to identify only a few that were substantively significant predictors. One, notably, was the ideology of the state justices deciding the dispute: the more left-of-center (“liberal”) the court, the more likely it was to apply intermediate scrutiny in a way favorable to the party alleging discrimination. Our model suggests that the most conservative court would find for the plaintiff alleging sex discrimination in only 26.1 % of the cases, while the most liberal court would do so in 70.9%. The government also plays an influential role. When it was the party defending discrimination, the court was far more likely to uphold the challenged classification (65.4% versus

33.0%).⁹⁷

Figure 4: Prevailing party by application of equal protection tests in constitutional sex discrimination litigation in state courts of last resort. N=416.⁹⁸



What these analyses tell us is that even after controlling for a range of relevant factors, a good deal of uncertainty remains about the conclusions state justices reach in constitutional sex discrimination disputes when they apply intermediate scrutiny—and those factors that

⁹⁷ We performed this analysis using a CLARIFY-like simulation. See *infra* note 109. The data and documentation necessary to reproduce these results are available on our web site.

⁹⁸ See *supra* note 69 (information on the database we used to generate this figure); Table 1 and *supra* note 89 (how we defined these standards); *supra* note 90 (how we treated case outcomes).

eliminate some of that uncertainty appear more related to politics than to the legal standard. If we move to the U.S. Supreme Court, however, even this degree of predictability vanishes. Since *Craig*, the justices have resolved 20 constitutional sex equality cases on equal protection grounds,⁹⁹ with the party alleging discrimination prevailing in fewer than half (nine out of 20, or 45%). While there may be some underlying explanation(s) to account for these outcomes, we were not able to identify a single one.¹⁰⁰ The Court does not seem to differentiate cases on the basis of whether a female litigant brought the claim, as some scholars have suggested,¹⁰¹ and it is not particularly deferential to the federal government (or the states, for that matter) when it attempts to defend a sex-based classification. Nor does the political ideology of the justices seem to exert much impact on their resolution of these cases. Though given the small number of cases we do not want to make too much of these findings, the latter is especially surprising in light of the large number of empirical studies ascribing a significant role for ideology in Court decision-making.¹⁰²

IV. Implications of the Analysis for Discrimination Based on Sex and Sexual Orientation

⁹⁹ See MEZEY, *supra* note 3.

¹⁰⁰ We considered four basic variables (under different measurements and specification): the ideology of the court, the presence or absence of women justices, whether a woman was claiming discrimination, and whether a government defended a sex-based classification. We measured the Court's ideology using the Segal & Cover scores. See Jeffrey A. Segal & Albert D. Cover, *Ideological Values and the Votes of U.S. Supreme Court Justices*, 83 AM. POL. SCI. REV. 557 (1989). See [<http://epstein.wustl.edu/research/sexdiscrimination.html>] (data and documentation necessary to reproduce these analyses).

¹⁰¹ See Baer, *supra* note 40.

¹⁰² See, e.g., studies cited in *supra* note 47 and *supra* note 84.

Given the results of our analyses, it would be difficult to take issue with Andrew Koppelman's conclusion about intermediate scrutiny. Even though he advocates treating laws that discriminate on the basis of sexual orientation as sex discrimination for purposes of equal protection analysis,¹⁰³ he recognizes that the sex discrimination argument is not free from indeterminacy. The question inevitably arises as to whether the state can offer an adequate justification for what it has done, and then a court must balance the interests involved in a way that will unavoidably allow for judicial discretion.¹⁰⁴

Our findings about the apparent indeterminacy of heightened scrutiny, not to mention our results reinforcing the relative predictability of rational basis and strict scrutiny, may lend support to Koppelman's conclusion. But what lessons should we take from our study? We see two as particularly important, one pertaining to the future of sex discrimination litigation, and the other, to the advancement of legal rights for gays.

We take up both in what follows. But before doing so, an important cautionary note is in order: Because we largely base these implications on analyses of state court decisions, we cannot state with any certainty the extent to which they transport to all American courts. More pointedly, using knowledge that we have gained from investigations of state cases to make inferences about federal litigation is a risky business indeed. The types of suits may differ,¹⁰⁵ as well as the parties,¹⁰⁶ to name just two points of distinction. Nonetheless, in light of the severe problems of addressing debates over equal

¹⁰³ See Koppelman, *supra* note 14.

¹⁰⁴ *Id.* at 535-36.

¹⁰⁵ For example, state courts typically do not resolve questions concerning federal immigration law or the military draft.

¹⁰⁶ For example, the U.S. government was not a participant in any of the 416 state cases in our database.

protection with reference only to federal courts, the state judiciary provides an antidote. For example, only by looking to the states—some of which have invoked strict scrutiny to resolve sex discrimination suits—were we able to get a handle on the counterfactual world: one in which federal courts deem sex a suspect class. And only by looking to the states, as we explain below, are we able to isolate those factors that could move sex-based litigation from the counterfactual to the factual: from a federal judiciary that now applies the intermediate standard to one that instead employs strict scrutiny.

A. Sex Discrimination Litigation

Throughout this article we have noted various expressions of dissatisfaction with intermediate scrutiny. While the critiques are many in number, one standing above virtually all others is the test's indeterminacy. From the vantage point of equality, the in-between approach may generate "better" outcomes than the traditional rational basis standard but it is highly unpredictable in application.

Our study confirms the veracity of this critique. As Figure 3 makes clear, if we believe it is desirable for courts to produce a larger number of equality-oriented outcomes, then heightened scrutiny better serves that objective than rational basis. Controlling for a host of other relevant factors, litigants challenging sex-based classifications are more than twice as likely to prevail now than they were prior to *Craig*. On the other hand, their odds, even under the intermediate test, are no better than 50-50, a far cry from the likelihood of victory under strict scrutiny—73%.

What these results underscore is a claim that advocates for women's rights have long made: the importance of elevating sex to a suspect class.¹⁰⁷ Until the

¹⁰⁷ See, e.g., MEZEY, *supra* note 3; Brake, *supra* note 37; Francis, *supra*

Supreme Court takes this step, federal courts will continue to employ the amorphous intermediate rule, sometimes upholding sex-based classification and sometimes voiding them—with little predictability.

On this much many agree. The question, of course, is how to alter the current standard. Our investigation into why state tribunals apply the equal protection tests supplies two answers. One is the existence of an equal rights amendment; the other is the presence of women on the bench.

1. Equal Rights Amendments (ERA)

Beginning with an ERA, scholars have long argued that the adoption of a federal ERA will force jurists to elevate sex to a suspect class, which in turn will lead them to eradicate virtually all sex-based classifications, as they now do in the case of race.¹⁰⁸

note 64. See also *infra* note 108.

¹⁰⁸ Elevating sex to a suspect class was a primary motivation for the drive for (and against) the ERA in the 1970s. See, e.g., WILLIAM N. ESKRIDGE, JR. & NAN D. HUNTER, *SEXUALITY, GENDER, AND THE LAW* (1997); HERMA H. KAY, *SEX-BASED DISCRIMINATION: TEXT, CASES AND MATERIALS* (2d ed. 1988); Mary E. Becker, *Obscuring the Struggle: Sex Discrimination, Social Security, and Stone, Seidman, Sunstein & Tushnet's Constitutional Law*, 89 COLUM. L. REV. 264 (1989); Brown, *supra* note 28; Mary Anne Case, *Reflections on Constitutionalizing Women's Equality*, 90 CALIF. L. REV. 765 (2002); Ruth Bader Ginsburg, *The Equal Rights Amendment is the Way*, 1 HARV. WOMEN'S L.J. 19 (1978); Kaufman, *supra* note 3; Catharine A. MacKinnon, *Unthinking ERA Thinking*, 54 U. CHI. L. REV. 759 (1987); Kathleen M. Sullivan, *Constitutionalizing Women's Equality*, 90 CALIF. L. REV. 735 (2002); Francis, *supra* note 66; Note, *Sex Discrimination and Equal Protection: Do We Need a Constitutional Amendment*, 84 HARV. L. REV. 1499 (1971). On other hand, some scholars argue that a formal rule, such as an ERA, will not effectively end the subordination of women by men at least in part because of the prevalence of male dominance in most facets of social, political and economic life. We do not attempt to assess this position, but our

Our exploration of the standards used in sex-based discrimination cases supplies some confirmation of the importance of an equal rights amendment, which over one-third of the states have now incorporated into their constitutions—with many containing similar language and purporting to carry analogous objectives as the federal ERA.¹⁰⁹ What we find, after controlling for all other relevant factors,¹¹⁰ is that the presence of an ERA significantly increases the odds of a court adopting a higher equal protection test. Specifically, when we set all other variables at their mean, the likelihood, on average, of a court invoking strict scrutiny to adjudicate a sex-based claim is just 11% in the absence of an ERA. That probability doubles to 23% when an ERA is in effect.

Of course, because this figure of .23 is relatively distant from 1.00, it is far from certain that an ERA will assure the application of strict scrutiny. But it does raise the probability of state jurists taking that step—and it may very well have the same effect on U.S. Supreme Court justices. In fact, in the early 1970s several declined to elevate sex to a suspect class at least in part because they thought it “inappropriate to ‘amend’ the Constitution while the ERA was pending.”¹¹¹

analysis does lend support to the claims of others who argue that formal equality provisions are not always inefficacious, but rather their effectiveness depends a good deal on who is interpreting them. Specifically, to foreshadow our results, we find that as the fraction of women serving on a state supreme court increases, the likelihood of the court adopting a higher standard of law also increases—and significantly increases at that. *See* Part IV.

¹⁰⁹ *See, e.g.,* Susan Crump, *An Overview of the Equal Rights Amendment in Texas*, 11 Hous. L. Rev. 136 (1973); Andrea J. Farone, *The Florida Equal Rights Amendment: Raising the Standard Applied to Gender Under the Equal Protection Clause of the Florida Constitution*, 1 FL. COASTAL L.J. 421 (2000).

¹¹⁰ *See supra* note 91.

¹¹¹ ESKRIDGE & HUNTER, *supra* note 108, at 78; *see also* Ginsburg, *supra* note 108.

The Court's declination came at a time when the ERA's passage looked promising. What about now, some thirty years later? What are the odds of adding an ERA to the U.S. Constitution? Addressing this question is beyond the scope of this article so we will only note here that, despite pronouncements in the 1980s to the contrary, the Equal Rights Amendment (ERA) may not be dead. Actually, there are signs that the battle may be heating up yet again.¹¹² For example, the "three-state" strategy deployed by organized interests in response to claims appearing in scholarly journals, policy memoranda, and the press that ratification of the 27th Amendment in 1992—over 200 years after it was proposed—may hold implications, if not promise, for the ERA.¹¹³ To be sure, this "reconstituted" drive for the ERA has generated substantial opposition (especially from Phyllis Schlafly and her Eagle Forum), but it may very well succeed in Illinois, where in 2003 65% of voters supported ratification and only 19% did not (17% had no opinion).¹¹⁴ Another sign is the increasing importance attached to the Amendment in academic and media treatments. By way of illustration, consider that in the first six months of 1993, just 186 news articles made mention of the ERA. For the same period in 2003, the number of news articles that mentioned the ERA was more than double (N=471) the amount of the corresponding period in 1993.¹¹⁵ Yet a final indication of the rising importance of the ERA comes from legal commentators. A passage from Judge Martha Daughtrey's

¹¹² Epstein et al., *supra* note 27.

¹¹³ See, e.g., Denning & Vile, *supra* note 88. But see Georgia Duerst-Lahti, *Time To Ratify the Equal Rights Amendment*, NEWSDAY, May 20, 2003, at A30; Ellen Goodman, *Equal Rights Amendment is Not Dead Yet*, NEW ORLEANS TIMES-PICAYUNE, Feb. 21, 2000, at B5.

¹¹⁴ Christi Parsons, *Not So Controversial to Voters: Poll Says Many Back Gay Rights, ERA, Gun Limits*, CHI. TRIB., June 17, 2003, at C1.

¹¹⁵ We obtained these figures from a LEXIS search (in the news group file) on the term "Equal Rights Amendment."

Madison Lecture, delivered at New York University, provides but one example:

In the course of cleaning out closets and drawers that had collected much too much stuff over a dozen years, I found this political button, brought home—as I recall—from an ABA meeting some years ago. It reads: “Happy Birthday E.R.A. 1923-1993, You Are Long Overdue!”

About the same time that I found the button, the ABA Journal published a cover story on the renewed efforts to amend the United States Constitution to prohibit discrimination on the basis of gender. As it turns out, the Equal Rights Amendment (ERA) which, if ratified, would have become the twenty-seventh amendment to the Federal Constitution—but which “died” for lack of ratification by three additional states in 1982—has been reintroduced in the current session of Congress. The prospect of a renewed effort to pass the ERA in Congress and to mount ratification campaigns in the fifty state legislatures raises a number of questions that I would like to explore with you this evening.¹¹⁶

Whether the “renewed effort” of which Judge Daughtrey speaks will succeed we cannot say. What does seem to be the case is that the bulk of contemporary commentary now suggests that the ERA may be as dead as the 19th Amendment, which took over 40 years to gain

¹¹⁶ Daughtrey, *supra* note 6, at 2-3 (citations omitted).

ratification.¹¹⁷

2. *Women on the Bench*

Certainly our results indicate that ERAs are important components in the quest for the eradication of sex-based discrimination because they increase the probability of a court applying a higher standard of law to adjudicate claims of sex discrimination. Furthermore, the application of a higher standard of law, even after controlling for other relevant factors, increases the probability of a court reaching a disposition favorable to litigants alleging a violation of their rights.

An ERA is not, however, the only factor that lifts the odds of the adoption of strict scrutiny. As some scholars have long speculated, the proportion of women on the deciding court also exerts a statistically significant effect.¹¹⁸ As that proportion increases, the probability of

¹¹⁷ See Francis, *supra* note 66 (parallels between the campaigns for ratification of the 19th Amendment and the ERA).

¹¹⁸ Virtually from the day Suzanna Sherry penned her classic work *Civic Virtue and the Feminine Voice in Constitutional Adjudication*, 72 VA. L. REV. 543 (1986) on the possibility of a “feminine” jurisprudence, scholars have hotly debated whether female judges “speak in a different voice.” For recent reviews of this literature, see Herma Hill Kay & Geraldine Sparrow, *Workshop on Judging: Does Gender Make a Difference?*, 16 WIS. WOMEN’S L.J. 1 (2001); Daniel M. Schneider, *Empirical Research on Judicial Reasoning: Statutory Interpretation in Federal Tax Cases*, 31 N.M. L. REV. 325 (2001). While the results of various research projects exploring whether male and female judges vote differently reach decidedly mixed results, those centering on jurisprudence (Sherry’s original target)—especially in the area of sex discrimination—are clearer. A consensus now exists that women have “pushed the law forward in sex discrimination cases” with their distinct approach to legal principles possibly altering the choices made by their male colleagues. Kay & Sparrow, at 11. See, e.g., Sherry; Sullivan, *supra* note 106. Our study lends empirical support to this growing consensus. At minimum, it seems rather clear that the presence of women on state judiciaries exerts an influence on how

applying a higher standard of law soars even after controlling for the presence (or absence) of an ERA.

To see the magnitude of the effect, consider a court composed exclusively of male justices. On average, the odds of that court using a rational basis standard, setting all other variables at their mean, is a hefty .50; the probability of that same court applying strict scrutiny is but .12. Now consider a court nearly equally divided between male and female judges: the probabilities nearly reverse. The odds of this court applying rational basis are, on average, but 14% while the probability for strict scrutiny jumps to 47%.

This result commends a rather pointed strategy for those seeking more equality-oriented outcomes in court cases, a campaign designed to bring more women to the federal judiciary. Surely the ultimate target would be the U.S. Supreme Court: with one more favorably disposed justice, a majority supporting strict scrutiny could emerge.¹¹⁹ Of course, history shows that the new justice need not be a woman; after all, it was William J. Brennan who was among the first to urge his colleagues to elevate sex to a suspect class.¹²⁰ On the other hand, it is perhaps no coincidence that it is Justices O'Connor¹²¹ and Ginsburg¹²² who continue to push the *United States v. Virginia* standard—a standard some say is more akin to strict

courts adjudicate sex-based claims.

¹¹⁹ This assumes that the four dissenters in *Nguyen*, 533 U.S. 53 (2001), the Court's most recent sex-discrimination case, who invoked the "exceedingly persuasive justification" of the *VMI* case to support their views, would be willing to elevate sex to a suspect class. Surely this is true of Justice Ginsburg, who may have viewed *VMI* as the first step in that direction. See Skaggs, *supra* note 7; Stobaugh, *supra* note 8; Morris, *supra* note 53. It is not so clear that Justices Breyer, O'Connor, and Souter viewed the *VMI* case in the same way.

¹²⁰ See *Frontiero*, 411 U.S. at 677 (Brennan's judgment for the Court).

¹²¹ See, e.g., *Nguyen*, 533 U.S. at 74 (O'Connor's dissent).

¹²² See, e.g., *United States v. Virginia*, 588 U.S. at 519 (Ginsburg writing the majority opinion of the Court).

scrutiny than it is to *Craig*.¹²³

Nonetheless, however important the U.S. Supreme Court, ignoring the lower appellate bench in any campaign designed to increase the number of female judges would be

¹²³ See *supra* note 8; *supra* note 9. See also Linda Greenhouse, *From the High Court, A Voice Quite Distinctly a Woman's*, N.Y. TIMES, May 26, 1999, at A1. Greenhouse writes that Ginsburg:

recounted in a 1997 speech to the Women's Bar Association . . . that a year earlier, as she announced her opinion declaring unconstitutional the all-male admissions policy at the Virginia Military Institute, she looked across the bench at Justice O'Connor and thought of the legacy they were building together.

Justice Ginsburg's opinion in the Virginia case cited one of Justice O'Connor's earliest majority opinions for the Court, a 1982 decision called *Mississippi University for Women v. Hogan* that declared unconstitutional the exclusion of male students from a state-supported nursing school. Justice O'Connor, warning against using "archaic and stereotypic notions" about the roles of men and women, herself cited in that opinion some of the Supreme Court cases that Ruth Ginsburg, who was not to join the Court for another 11 years, had argued and won as a noted women's rights advocate during the 1970's.

Addressing the women's bar group, Justice Ginsburg noted that the vote in Justice O'Connor's 1982 opinion was 5 to 4, while the vote to strike down men-only admissions in Virginia 14 years later was 7 to 1.

"What occurred in the intervening years in the Court, as elsewhere in society?" Justice Ginsburg asked. The answer, she continued, lay in a line from Shakespeare that Justice O'Connor had recently spoken in the character of Isabel, Queen of France, in a local production of "Henry V": "Haply a woman's voice may do some good."

See also Daughtrey, *supra* note 6, at 21-22.

in error. Since change from the “bottom up” is not unknown in American legal history,¹²⁴ it is always possible that women serving in the circuits could exert “hydraulic pressure” on the Supreme Court, forcing it to reevaluate its current standard.¹²⁵ Yet, at the time of this writing, less

¹²⁴ An interesting example along these lines comes from early legal debates over how to define obscenity. While the U.S. Supreme Court until the 1950s clung to a highly restrictive definition developed in the British case of *Regina v. Hicklin*, L.R. 3 Q.B. 360 (1868), the lower federal courts were liberalizing or even rejecting that definition. Among the most prominent exemplars is Judge Augustus Hand’s opinion in *United States v. One Book Entitled “Ulysses” by James Joyce*, 72 F.2d 705 (1934). See EPSTEIN & WALKER, *supra* note 3, at 359-60.

¹²⁵ In the legal annals, the term “hydraulic pressure” (usually associated with public pressure on the Court) has taken on a negative connotation owing to its use by Justice Holmes in his dissent in *Northern Securities Company v. United States*:

Great cases like hard cases make bad law. For great cases are called great, not by reason of their real importance in shaping the law of the future, but because of some accident of immediate overwhelming interest which appeals to the feelings and distorts the judgment. These immediate interests exercise a kind of hydraulic pressure which makes what previously was clear seem doubtful, and before which even well settled principles of law will bend.

193 U.S. 197, 400-01 (1903). See also *Furman v. Georgia*, 408 U.S. 238, 405 (1972) (Burger, C.J., dissenting) (“The ‘hydraulic pressure[s]’ that Holmes spoke of as being generated by cases of great import have propelled the Court to go beyond the limits of judicial power, while fortunately leaving some room for legislative judgment.”); *Payne v. Tennessee*, 501 U.S. 808, 867 (1991) (Stevens, J. dissenting) (“The great tragedy of the decision, however, is the danger that the “hydraulic pressure” of public opinion that Justice Holmes once described—and that properly influences the deliberations of democratic legislatures”); *Nixon v. Administrator of General Services*, 433 U.S. 425, 505 (Burger, C.J., dissenting) (“Well-settled principles of law are bent today by the Court under that kind of ‘hydraulic pressure.’”). But pressure can, of course, come from sources other than public opinion,

than 25% of the seats on the federal circuit courts are occupied by female judges (32 of 134),¹²⁶ meaning that the odds of attaining a panel with two women, much less three, are rather small— and, for some circuits, border on trivial, as Table 2 indicates. A strategy aimed at increasing these figures, if successful and if our analysis of the state courts transports to the federal judiciary, would likely help, and not impede, the goal of elevating sex to a suspect class.

and it is equally as certain that pressure to change problematic principles of law, even if well-settled, is hardly the tragedy that some of these statements suggest.

¹²⁶ Data are available at

http://www.allianceforjustice.org/judicial/judicial_selection_resources/selection_database/activejudges.asp (last accessed on December 31, 2003).

Table 2: The gender composition of the U.S. Courts of Appeals. Each cell represents the probability of panel composed of a particular combination of male (M) and female (F) judges across the appellate courts, assuming random assignment of three-judge panels.¹²⁷

Circuit (Gender Composition)	Probability of a Panel of:		
	Three Female	Two Females	One Female
1 st (4 M; 1 F)	0.00	0.00	0.60
2 nd (7 M; 2 F)	0.00	0.08	0.50
3 rd (7 M; 4 F)	0.02	0.26	0.51
4 th (9 M; 2 F)	0.00	0.06	0.44
5 th (8 M; 3 F)	0.01	0.15	0.51
6 th (5 M; 5 F)	0.08	0.42	0.42
7 th (8 M; 2 F)	0.00	0.07	0.47
8 th (7 M; 1 F)	0.00	0.00	0.38
9 th (15 M; 5 F)	0.01	0.13	0.46
10 th (8 M; 3 F)	0.01	0.15	0.51
11 th (8 M; 1 F)	0.00	0.00	0.33
D.C. (6 M; 1 F)	0.00	0.00	0.43
Mean	0.01	0.11	0.46

B. Discrimination Against Gays and Lesbians

In an effort to eradicate discrimination against gays and lesbians, commentators have proposed a number of doctrinal avenues.¹²⁸ One, the Due Process Clause, proved

¹²⁷ The data on gender composition are as of January 1, 2003 (derived from http://www.allianceforjustice.org/judicial/judicial_selection_resources/selection_database/activejudges.asp and do not include vacancies. Assuming random assignment of federal appellate judges to panels, we calculated the probabilities in accord with simple probability rules. See generally SHELDON ROSS, A FIRST COURSE IN PROBABILITY 24-63 (6th ed. 2002).

¹²⁸ For reviews, see ESKRIDGE & HUNTER, *supra* note 108; Massaro, *supra* note 14; ANDREW KOPPELMAN, THE GAY RIGHTS QUESTION IN AMERICAN LAW (2002). For a critique of many of these arguments, see

successful in *Lawrence v. Texas*.¹²⁹ That has not, however, diminished the importance of equal protection in the battle to eradicate classifications based on sexual orientation.¹³⁰

Along these lines, scholars have suggested two chief courses of action. Some urge the application of heightened scrutiny to laws that discriminate between homosexuals and heterosexuals, that is, these commentators “seek to garner intermediate scrutiny for gays as gays.”¹³¹ Another group suggests that discrimination against gays and lesbians is, in fact, discrimination based on sex. Hence, courts should apply the same level of scrutiny to classifications based on sexual orientation as they now do for laws that amount to sex discrimination.¹³²

Cass Sunstein, *Homosexuality and the Constitution*, 70 IND. L.J. 1 (1994); *but see* Massaro’s cogent response, *supra* note 14.

¹²⁹ 539 U.S. 558, 123 S.Ct. 2472. *But see* Post, *supra* note 24. Post claims that passages in the opinion are framed in the language of equal protection. Indeed, the Court itself seemed to “meld” equal protection and due process when it wrote that “[e]quality of treatment and the due process right to demand respect for conduct protected by the substantive guarantee of liberty are linked in important respects,” though it added that “a decision on the latter point advances both interests.” 123 S.Ct. at 2482.

¹³⁰ *See* Massaro *supra* note 14; WILLIAM N. ESKRIDGE, JR., *THE CASE FOR SAME-SEX MARRIAGE* (1996).

¹³¹ *See* Yoshino, *supra* note 14; Shalleck, *supra* note 14. *See generally* JOHN HART ELY, *DEMOCRACY AND DISTRUST* (1980). Along similar lines, some scholars and judges have proposed applying strict scrutiny to discrimination based on sexual orientation. *See, e.g.*, Watkins v. U.S. Army, 875 F.2d 699, 711-28 (9th Cir. 1989) (Norris, J., concurring), *cert. denied*, 498 U.S. 957 (1990); Renee Culverhouse & Christine Lewis, *Homosexuality as a Suspect Class*, 34 S. TEX. L. REV. 205 (1993); Note, *The Constitutional Status of Sexual Orientation: Homosexuality as a Suspect Classification*, 98 HARV. L. REV. 1285. Our analysis below speaks to this proposal as well.

¹³² *See* Clausen, *supra* note 14; Koppelman, *supra* note 14; Law, *supra* note 14; Massaro, *supra* note 14; Pharr, *supra* note 14. *See also* ESKRIDGE, *supra* note 130, at 162 (noting that in *Baehr v. Lewin*, the court adopted the argument that “the state’s refusal to give marriage licenses to same-sex couples is sex discrimination . . . [T]he Hawaii

This last argument has some appeal: if courts treat sex as a suspect class and if they place discrimination on the basis of sexual orientation under the rubric of sex discrimination, then the odds are high of eradicating whatever classification is at issue. This is the central message of our study; and it is the lesson of *Baehr v. Lewin*, as well.¹³³ In that case, the Supreme Court of Hawaii treated the denial of same-sex marriages as sex discrimination and applied strict scrutiny—the standard it uses, owing to the presence of state ERA, to assess sex-based classifications. But the argument for treating discrimination against gays and lesbians as sex discrimination—as well as, of course, proposals seeking to “garner intermediate scrutiny for gays as gays”—loses some of its appeal in the current federal context, as well as in most states. In those arenas, sex is not treated as a suspect class, but subject to intermediate scrutiny, which, as we have demonstrated throughout, is far less likely to lead to equality-oriented outcomes.

This demonstration, however, is not meant to suggest that advocates for gay rights should eschew an equal protection strategy designed to attain heightened constitutional scrutiny. Actually, we, along with many legal scholars, see benefits to this approach—some of which are symbolic, but others, quite practical. For one thing, as our data suggest, moving from rational basis to intermediate scrutiny will, in all likelihood, further the cause of gay rights. Indeed, if the results presented here generalize across the judiciary and transport from sex discrimination to sexual orientation, the probability of success in court will double. For another, as Ruth Bader Ginsburg points out, the elevation of sex from rational basis to heightened scrutiny has had salutary effects that

court made the right decision.”).

¹³³ 852 P.2d 44 (Haw. 1993).

transcend the courtroom:

The Supreme Court, since the 1970s, has effectively carried on in the gender discrimination cases a dialogue with the political branches of government. The Court wrote modestly, it put forth no grand philosophy. But by forcing legislative and executive branch re-examination of sex-based classifications, the Court helped to ensure that laws and regulations would catch up with a changed world.¹³⁴

On the other hand, there are costs associated with the strategic pursuit of heightened scrutiny in the name of advancing the legal rights of gays. Primarily, if courts began to apply the mid-level test to classifications based on sexual orientation—either by treating them as a separate class or folding them into sex discrimination—they would, in all likelihood, abide by that standard for the foreseeable future if constitutional sex discrimination litigation is any indication. To see this danger, we need only to recall that the test established in *Craig* nearly 30 years ago remains the test that the Court applies today despite efforts on the part of Justice Ginsburg and others to “ratchet it up.”¹³⁵

That principles of law endure is no great surprise. As we noted earlier,¹³⁶ if courts do not follow previously established rules of law, or do so in unpredictable ways, they risk undermining their fundamental efficacy, for members of legal and political communities base their future expectations on the belief that others will follow

¹³⁴ Ruth B. Ginsburg, *Constitutional Adjudication in the United States as a Means of Advancing the Equal Stature of Men and Women Under the Law*, 26 HOFSTRA L. REV. 263, 270 (1997).

¹³⁵ See *supra* note 6.

¹³⁶ See *supra* note 42.

existing rules. However, this phenomenon does not appear to be the case with regard to *Craig*, which has generated unpredictable results in the sex discrimination area and may very well do the same in constitutional cases centering on sexual orientation. Advocates for gay rights can expect, if our results are any indication, to lose as many cases as they win, with the equal protection test itself providing little guidance to differentiate between the two.

By explicating these problems, we emphasize once again that our goal is not to deter advocates from following the equal protection path (especially in the event of the Court elevating sex to a suspect class). It is rather to persuade members of the legal community to undertake an analysis of its particular costs and benefits—whether with our data or with other, more tailored observations.

V. Conclusion

For nearly thirty years now, the U.S. Supreme Court has employed a “heightened scrutiny” test to adjudicate constitutional claims of sex discrimination. While some commentators endorse this approach, far more have questioned it. Their critiques are varied in message and many in number, but chief among them is the test’s seeming lack of determinacy: because the test is so “amorphous” it fails to establish reliable expectations about the results of sex-discrimination litigation.

Our empirical results put this normative critique on stronger footing. We find that when courts apply the intermediate standard, litigants alleging sex discrimination are nearly as likely to win as they are to lose. This finding is in marked contrast to the relatively predictable outcomes generated when courts apply strict scrutiny, under which most parties challenging sex-based classifications prevail.

For those desiring a larger number of equality-oriented outcomes, the task is to convince courts to elevate

sex to a suspect class. We supplied several strategies for accomplishing this objective, and surely others exist. Until one or more succeeds, however, Ginsburg's cautionary remark of two decades ago remains apt: "variance," and not uniformity, "within the federal judiciary will persist."¹³⁷

¹³⁷ Ginsburg, *supra* note 58.