Michigan Law Review

Volume 76 | Issue 5

1978

Equal Protection: A Closer Look at Closer Scrutiny

Michigan Law Review

Equal Protection: A Closer Look at Closer Scrutiny

TABLE OF CONTENTS

I. INTRODUCTION

The equal protection clause, with its social and analytic complexities, has stirred courts and commentators perhaps more than any other constitutional provision. Conventional wisdom now holds

that the two-tiered equal protection test[1] provides opaque analysis and mechanical results.[2] Different remedies have been proposed. Some commentators attempt to supply a governing, substantive principle for equal protection adjudication.[3] Others, notably Professor Gunther,[4] have advocated strengthening the rational basis test by requiring substantial factual support for a classification and by subjecting the classificatory means to greater scrutiny.[5] However, this second approach has been profoundly shaken by the criticism, expressed in an influential student note,[6] that any means-end rationality requirement can be tautologically satisfied.[7] Consequently, in many eyes, equal protection lacks both a governing principle and a sound method of argument.

This Note proposes to restore means-end analysis to legal respectability through a comprehensive integrated approach to purpose, misfit, and balancing. The search for a rational basis is meaningless if there are no constraints on the kind of purpose which may justify a classification.[8] Therefore, this Note initially explores ways in which a court can more rigorously scrutinize statutory purpose. The next significant question is how a court should evaluate the degree of coincidence between the class picked out by the law and the class which would be picked out if the law were to achieve its goals. Such "misfit" analysis is underdeveloped, largely because the liberality of the "purpose" requirement often renders it unnecessary. A major part of this Note analyzes misfit (or as it is commonly described, overinclusion and underinclusion).

After a legitimate equal protection purpose has been identified and the misfit it creates analyzed, the classification must be subjected to a balancing test. This Note describes the dynamics of that balance, especially the role of misfit and the relevance of reasonable alternatives. The difference between the substantive due process and equal protection rationality tests is also explained.

---

1. A helpful general exposition of the test can be found in *Developments in the Law—Equal Protection*, 82 HARV. L. REV. 1065, 1076-132 (1969) [hereinafter cited as *Developments*].

2. *See* Gunther, *The Supreme Court, 1971 Term—Foreword: In Search of Evolving Doctrine on a Changing Court: A Model for a Newer Equal Protection*, 86 HARV. L. REV. 1, 8 (1972).

3. *See, e.g.*, Nowak, *Realigning the Standards of Review Under the Equal Protection Guarantee—Prohibited, Neutral, and Permissive Classifications*, 62 GEO. L.J. 1071 (1974); Wilkinson, *The Supreme Court, the Equal Protection Clause, and the Three Faces of Constitutional Equality*, 61 VA. L. REV. 945 (1975); Karst, *The Supreme Court, 1976 Term—Foreword: Equal Citizenship Under the Fourteenth Amendment*, 91 HARV. L. REV. 1 (1977).

4. Gunther, *supra* note 2.

5. *See* Nowak, *supra* note 3, at 1094; Simson, *A Method for Analyzing Discriminatory Effects Under the Equal Protection Clause*, 29 STAN. L. REV. 663, 665-66 (1977).

6. Note, *Legislative Purpose, Rationality, and Equal Protection*, 82 YALE L.J. 123 (1972).

7. For another expression of this view, see Sandalow, *Racial Preferences in Higher Education: Political Responsibility and the Judicial Role*, 42 U. CHI. L. REV. 653, 654-63 (1975).

8. *See* Note, *supra* note 6.

No simple formula can resolve equal protection claims, and this Note does not suggest one. Nor is this Note, properly speaking, an apology for closer scrutiny. Rather, it attempts to describe a mode of argument which should be used in equal protection cases, a framework general enough[9] to accommodate a broad range of substantive views about the material values which should guide equal protection analysis.

The argument in this Note is long and complex, and by its nature impossible to summarize in a neat phrase such as "rational basis," "purpose capable of withstanding analysis,"[10] or "means-focused scrutiny."[11] Nevertheless, a general exposition of the model may be helpful. Much of the following outline may be somewhat obscure at this stage, but it should serve as a useful point of reference whenever the reader is unsure of the argument's direction.

Part I.A introduces the main themes of the equal protection clause. Equal protection proscribes both "hostile" and "unfair" discriminations. A hostile discrimination is a law which is intended to burden or benefit a group because of an obviously irrelevant characteristic and which produces stigma. This Note deals exclusively with unfair discriminations, namely, laws whose consequences are unjust in light of the characteristics of the classes created and of the purposes of the law. These two different prohibitions respond to different kinds of injustice and require separate analysis. Thus, the prohibition against "unfairness" does not require a court to examine legislative motive, but to evaluate the effect of a law on the interests of those classified.

The fairness of a law can be judged, as a first approximation, by asking whether the class picked out adequately matches the class that should have been picked out, given the law's purpose (or, to use the current expression, by whether "misfit" is small enough). Some constraints on "purpose" are necessary, however, if the law's rationality is to be judged. Otherwise the means—picking out class A—is always rationally (but tautologically) related to the "end"—picking out class A.

Part III's discussion of purpose demonstrates that misfit analysis is feasible despite the tautology objection. To refute that objection, we must distinguish between goals which are legitimate, impermissible, or neither. A legitimate goal, as specially defined in this Note, is one which so clearly differs in kind from the most nearly comparable goal that the legislature is entirely free to pursue it. That is, the

9. The Note's analysis does not extend to the problems of unconstitutional motivation and the nature of the classifying trait (*i.e.*, the problem of "suspect" classes), for reasons suggested in Part I.A *infra*.

10. Nowak, *supra* note 3, at 1094.

11. Gunther, *supra* note 2, at 20-24.

state need not justify such a goal in terms of its rational relationship to some *further* end. Every classification should be judged according to a legitimate goal, and not simply a *possibly* legitimate goal. This constraint does much to avoid the tautology trap, for the state must ordinarily justify a law by a somewhat general purpose, not by a narrow, tautological interest in burdening those whom the law in fact burdens. In a few cases, however, which we may call "self-justifying" classifications, the tautological justification suffices.

Impermissible goals are those goals which cannot be invoked at all to justify a law, however "rationally" they are served in any given case. It is difficult to establish clear standards for "illegitimacy," but for several reasons later noted, this difficulty may not seriously jeopardize equal protection review. The last, catch-all category comprises those goals which may be invoked only tentatively to justify a law. Even if they are quite "rationally" served in a given case, they do not conclusively establish the fairness of the law, since they might not be rationally related to any purpose which *is* clearly legitimate.

Once a legitimate purpose has been isolated, we can meaningfully analyze the "misfit" (that is, overinclusion and underinclusion) created with respect to that purpose. Part IV describes in some detail the requirements of misfit analysis, since courts and commentators have badly neglected this topic. The most basic equal protection requirement is the minimum rationality constraint, namely, that the burdened (or benefitted) class pose a higher social harm (or need), be less costly to apply the burden (or benefit) to, or have less of an interest in avoiding the burden (or obtaining the benefit) than a class not picked out for special treatment. These comparisons refer to the average (not the aggregate) harm or cost or interest for each class. In other words, this weakest rationality constraint prohibits the state from striking at a large part of the harm, unless, in that part, the harm is more acute.

Properly understood, this constraint helps dispel several misconceptions about misfit analysis. By emphasizing *differential* harm, it avoids the popular illusion that the *absolute* amount of misfit is the measure of a law's irrationality. The constraint also illustrates that administrative convenience is a basic factor in the analysis of *any* law, not just another goal which the state may or may not be pursuing. In the sex discrimination cases especially, the Court has improperly distinguished between presumptions whose "only" justification is administrative cost and those which are "rational." This distinction is false because an imperfect but reasonable generalization can be described either as "rationally" distinguishing classes of unequal average harm, *or* as distinguishing classes of unequal cost but equivalent harm, for it is simply more expensive to "find" the harmful individuals in the class posing a lower average harm.

The elements of the minimum rationality constraint merit further discussion, since they are also the factors ultimately balanced in a "closer scrutiny" test. Harm can be described as continuous, intermediate or discrete.[12] However, the intermediate characterization, that only those meeting a threshold level of harm deserve a burden, can be an artificial device for evading misfit review. Cost can be separated into (1) the cost of identifying to whom the burden shall apply and (2) the cost of applying the burden.[13] The discussion will indicate that even the second type of cost has equal protection relevance, despite initial appearances to the contrary; and that "individualized hearings" are *not*, again despite appearances, the costly but "perfectly rational" solution to classificatory inaccuracy. Finally, the burdened and the unburdened will sometimes have a different personal interest in avoiding the burden.[14] Of course, the personal interest most often has equal protection significance simply because a law differentially infringes it, not because different classes value it differently.

A major failing of contemporary misfit analysis is its inability to explain the difference between the rationality requirements of substantive due process and equal protection. Part IV.C.1 demonstrates that due process analyzes the overbreadth, *i.e.*, overinclusiveness, of a law, while equal protection analyzes the degree to which the burdened and unburdened differ in posing a harm. The two methods are independent of each other—a due process overbreadth violation is not necessarily an equal protection violation, and vice versa—but they may sometimes be profitably read together to invalidate a law which does not quite violate either test. (Procedural due process also involves some "misfit" analysis of classificatory accuracy, and can be integrated with substantive due process/equal protection in a generalized test.)

The distinction between substantive due process and equal protection suggests an answer to the question whether overinclusion and underinclusion should be analyzed differently by a court.[15] That answer is a complicated "yes." Neither overinclusion nor underinclusion is directly relevant to equal protection. Each is *indirectly* relevant insofar as it affects the size of the differential in harm that the classes of the burdened and the unburdened threaten. Overinclusion is, however, directly relevant to substantive due process analysis, and therefore we might say that courts should be somewhat more tolerant of underinclusion than of overinclusion. (This tolerance is a

12. *See* Part IV.B.1 *infra*.
13. *See* Part IV.B.2 *infra*.
14. *See* Part IV.B.3 *infra*.
15. *See* Part IV.C.2 *infra*.

sounder formulation of the view that the state may attack "part of a harm.")

The previous analysis is valid whether overinclusion and underinclusion are defined with respect to an affirmative burden or with respect to the "burden" of denying a benefit—with one important exception. Unlike equal protection, due process cannot treat benefits and burdens uniformly, because denial of a state benefit in which the plaintiff does not yet have a property interest is not a "deprivation of . . . property" under the due process clause. But this lack of uniformity is remedied somewhat by the irrebuttable presumption doctrine. That doctrine seems to be essentially a mild substantive due process (rather than equal protection) test and does *not* seem to require that the plaintiff have an actual property interest in the benefit which an irrebuttable presumption has denied him.

The last topic under misfit analysis asks how a court should evaluate laws with multiple goals.[16] When the state classifies persons who more acutely pose *each* of the multiple harms, the analysis is simple: a court compares the burdened class with the class which poses multiple harms and evaluates misfit in the usual way. But in other cases, a court must justify its decision to burden those posing multiple harms and not those threatening only one of the harms, since otherwise the law would violate the minimum rationality constraint. A court might reason that the possession of one harm implies a higher degree of another; that pursuit of one goal inevitably aids or inhibits pursuit of a second; that a single legitimate purpose can be inferred from the several purposes; or that the classification is part of a general policy of attacking one of the harms.

Misfit analysis is impossible when a court cannot reasonably ascertain the rough dimensions of the "purpose" class, *i.e.*, those whom the law should have picked out given its purpose.[17] Two broad categories of goals are inevitably arbitrary in this sense and thus are exceptions to misfit analysis. First, some goals have inherently imprecise standards (*e.g.*, aesthetic and moral judgments). Second, some goals are "neutral," for they are indifferent to *who* is picked out; the only purpose is to burden or benefit a limited class of persons indistinguishable from a larger class (*e.g.*, lotteries, or attempts to reduce total program cost). Analysis of neutral goals is somewhat complex. Such goals are most sustainable if they further a substantive purpose (other than the mere numerical diminution of a larger class) which would *not* be served by rationally excluding some members of the class. (An example is a bona fide social experiment.) Absent such a purpose, the state usually must employ a rational criterion of exclusion and must *proportionally* diminish a benefit or

16. *See* Part IV.D *infra*.

17. *See* Part V *infra*.

burden which is not discrete. If the state may pursue a neutral goal, it should use as random a process as is practical to reduce the class, especially if the deprivation imposed is significant or the classificatory trait would stigmatize those classified.

Once a court has identified a legitimate purpose and evaluated the misfit it creates, it should subject the law to a balancing test.[18] Unfortunately, the present tripartite equal protection test's characterization of balancing is sketchy and possibly misleading. The court should initially ponder whether the differential in harm between the burdened and the unburdened is great enough to justify the differential treatment in light of the significance of the personal and governmental interests, the severity of the burden, and the extent to which the governmental interest is served. A court might be able to decide this question immediately if the law is obviously either prima facie fair or unfair. If the law is not at one of those extremes, then the court should proceed to examine "reasonable," less restrictive alternatives, where reasonableness is judged according to the prima facie fairness of the law, the social costs of the alternative, and how much less restrictive the alternative is. In some cases, even perfect or near-perfect means-end fit will not preserve a law, if the personal interest infringed is so important that only an unusually strong state interest can justify it. However, this more radical balancing approach should be used cautiously, for it often approximates the substantive due process test.

Due process balancing is similar in some respects to equal protection balancing. It weighs the same elements as equal protection, but is seeking to justify overbreadth, a different kind of inequality than the differential harm which equal protection seeks to justify. Again, the court makes a prima facie judgment about overbreadth and examines less restrictive alternatives if the overbreadth is neither obviously fair nor obviously unfair. Due process more often entails the more radical form of balancing in which perfect fit is not enough, and misfit analysis is generally less important under due process than under equal protection.

A. *The Function of the Equal Protection Clause*

Any analysis of the equal protection clause makes assumptions about that clause's proper function. This section reports several uncontroversial but not unimportant assumptions which underlie the Note's analysis.

One traditional view of the predominant theme of the equal protection clause is that it protects "discrete and insular minorities"[19]

18. *See* Part VI *infra.*

19. This phrase first appeared in United States v. Carolene Prods. Co., 304 U.S. 144, 152 n.4 (1938).

from oppression in the political process.[20] But this view fails to explain why some losers in that process deserve judicial protection and others do not. Logically the courts should intervene whenever an important relevant interest has not been voiced in the legislative process, whether or not that interest can be ascribed to an "insular minority." Surely courts cannot and should not fill all the lacunae of the legislative process.

A court enforcing the equal protection guarantee is not a surrogate legislator. It should however shield people (usually, though not invariably, members of minority groups) from hostile discriminations and discriminations which treat them unfairly.[21] This dual protection does not require the courts to re-evaluate the costs and benefits already evaluated by the legislature. It does require courts to decide whether hostility motivated discriminatory treatment of a group and whether a class has been treated unfairly in light of its characteristics and the purposes of the law. Judicial interest-balancing is legitimate only insofar as it helps answer these questions. While these questions are difficult, courts in equal protection analysis need not balance as often or as thoroughly as is implied by the traditional argument that the judiciary should bolster the political strength of minorities.

The notions of "hostility" and "unfairness" deserve comment. "Hostile discrimination" describes a law which would not have been passed but for the legislature's desire to help or hurt a class of persons because that class possesses an obviously irrelevant characteristic.[22] For example, we "suspect" that racial classifications are hostile discriminations because our national experience teaches that they are probably motivated solely by a desire to penalize people on the basis of their skin color, a characteristic generally irrelevant to legitimate legislative purposes.[23]

---

20. *See* Ely, *The Constitutionality of Reverse Racial Discrimination*, 41 U. CHI. L. REV. 723 (1974). Sandalow, *Judicial Protection of Minorities*, 75 MICH. L. REV. 1162 (1977), offers a sophisticated criticism of this traditional view. As Sandalow notes, the view that minorities merit special judicial solicitude is not confined to cases interpreting the equal protection clause. (United States v. Carolene Prods. Co., 304 U.S. 144 (1938), itself was a due process case.)

21. "Hostile" discriminations are analyzed by the Court under the rubric of "suspect" traits. "Unfair" discriminations include classifications burdening "fundamental interests" or lacking a "rational basis." The hostile/unfair distinction, while useful, does not precisely describe the Court's practice. *See generally* text at note 27 *infra*.

22. This definition ignores several difficult questions. Does "obviously irrelevant" mean irrelevant to *any* legitimate purpose (in any context), or simply irrelevant to any conceivable purpose of the challenged law? Must the legislators *know* that the characteristic is obviously irrelevant (*i.e.*, is it a defense that they believe blacks, women, or illegitimates are inferior or unqualified)? How obvious must the irrelevancy be? If a characteristic *may* be relevant, and if the legislators intend to burden those possessing it *regardless* of its relevance, is the discrimination "hostile"? Assume, for example, that a legislature requires women (but not men) to prove their fitness for physically taxing jobs simply to discourage women from unseemly, masculine tasks. Does the fact that the requirement is a barely rational generalization about relative fitness mean the discrimination is not hostile?

23. Race may be a relevant characteristic but still be a forbidden legislative criterion sim-

If the irrelevant characteristic appears on the face of the statute, the illicit motivation can perhaps be presumed, although the presumption can be rebutted, especially if the facial use of the characteristic is not stigmatizing. A classification by race will almost invariably fail because of the stigma such classifications create and sustain.[24] A classification by sex will more commonly be upheld, partly because sex is not so often "obviously irrelevant" as race, and partly because classification by gender is thought less stigmatic.[25] Where hostility cannot be read on the face of the law, an examination of legislative motivation may be required. Stigma will again be relevant. Divining the legislative motive will, of course, be more difficult.

"Unfair" discrimination need not result from hostile legislative motivation. It arises when the characteristics of a statutorily created class are insufficiently related to the purposes of the law. Such discrimination constitutes the conceptually difficult portion of equal protection analysis. Determining whether discrimination is hostile requires proving legislative intent, to be sure, but once we identify classifications which are "suspect"—*i.e.*, probably irrelevant to legitimate state ends—the legal standard is straightforward. Identifying unfair discrimination, on the other hand, necessitates a complex analysis of the legitimacy of state ends, the relevance of a classification to those ends, and the significance of the amount of "misfit"

---

ply because it stigmatizes. Consider Professor Brest's example of an employer who preferred whites to blacks because blacks had a higher absenteeism rate: even if all other rules for predicting absenteeism were less accurate than a racial criterion, such a criterion would certainly be invalidated. *See* Brest, *The Supreme Court, 1975 Term—Foreword: In Defense of the Antidiscrimination Principle*, 90 HARV. L. REV. 1, 6 (1976); Posner, *The Defunis Case and the Constitutionality of Preferential Treatment of Racial Minorities*, 1974 SUP. CT. REV. 1, 21.

The constitutionality of reverse discrimination in some circumstances, *see* Regents of the Univ. of Cal. v. Bakke, 98 S. Ct. 2733, 2738, 2766 (1978) (opinions of Powell & Brennan, JJ.), may seem to contradict the statement in the text. However, upon analysis, the contradiction disappears. Race is "relevant" to compensating minorities for past discrimination; it is a separate question whether this purpose is legitimate. Although the "rationality" of a statute is not conclusive, there may be reasons (including the significance of stigma) for being more "suspicious" of a rational classification disadvantaging blacks than of a rational classification disadvantaging whites. *See* 98 S. Ct. at 2783-85 (opinion of Brennan, J.).

24. Affirmative action programs, as was remarked in note 23 *supra*, may be consistent with the stigmatic effects analysis. That analysis does, however, provide some support for a constitutional preference for "goals" rather than "quotas," since whites might find goals less stigmatizing. *See* Regents of the Univ. of Cal. v. Bakke, 98 S. Ct. 2733, 2751 n.34 (1978) (opinion of Powell, J.).

25. Significant stigmatic harm might be a prerequisite to a finding of "hostile" discrimination. Consider, for example, United States Dept. of Agriculture v. Moreno, 413 U.S. 528 (1973), in which the Supreme Court conceded that the probable purpose of the statute (to deny food stamps to "hippies") was illegitimate, yet invalidated the law on a different ground. The Court may have believed the stigmatic harm was not great, since "hippies" are not a traditionally disfavored minority. (Of course, one may dispute the correctness of such reasoning, which implies that if a class is only temporarily disfavored, it is not suspect.)

between the classification and the ends in light of the personal interests impaired and the governmental interests furthered by the statute. Most laws are at worst "unfair"—they classify people according to characteristics which are not *obviously* irrelevant to legislative ends. This Note presents a framework for analyzing the fairness of statutes; it addresses only indirectly the problem of hostile discrimination.[26]

To summarize, the equal protection clause performs two important functions: it prohibits hostile discrimination and unfair discrimination. The former prohibition essentially concerns stigmatic effects, though it may also guarantee purity of motive for its own sake.[27] The prohibition against unfairness is less concerned with motive or stigma and more concerned with classificatory inaccuracy and with the law's impact on the interests of those it classifies.

These two functions of equal protection are not, of course, mutually exclusive. Classifications by gender exemplify laws whose validity is judged under both standards. Indeed, most sex classifications which have been found unconstitutional would probably withstand scrutiny under either standard alone, but stigma and classificatory inaccuracy together justify their invalidation.[28] In general, any classification which barely satisfies the fairness test developed in subsequent sections might nevertheless violate the equal protection clause if it contains enough indicia of hostile discrimination. Because the prohibitions against hostile and unfair discriminations serve distinct policies, however, one cannot simply add stigma or any other feature of hostile discrimination to the equal protection balance later described in this Note.

B. *Introduction to Misfit Analysis*

Under one formulation of equal protection fairness, a legislative classification "must rest upon some ground of difference having a fair and substantial relation to the object of the legislation, so that all

26. For an illuminating discussion, see Brest, *supra* note 23. Useful general discussions of legislative motive include Brest, *Palmer v. Thompson: An Approach to the Problem of Unconstitutional Legislative Motive*, 1971 SUP. CT. REV. 95; Eisenberg, *Disproportionate Impact and Illicit Motive: Theories of Constitutional Adjudication*, 52 N.Y.U.L. REV. 36 (1977); Ely, *Legislative and Administrative Motivation in Constitutional Law*, 79 YALE L.J. 1205 (1970).

27. *See* Tussman & tenBroek, *The Equal Protection of the Laws*, 37 CALIF. L. REV. 341, 358 (1949).

28. For an early analysis suggesting that the Supreme Court evaluates both (what I have called) hostility and unfairness, see Cox, *The Supreme Court, 1965 Term—Foreword: Constitutional Adjudication and the Promotion of Human Rights*, 80 HARV. L. REV. 91, 95 (1966): "The [new equal protection] decisions appear to rest upon two largely subjective judgments . . . . One element is the relative invidiousness of the particular differentiation . . . . The second element is the relative importance of the subject with respect to which equality is sought. . . ." *Developments, supra* note 1, at 1120-21, suggests that the Court's judgment is determined by the interaction of these two elements.

persons similarly circumstanced shall be treated alike."[29] As this suggests, equal protection requires courts to identify the criterion of similarity:[30] with respect to *what* common feature must people be treated equally?[31] In the context of the equal protection clause, the relevant feature is the purpose of the law.[32] More precisely, two individuals must be treated equally if they share some quality and if the purpose[33] of the law is to benefit or burden individuals who possess that quality.[34] For example, a trespassing law which applied only to the left-handed would probably violate equal protection, for right-handed as well as left-handed trespassers violate owners' property rights, and it is the law's apparent purpose to prevent that harm.

Legislative classifications simplify the process of determining whether individuals possess this relevant quality or feature by stipulating that an easily identifiable trait shall be the criterion for apply-

29. Royster Guano Co. v. Virginia, 253 U.S. 412, 415 (1920). This formulation of the test is less deferential to the state than some other formulations. *Cf.* McGowan v. Maryland, 366 U.S. 420, 425 (1961) (equal protection "is offended only if the classification rests on grounds wholly irrelevant to the achievement of the State's objective"). This Note will argue, however, that if the rational basis test is to have *any* application, some analysis of the "substantiality" of the means-end relationship is necessary. The *Royster* test is employed in the text to anticipate this analysis. Concededly, that test begs the question at this early stage of the argument.

30. *See* Tussman & tenBroek, *supra* note 27. In addition, see Benn, *Equality, Moral and Social*, III ENCYCLOPEDIA OF PHILOSOPHY (1967); Sandalow, *supra* note 7, at 654-63.

31. The requirement is usually stated as demanding *similar* treatment for the similarly situated, which implies that somewhat unequal treatment might suffice. This is misleading, for if we could determine easily which individuals are the *same* with respect to the attribute that the law is intended to burden or benefit, then those individuals should be treated equally. If, however, "similar treatment" means treatment in *proportion* to possession of the relevant trait, the "similar" formulation states the requirement properly.

Courts that use the "similar" formulation probably mean to suggest that the legislature need not create mathematically precise classifications. However, it would be more accurate to say that those who are in the same position with respect to the purpose of a law should be treated equally, but that the necessity of line-drawing and the impossibility of precisely identifying the "purpose" class justify some imprecision.

See Bedau, *Egalitarianism and the Idea of Equality*, in EQUALITY 3 (Pennock & Chapman eds. 1967).

32. McLaughlin v. Florida, 379 U.S. 184, 191 (1964); Tussman & tenBroek, *supra* note 27, at 346. *But see* text at and following notes 75-77 *infra* for discussion of the claim that the fundamental interest branch of active equal protection review substitutes "exercise of a fundamental interest" for "the purpose of the law" as the criterion of similarity.

33. This is not to say, however, that a simple desire to discriminate is a permissible purpose. *See* Part III.A.2 *infra*.

34. The principle that the similarly situated must be treated equally entails the corollary that the differently situated must be treated differently (unless, of course, they are similarly situated in some other respect that would justify the same treatment). If *A* and *B* are differently situated, then by definition *A* has some quality which *B* lacks by virtue of which he deserves to be benefitted (or burdened). But it is improper also to benefit *B*, who lacks the quality and thus does not deserve the benefit. Of course *B* might deserve the benefit by virtue of some *other* quality.

This corollary is not merely of intellectual interest. If a statute applies to *all* persons but should only apply to a few, those unfairly included cannot point to another class, the "unclassified," to whom they are similarly situated. The corollary principle is simply the doctrine of due process overbreadth. A fuller discussion of the difference between the equal protection "principle" and its due process "corollary" is postponed until Part IV.C.1 *infra*.

ing the burden or benefit. In the classic formulation of Tussman and tenBroek,[35] we may define class T as consisting of all individuals possessing the legislatively declared trait, and class M as consisting of all individuals possessing the mischief which the law seeks to eliminate.[36] As thus defined, M is the class of those similarly situated with respect to the purpose of the law. Under this formulation, the essential question for analysis of equal protection rationality or "misfit" is whether the classes M and T sufficiently coincide.

Assume, for example, that a statute conditions the privilege of driving upon passing a driving test. The mischief aimed at is dangerous driving, and the trait identified is failure to pass the test. The "fit" between M and T is doubly imperfect. First, some dangerous drivers will pass the test—thus, the class T is "underinclusive," for it does not include all members of M. Second, some safe drivers will fail the test—thus, T is also "overinclusive," for it includes persons who are not members of M.[37]

The Court has developed several measures of how much statutory misfit will be tolerated. Under even the most tolerant standard, a classification must at least yield "some" fit. Thus, if M and T have no members in common—*i.e.*, if no one who fails the test is a dangerous driver—then the classification must be rejected as completely irrational. On the other hand, the rational basis standard does not demand "perfect rationality"—*i.e.*, that T and M be equivalent and thus that all and only dangerous drivers fail the test. Consequently, if the misfit is not too great, the classification created by the driving test statute would probably be sustained.[38]

Since misfit measures the extent to which M and T do not coincide, misfit analysis is unnecessary when they coincide perfectly. Of course, M and T can be made to coincide by defining T in terms of M or defining M in terms of T. Neither device, however, is usually acceptable. If the legislature used the first device and simply provided that dangerous drivers be denied licenses, we would still be left

35. *See* Tussman & tenBroek, *supra* note 27, at 346-48.

36. "Mischief" is used by Tussman and tenBroek. They did not coin an analogous term for the quality by virtue of which a person deserves a benefit rather than a burden, other than to note that the "purpose of a law may be either the elimination of a public 'mischief' or the achievement of some positive public good." *Id.* at 346. This Note uses the term "need" for this quality of deserving a benefit, even though the term is somewhat misleading insofar as it suggests that only deprivations, and not other personal "merits" such as excellence and importance to the economy, deserve benefits.

37. The following terminology may be helpful. "Properly included" will refer to the class that should be burdened or benefitted and that in fact is burdened or benefitted— *i.e.*, the intersection of M and T: ($M \cap T$). "Improperly included" means overincluded: ($\sim M \cap T$). "Improperly excluded" means underincluded: ($M \cap \sim T$). "Properly excluded" means correctly excluded from the burden or benefit: ($\sim M \cap \sim T$).

38. This statement is a gross oversimplification, as the remainder of this Note illustrates. But it is true that courts generally accept "small" amounts of misfit and that such misfit is often justifiable (though assuredly *not* simply because it is small). *See* text at notes 183-85 *infra*.

to devise a method of determining who is dangerous. If the legislature granted an administrator the authority to identify dangerous drivers, we would have an ascertainable class T—the class of individuals adjudged dangerous by the administrator—but no assurance of a perfect fit. Of course, the more clearly the standards for the administrator's decision are specified, the more closely they resemble a legislative classification based simply on the possession of certain traits, and we are back where we started. In short, insofar as this device produces perfect fit, it entirely undercuts the purpose of legislative classifications—namely, to simplify the process of identifying those who deserve the benefit or burden.

Perfect fit may also be achieved by defining M in terms of T. A court might define the law's purpose as denying licenses to individuals who fail driving tests, not as keeping dangerous drivers off the road. One is inclined to ask why such a purpose should be given credence; the answer must be that such individuals are likely to be dangerous drivers. This method of eliminating misfit seems transparently improper, but it does show that without an independent constraint on judicial definitions of statutory purpose, any standard for classificatory accuracy can be tautologically satisfied. Fortunately, as we will see, such a constraint exists.[39]

The foregoing discussion of misfit is greatly simplified. For instance, we shall see that it is fundamentally misconceived to treat misfit analysis as simply a question of the degree to which the "purpose" and "trait" classes (M and T) coincide. Rather, the proper focus is the difference in the degree to which classes T and $\sim$T pose the harm, where $\sim$T is the class of persons lacking the classificatory trait.

Let us complete this introduction to misfit analysis by glancing at the typical mode of argument in equal protection cases. Ordinarily, the plaintiff complains that he is burdened even though his situation is similar to that of others who are not burdened.[40] Instead of asking whether the asserted misfit is justifiable, courts usually respond by finding a respect in which plaintiff is *differently* situated from those allegedly similarly situated. Where the plaintiff claims to belong to a class with a low enough level of harm that the purpose of the law does not require imposition of the statutory burden, the court usually finds that the plaintiff belongs to another relevant class with respect to which it *is* proper to impose the burden. Of course, in finding "another relevant class," the court has simply interpreted the purpose of the classification in a manner unfavorable to the plaintiff.

An example will illustrate this process. In *McGinnis v. Royster*,[41]

39. *See* Part III.A.1 *infra*.

40. Alternatively, the plaintiff may complain that he is not benefitted even though his situation is similar to that of others who are benefitted.

41. 410 U.S. 263 (1973).

a state statute granted "good time" credit against prison sentences in proportion to the amount of time the prisoner had been incarcerated since sentencing. Although time served before sentencing was credited to the *sentence* imposed, it was *not* included in the computation of good time.[42] Thus, a defendant unable to make bail ultimately served longer than a bailed defendant (since the former could not receive good time credit for pretrial incarceration). The pretrial inmates who challenged the statute implicitly claimed that its purpose was to award credit based on time served, that with respect to this purpose, they resembled inmates who had made bail, and that therefore they also deserved credit according to total time served.[43]

The Supreme Court rejected this equal protection challenge by interpreting the statute's purpose[44] as giving credit for incarceration time during which rehabilitative programs were offered.[45] Since jails are less likely than prisons to offer such programs, the Court reasoned, it was rational to treat prisoners who had spent pretrial time in jail differently than persons incarcerated only in a prison. Note that the Court did not justify the original classification by adopting the plaintiffs' interpretation of the statute's purpose and accepting that misfit as tolerable. Rather, the Court defined the purpose so as to explain the difference in treatment of which the plaintiffs had complained.

The Court's analysis in *McGinnis* highlights a significant distinction between "specific" and "general" ways of justifying misfit. By a "specific justification" a court may explain away the misfit entirely as to the complaining party. Thus, if the party complains that he has been improperly overincluded (with respect to a burden) or underincluded (with respect to a benefit), the court may identify a legitimate purpose as to which this specific party is properly classified. In *McGinnis*, for example, the Court "specifically" justified the differ-

---

42. The actual credit system was somewhat more complex. For example, the credit could be forfeited for "bad conduct" or for violation of prison rules, and in no case could the credit win release from prison before a year. "Good time" earned affected the minimum and maximum parole dates somewhat differently, and pretrial time *was* credited in computing the latter. 410 U.S. at 264-68.

The plaintiffs also complained that the statute's pretrial-time exclusion was not applied to county penitentiary inmates (nonfelons with less than one-year sentences). The Court rejected this equal protection complaint on the ground that these inmates might have required "quantitatively and qualitatively less rehabilitation" than felons. 410 U.S. at 274.

43. More precisely, the plaintiffs complained that the classification discriminated against those unable to afford or to qualify for bail. 410 U.S. at 265, 268. That is, since pretrial jail time *was* credited to the *sentence* imposed, a bailed defendant differed from a jailed defendant only in that he acquired more good-time credit.

44. The Court was careful to note that this was only one of the purposes of the statute and that equal protection analysis does not require that the statute be rationally related only to a "primary" purpose. 410 U.S. at 275-77.

45. The Court added that state prison officials could not observe and evaluate the performance of a jail inmate. 410 U.S. at 273.

ential treatment of the plaintiffs by identifying one purpose of the statute as crediting time served where rehabilitative programs were offered.

The second alternative available to a court faced with misfit is to justify it "generally." In a "general" justification of misfit, the court concedes that the complaining party has been burdened without possessing the corresponding need, but the court finds grounds which justify the misfit. In *McGinnis*, for example, the plaintiffs might have responded to the Court by arguing that they did make rehabilitative progress in jail and therefore were improperly denied the benefit of good time credit. The Court might have replied by generally justifying the misfit: jail inmates are less likely to achieve such progress than prison inmates (who are offered special programs), and it is unreasonable to require the state to evaluate each prisoner's progress, even if such an inquiry would show that the *plaintiffs* had progressed.

Unlike a general justification, which leaves misfit unexplained, a specific justification seems simply to eliminate misfit. This appearance is deceiving, however. Although a specific justification asserts a purpose with respect to which the complainant is properly included, it may create misfit as to other individuals. Specific justifications may offer an intuitively satisfying reason for burdening or not benefitting the complainant before the court, but the misfit they create still must be justified "generally." Otherwise the state could invariably justify its laws simply by finding *any* purpose rationally furthered by burdening the immediate plaintiff. For example, suppose that all the plaintiffs in *McGinnis* had threatened witnesses. The statute surely could not have been justified by the state's interest in denying good time credit to those who have obstructed justice; the plaintiffs would have been almost the only pretrial inmates properly included with respect to that interest. The requirements for general justifications of misfit are discussed in detail later.[46] Nevertheless, it should be emphasized here that a court's use of a specific rather than general justification should be irrelevant to the success of an equal protection claim.[47] Whether misfit can be specifically justified, that is, whether the *plaintiff* deserves the benefit or exemption, should be irrelevant because even if he does not, the plaintiff should have both standing to raise the equal protection claim and a right to a remedy if he succeeds on the merits. The following discussion will explain

46. *See* Part IV *infra*.

47. The state may not justify its treatment of a class by reference to its "proper" treatment of only a part of that class; by the same token, however, the state need not "properly" treat *all* members of a class to justify distinguishing that class from another. In both the former case (special justification) and the latter case (general justification), a *class* differential must be explained.

why equal protection standing and remedy doctrines have this fortunate consequence.

Equal protection standing doctrine seems to produce some curious results. Why can a confessed thief complain of discriminatory enforcement of the criminal laws?[48] Why can a widower who admittedly cannot prove he was dependent on his deceased spouse successfully demand benefits when an obvious purpose of the law was to distribute benefits according to dependency?[49] The answer, in each case, is that the plaintiff suffered legal injury simply in being treated differently than others who are similarly situated. Discrimination claims involve more than an individual's relationship to the government; they involve the fairness of disparate treatment of classes of individuals.[50] The essence of an equal protection claim is not that the plaintiff "deserves" a benefit, but rather that he belongs to a class that is situated similarly and therefore should be treated similarly to the benefitted class.

This "relative desert" analysis allows plaintiffs to avoid issues of standing in equal protection cases. If, for example, the law benefits members of class B but not of class A, plaintiffs who belong to class A will allege that they are relevantly similar to the members of B.[51] Plaintiffs will then assert that whatever reasons the state had for benefitting members of B must also logically require benefitting members of A; all are, by hypothesis, equally deserving. Because plaintiffs need not argue that they should be benefitted though they are undeserving, courts invariably confront a claim which on its face confers standing.[52]

---

48. *See generally* Comment, *The Right to Nondiscriminatory Enforcement of State Penal Laws*, 61 COLUM. L. REV. 1103 (1961).

49. *See* Califano v. Goldfarb, 430 U.S. 199 (1977).

50. *See* Ross v. Moffitt, 417 U.S. 600, 609 (1974): " 'Due process' emphasizes fairness between the State and the individual dealing with the State, regardless of how other individuals in the same situation may be treated. 'Equal protection,' on the other hand, emphasizes disparity in treatment by a State between classes of individuals whose situations are arguably indistinguishable."

51. *See* Califano v. Goldfarb, 430 U.S. 199 (1977). Leon Goldfarb claimed that he should receive benefits because he was as deserving as women who could not prove dependency, not because he was as *un*deserving.

52. In Allied Stores of Ohio v. Bowers, 358 U.S. 522 (1959), the Supreme Court seemed to hold that an equal protection claimant has standing even where state law requires that the discrimination, if proved, be "equalized" by extending the burden to others rather than by exempting the plaintiff from the burden. The plaintiff, a resident taxpayer seeking reduction of a tax on his stored merchandise, argued that similar property of nonresidents was exempted in violation of the equal protection clause. The state court had held that only the legislature could extend the exemption to plaintiff and that it was unnecessary to consider the constitutional claim, since plaintiff's property would be taxable regardless of the merits of his claim. The Supreme Court disagreed and reached the merits, reasoning that since the state court had not invalidated the exemption, it was immaterial that plaintiff's claim would be fruitless if the exemption *had* been stricken. (On the merits, the Supreme Court denied the claim.)

The Court's holding in *Bowers* is somewhat ambiguous. If indeed the state court had no power to remedy the alleged discrimination by extending the exemption to plaintiff, then why

Only when devising a remedy for an equal protection violation do courts directly confront the dilemma of protecting the "guilty" or benefitting the undeserving. Even here, courts may try to avoid the embarrassment of such a result by concluding that the legislature intended to benefit or excuse the class to which the plaintiff belongs, as well as the class actually benefitted. But such statutory interpretation must often be dishonest. The Court awarded benefits to Leon Goldfarb not because Congress intended that all surviving spouses be benefitted, for it surely did not, but because automatically granting benefits to widows but not widowers was unfair.[53] Although Congress intended to benefit *dependent* surviving spouses (and believed that widows are usually dependent while widowers are not), the Court properly granted benefits to Leon Goldfarb even though he was not dependent. Had it refused to do so, the Court could have avoided protecting the equal protection rights of the undeserving, but it would have denied Goldfarb adequate relief.

Of course, the legislature ultimately can remedy the discrimination in either direction; the presumption in *Goldfarb* that only widows are dependent can either be eliminated or extended to both widows and widowers. But a stingy judicial remedy would thwart the purposes of the equal protection clause by removing the incentive for bringing claims and by effectively making some claims nonjusticiable. Although the claimant must take his chances that the legislature will ultimately resolve the inequality to his detriment, he should not have to take that chance in court.[54]

---

should it have to address the merits? Perhaps the Court believed that the plaintiff would have been satisfied even with the "remedy" of eliminating the exemption for nonresidents. Read more broadly, *Bowers* may suggest that a state court must provide an effective remedy for equal protection violations and cannot simply leave the issue to the state legislature. *See* Iowa-Des Moines Natl. Bank v. Bennett, 284 U.S. 239 (1931). In any event, *Bowers* does suggest a liberal approach to standing in equal protection cases.

53. Califano v. Goldfarb, 430 U.S. 199, 216-17 (1977) (plurality opinion).

54. It should be noted that most courts and commentators disagree with this position and argue instead for a case-by-case determination of the wisdom of extending rather than eliminating a presumption. *See* Califano v. Jobst, 434 U.S. 47, 56 n.14 (1977); Stanton v. Stanton, 421 U.S. 7, 17-18 (1975); Skinner v. Oklahoma, 316 U.S. 535, 542-43 (1942); Arp v. Workers' Compensation Appeals Bd., 19 Cal. 3d 395, 407-11, 563 P.2d 849, 856-58, 138 Cal. Rptr. 293, 300-02 (1977); Note, *Extension versus Invalidation of Underinclusive Statutes: A Remedial Alternative*, 12 COLUM. J.L. & SOC. PROB. 115 (1975); *Developments, supra* note 1, at 1136-37. The issue is admittedly complex: the decision to extend benefits may drastically affect state budgets (*see* Shapiro v. Thompson, 394 U.S. 618 (1969), and companion cases); courts are naturally reluctant to "rewrite" legislation; and additional federalism concerns are implicated when a federal court devises a remedy for a state law equal protection violation. Nevertheless, it is important to preserve incentives for bringing equal protection challenges whether plaintiff is burdened or benefitted, deserving or undeserving. Courts might accomplish this by at least granting a prevailing *plaintiff* the relief he requests, even if wholesale extension of a benefit (or of an exemption from a burden) would require legislative action. This approach produces some unfairness itself, however, and seems unworkable for class actions. Alternatively, courts might delay relief until the legislature has had a reasonable time to react to the finding of an equal protection violation.

This introduction to misfit analysis has done little more than explain terms and describe how the misfit issue arises in equal protection litigation. This Note next examines existing techniques of *evaluating* misfit, that is, of determining when misfit violates equal protection.

II. Existing Approaches to Misfit

Before propounding a method of evaluating misfit, this Note will study the Supreme Court's current treatment of misfit. While that treatment has been doctrinally murky, generalizations can be derived from it by scrutinizing the standards the Court purports to apply. The first part of this section discusses the traditional "rational basis" test. The remaining subsections investigate the relationship of misfit to two kinds of "closer scrutiny": fundamental interests and irrebuttable presumptions. Principles from these areas will be tapped for precedent for a more sensitive equal protection scrutiny than the traditional rational basis test provides.

A. *Rational Basis Test*

The Supreme Court has recognized that the rational basis test demands an analysis of misfit, but it has not enunciated clearly its standards for finding misfit justifiable. None of the three broad standards which the Court's decisions suggest is unqualifiedly adequate.

In a frequently cited passage, the Court reasoned that a classification does not offend the equal protection guarantee "merely because it is not made with mathematical nicety or because in practice it results in some inequality."[55] To the extent that this first standard merely emphasizes the inevitable imprecision of legislative lines it is unobjectionable. However, it may also imply that a little inequality is always tolerable. Such an unqualified statement is inaccurate.[56] Slight inequalities should be allowed if it is too costly to remedy the misfit by identifying the members of the overincluded and underincluded classes. However, better reasons for permitting the inequality must be given than the flat assertion that the misfit is too slight to require judicial scrutiny. A complaint that it was unfair for the government to accept Spiro Agnew's nolo contendere plea is not sufficiently answered by the response that since almost all *other* tax

55. Lindsley v. Natural Carbonic Gas Co., 220 U.S. 61, 78 (1911). *See* Metropolis Theatre Co. v. Chicago, 228 U.S. 61, 69-70 (1913) ("The problems of government are practical ones and may justify, if they do not require, rough accommodations—illogical it may be, and unscientific"); Patsone v. Pennsylvania, 232 U.S. 138, 144 (1914) ("abstract symmetry" not required); Bain Peanut Co. v. Pinson, 282 U.S. 499, 501 (1931) ("the machinery of government would not work if it were not allowed a little play in its joints").

56. *See* Part IV.A *infra*.

evaders are found guilty and punished, the government's policy is largely "accurate" in punishing the guilty.

The Court frequently justifies misfit by a second standard, namely, that the legislature may strike where the harm is most acute.[57] Unlike the first standard, the second purports to *justify* misfit, or at least underinclusiveness. For example, the state could reasonably keep only the *most* dangerous drivers off the road. Thus, although a safe driver who fails the driving test may not be comforted to learn that legislative lines are inevitably imprecise, a very dangerous driver who fails the test might be convinced that the classificatory scheme is reasonable even though it does not screen out the marginally dangerous. However, the degree of harm standard is an incomplete test of classificatory inaccuracy; it is a necessary but not a sufficient condition of adequacy of fit.

Thus, *perfect* fit is not achieved just because the average harm in the burdened class exceeds by the slightest degree that in the class not burdened. A classification that picks out all and only harmful individuals is surely more "rational"—that is, it produces less misfit—than one that selects a class only marginally more "harmful" than the unburdened class and containing many harmless individuals. For example, a regulation requiring policemen to retire who individually have been found unable to perform their tasks causes less misfit than one requiring all policemen to retire at the age of fifty.[58] Some cases might satisfy the degree of harm standard yet still violate equal protection because of the ease with which the overincluded or underincluded could be properly treated. For example, no court would condone a criminal statute exempting redheaded persons on the ground that T (criminals without red hair) has a higher average harm than ∿T (non-criminals plus criminals with red hair). Of course, a "degree of harm" justification can be helpful, for if it is *not* possible inexpensively to identify similarly situated subclasses, the level of misfit might be permissible. Nevertheless, it cannot be conclusively presumed that misfit is justifiable merely because the average harm in the burdened class exceeds the average in the nonburdened class.

Under the Court's third misfit justification, the legislature may remedy parts of the mischief:[59] it may address problems "one step

57. Williamson v. Lee Optical, Inc., 348 U.S. 483, 489 (1955); Minnesota v. Probate Court, 309 U.S. 270, 275 (1940); Truax v. Raich, 239 U.S. 33, 43 (1915).

58. *See* Massachusetts Bd. of Retirement v. Murgia, 427 U.S. 307 (1976), which upheld mandatory retirement. Marshall's dissent correctly states that individualized review is more rational than a flat rule, but of course that fact alone should not invalidate the rule, for the cost of precision might justify a less precise rule. *See* Part IV.B.2 *infra*.

59. Jefferson v. Hackney, 406 U.S. 535, 546 (1972); Dandridge v. Williams, 397 U.S. 471, 486-87 (1970) ("the Equal Protection Clause does not require that a State must choose between attacking every aspect of a problem or not attacking the problem at all"); Williamson v. Lee Optical, Inc., 348 U.S. 483, 489 (1955) ("The legislature may select one phase of one field and

at a time"[60] and "need not run the risk of losing an entire remedial scheme simply because it failed, through inadvertence or otherwise, to cover every evil that might conceivably have been attacked."[61] This standard is by far the most troublesome of the three, since it appears to sanction completely irrational distinctions. A state which denied licenses to left-handed drivers to reduce auto pollution[62] would, after all, be proceeding "one step at a time." Almost any equal protection challenge would fail if this standard were adopted without qualification,[63] for any classification which confers a benefit or burden only on some of the similarly situated could be described as a legislature's partial remedy. As we shall see, the justification may be acceptable where the classification is inevitably arbitrary—for example, where genuine phase-in problems exist or where burdens or benefits are distributed as part of a bona-fide social experiment.[64] Nevertheless, by failing to distinguish these special cases the Court offers an analysis which, taken seriously, would eviscerate the rational basis test.[65]

Despite the availability of these three putative misfit justifications, the Court has in some cases found that classifications failed to

---

apply a remedy there, neglecting the others"); Central Lumber Co. v. South Dakota, 226 U.S. 157, 160-61 (1912) ("If a class is deemed to present a conspicuous example of what the legislature seeks to prevent, the fourteenth amendment allows it to be dealt with although otherwise and merely logically not distinguishable from others not embraced in the law").

60. Williamson v. Lee Optical, Inc., 348 U.S. 483, 489 (1955).

61. McDonald v. Board of Election Commrs., 394 U.S. 802, 809 (1969). *See* Queenside Hills Realty Co. v. Saxl, 328 U.S. 80, 84 (1946) (a legislature "need not take account of new and hypothetical inequalities that may come into existence as time passes or as conditions change").

62. The example is from Posner, *supra* note 23, at 29 n.56. Posner cites it as one of the few nonracial classifications so "palpably inconsistent" with equal protection as to be invalid. But there is a disturbing similarity between his example and other arbitrary classifications which the Court has upheld. In Geduldig v. Aiello, 417 U.S. 484 (1974), the Court upheld against an equal protection challenge California's exclusion of pregnancy from the definition of "disability" in its mandatory employee-funded disability program. The Court reasoned that the state need not insure against *all* risks of employment disability and that women were not discriminated against because "[t]here is no risk from which men are protected and women are not." 417 U.S. at 495-97. But to say that all risks *common* to both groups are covered is not to say that a classification is rational which covers all gender-specific risks of men yet fails to cover all such risks for women.

Of course, there might be other ways to explain *Geduldig*—for example, treating pregnancy as a special kind of disability meriting exclusion, or viewing the discrimination as serving the goal of having identifiable groups such as men and women receive benefits proportional to their contributions. (Apparently women are more "expensive" than men even without pregnancy coverage, 417 U.S. at 497 n.21.)

63. A few classifications would not pass judicial review even under this standard, *e.g.*, those drawn on explicit racial or ethnic lines or motivated by racial animus. No court would sustain a law denying licenses to blacks to reduce auto pollution. But any classification that is now judged under the rational basis test and not the compelling state interest test could be upheld under this "one step at a time" standard.

64. *See* Part V *infra*.

65. On a subtler interpretation, this test *is* valid. *See* text accompanying note 284 *infra*.

meet the rational basis test. Justice Marshall,[66] Professor Gunther,[67] and several other commentators[68] have recognized that a tacit balancing test has been used in these cases. But because the test is tacit, clear standards for gauging the significance of misfit cannot be directly derived from the cases. Consequently, these cases will not be separately examined, although many of them will be discussed by way of illustration throughout this Note. It is instructive, however, to investigate two doctrines by which the Court explicitly subjects classification to "closer scrutiny" than the rational basis test (as traditionally described) would warrant.

B. *Fundamental Interests*

In a series of cases, the Court has held that when certain "fundamental interests" (such as the right to vote or to travel)[69] are *differentially* infringed by a classification, the classification must fall unless it is "necessary to promote a *compelling* governmental interest."[70] In effect, the Court decides that, absent such an overriding state interest, a statute which satisfactorily classifies people with respect to a non-compelling interest must be invalidated if in some other, more important, respect it treats people unequally.

Equal protection is often substantive due process in disguise, for in many cases the Court objects not so much to the *differential* infringement of a fundamental interest as to the fact of the infringement itself.[71] In *Shapiro v. Thompson*,[72] for example, the Court invalidated a welfare residency statute because it penalized the right

66. *See* Justice Marshall's dissenting opinions in Massachusetts Bd. of Retirement v. Murgia, 427 U.S. 307, 317-27 (1976); San Antonio Independent School Dist. v. Rodriguez, 411 U.S. 1, 98-110 (1973); and Dandridge v. Williams, 397 U.S. 471, 519-21 (1970).

67. *See* Gunther, *supra* note 2.

68. *See* Perry, *Constitutional "Fairness": Notes on Equal Protection and Due Process*, 63 VA. L. REV. 383, 402-03 (1977); Simson, *supra* note 5, at 668-81.

69. The following interests have been denominated fundamental: the right to travel, Shapiro v. Thompson, 394 U.S. 618 (1969); the right of privacy, Carey v. Population Servs. Intl., 431 U.S. 678 (1977); the right to vote, Reynolds v. Sims, 377 U.S. 533 (1964); the right to marry, Zablocki v. Redhail, 434 U.S. 374 (1978); possibly the right to procreate, Skinner v. Oklahoma, 316 U.S. 535 (1942); and possibly the right of fair access to the criminal process, Douglas v. California, 372 U.S. 353 (1963), and Griffin v. Illinois, 351 U.S. 12 (1956). *See* People v. Olivas, 17 Cal. 3d 236, 551 P.2d 375, 131 Cal. Rptr. 55 (1976) (the right to liberty is an explicit equal protection fundamental interest). Among the "non-fundamental" interests are the right to an education, San Antonio Independent School Dist. v. Rodriguez, 411 U.S. 1 (1973); and the right to housing, Lindsey v. Normet, 405 U.S. 56 (1972).

70. Shapiro v. Thompson, 394 U.S. 618, 634 (1969) (emphasis original).

71. *See Developments, supra* note 1, at 1131-32; Shapiro v. Thompson, 394 U.S. 618, 659-63 (1969) (Harlan, J., dissenting). Justice Marshall has noted that the Court's limitation of fundamental interests in San Antonio Independent School Dist. v. Rodriguez, 411 U.S. 1 (1973), to rights explicitly or implicitly guaranteed in the Constitution is open to the objection that violations of such rights could be strictly scrutinized simply because of their constitutional status, thus rendering the fundamental interest doctrine superfluous. 411 U.S. at 100 n.59.

72. 394 U.S. 618 (1969).

of welfare recipients to travel. But a more equal infringement of that right, for instance, one which imposed qualitatively similar disabilities (such as a general residency requirement for employment) on new residents other than welfare recipients, probably would also be harshly received by the Court.

Nevertheless, there may well be interests which may be universally, but not differentially, infringed. In such cases the Court, in effect, heavily weights the individual's interest in not being burdened, and any overinclusion is assigned more "disutility" than it would ordinarily receive. Thus, in *Skinner v. Oklahoma*[73] the Court assumed that the state may sterilize all felons and that it may ordinarily draw somewhat arbitrary lines between offenses like larceny and embezzlement. A state may imprison embezzlers longer than larcenists or vice versa. Nonetheless, the Court found the arbitrariness of sterilizing larcenists but not embezzlers unacceptable because sterilization implicates the fundamental interest in procreation.[74]

Moreover, even a classification which produces perfect fit might be invalid if it infringes a fundamental interest. While this aspect of the fundamental interest doctrine may also appear to be disguised substantive due process, closer analysis reveals that it is not. If a burden substantially impairs a valuable personal right, then even though only those who pose the social harm in question are burdened, we are reluctant to burden them and not those who lack the harm. Therefore, a state may not condition the right to vote on payment of a poll tax,[75] even though (1) there is no constitutional right to vote in state elections and therefore no substantive due process objection[76] and (2) the tax may be rationally related to the state's interest in an informed electorate.[77] A poll tax would be unconstitutional even if it were perfectly rational, picking out all and only the best informed voters, since the state must produce a compelling justification for denying to some a fundamental right which is granted to others. In effect, the Court has stipulated that a state *must*, absent a compelling state interest, treat equally those who are similarly situated with respect to the exercise of a fundamental interest. "Purpose" (the usually permissive criterion of similar situation) has been replaced by this mandatory criterion.

Unlike the rational basis test, fundamental interest analysis may, as was suggested above, require even more than a perfect "fit." We shall see, however, that this surprising requirement can also be sensibly applied under the rational basis test. Misfit should be evaluated

73. 316 U.S. 535 (1942).

74. 316 U.S. 540-41.

75. Harper v. Virginia Bd. of Elections, 383 U.S. 663 (1966).

76. 383 U.S. at 665.

77. 383 U.S. at 684-85 (Harlan, J., dissenting).

with respect to state interests which are more than trivial compared to the burdened personal interests: perfect fit with respect to trivial state interests should not be enough.[78]

The fundamental interest doctrine also implies a second principle that can strengthen the rational basis test. Not *every* infringement of a fundamental interest does or should invoke "strict scrutiny," only those which significantly penalize the exercise of the interest. Thus, the courts have found that very short residency requirements do not "penalize" the fundamental right to travel[79] and that a state's refusal to provide counsel to indigents for discretionary state appeals does not "infringe" the fundamental right of access to the courts.[80] These cases might be described as recognizing only a right to a "meaningful" appeal or only a conditional right to travel. Perhaps more realistically, they represent attempts by the judiciary to manipulate the rigid fundamental interest categorization to avoid harsh results. Under either reading, the Court implicitly balances personal and governmental interests. A modified rational basis test should explicitly incorporate some of the balancing implicit in the fundamental interest doctrine—a trivial state interest should not justify serious infringement of a nearly "fundamental" personal interest. Thus, more than a weak administrative convenience rationale would be required for a classification which, by rendering unrelated households ineligible for food stamps, impinges on the personal interest in free association.[81]

Although the fundamental interest branch of the strict scrutiny test, as articulated by the Court, poses unresolved problems,[82] it is

78. *See* Part VI.B.3 *infra*.

79. *See* Starns v. Malkerson, 326 F. Supp. 234 (D. Minn. 1970), *aff'd. mem.*, 401 U.S. 985 (1971).

80. Ross v. Moffitt, 417 U.S. 600 (1974).

81. *See* United States Dept. of Agriculture v. Moreno, 413 U.S. 528 (1973).

82. First, it is quite unclear how a fundamental interest is recognized. Although the Court has recently declared that only those interests "explicitly or implicitly guaranteed by the Constitution" are fundamental, San Antonio Independent School Dist. v. Rodriguez, 411 U.S. 1, 33-34 (1973), this declaration fails to explain the cases, *see* 411 U.S. at 99-100 (Marshall, J., dissenting), and is better viewed as an indication that the Court is disinclined to expand the list of fundamental interests.

Second, this mode of analysis sometimes merely weighs misfit more heavily, as in *Skinner*, and sometimes demands that any infringement of the burdened class's interest be minimal regardless of the degree of fit. *See* Harper v. Virginia Bd. of Elections, 383 U.S. 663 (1966), discussed in text at notes 75-77 *supra*. The fundamental interest test as now applied might imply both procedures in every case. But *Skinner* suggests otherwise; although it weighed misfit more heavily, the Court intimated that more rational classifications would be valid, for it approvingly cited Buck v. Bell, 274 U.S. 200 (1927), a case that sustained the compulsory sterilization of institutionalized mental defectives. For further discussion, see Part VI.B.3 *infra*.

Third, it is unclear how much the misfit (or burden) must be reduced to satisfy the "less restrictive alternative" test. Indeed, the dual requirement that a classification be "necessary" to achieve a "compelling" state interest has been criticized as impossibly strict; rational-relation-to-a-compelling-interest and less-restrictive-means-to-a-legitimate-interest tests have been

consistent with and in part a natural development of the rational basis test. More importantly, as was suggested above, two elements of the fundamental interest doctrine—"perfect fit is not enough" and "implicit balancing"—can be recruited to fortify the rational basis test, thereby creating a full continuum between the restrained and active levels of equal protection review.[83]

C. *Irrebuttable Presumptions*

A remarkable due process doctrine with equal protection implications has recently surfaced which applies a stricter scrutiny to certain classifications than does the rational basis test. Under the irrebuttable presumption doctrine,[84] courts implicitly assume that a legislative classification which grants a benefit to members of T is intended to benefit another identifiable class T*.[85] Courts then find it a denial of due process for the legislature to establish a "permanent and irrebuttable presumption" that persons who are not within T are not within T*, when this presumption is "not necessarily or universally true in fact, and when the State has reasonable alternative means" of determining if a person is within T*.[86] The remedy is not to invalidate the classification, but to allow the complainant to rebut the presumption and establish his membership in T*, typically through a hearing. For example, the Court in *Stanley v. Illinois*[87] held that due process was violated by a statute which presumed that all unmarried fathers are unqualified to raise their children and which therefore required the state to take care of such children upon the mother's death. Even if most unmarried fathers are unqualified, the Court reasoned, due process requires a hearing on that issue.[88]

The irrebuttable presumption doctrine can be viewed as an attempt to remedy the misfit created by a statutory classification which is too underinclusive with respect to a benefit or too overinclusive

proposed as substitutes. *See generally* Dunn v. Blumstein, 405 U.S. 330, 363-64 (1972) (Burger, C.J., dissenting); Bullock v. Carter, 405 U.S. 134, 144 (1972); Singer, *Sentencing Men to Prison: Constitutional Aspects of the Burden of Proof and the Doctrine of the Least Drastic Alternative as Applied to Sentencing Determinations*, 58 CORNELL L. REV. 51, 55-58 (1972). For further discussion, see Part VI *infra*.

83. *See* Part VI.A *infra*.

84. *See, e.g.*, Cleveland Bd. of Educ. v. LaFleur, 414 U.S. 632 (1974); United States Dept. of Agriculture v. Murry, 413 U.S. 508 (1973); Vlandis v. Kline, 412 U.S. 441 (1973); Stanley v. Illinois, 405 U.S. 645 (1972).

85. Note, *The Irrebuttable Presumption Doctrine in the Supreme Court*, 87 HARV. L. REV. 1534 (1974).

86. Vlandis v. Kline, 412 U.S. 441, 452 (1973).

87. 405 U.S. 645 (1972).

88. In terms of T and T*, the statute in *Stanley* defined T as the class consisting of married parents and unmarried mothers. The Court, apparently defining T* as the class of suitable parents, required the state to allow unmarried fathers, who, by definition, were not in T, to demonstrate that they were in fact suitable parents and thus in T*.

with respect to a burden. If this doctrine states that any overinclusive or underinclusive classification demands individual hearings despite administrative cost, its failing is obvious. Almost every classification might be challenged, and without some constraint the doctrine represents "nothing less than an attack upon the very notion of lawmaking itself."[89] The Court has therefore evidently limited[90] the doctrine to classifications in which "basic human liberties are at stake,"[91] such as the rights to conceive and raise[92] one's children and to exercise freely personal choice in matters of marriage and family life.[93] As one commentator has suggested, the doctrine offers an intermediate equal protection[94] standard for "interests and classifications felt to be too significant or untrustworthy to be reviewed under the relaxed standard of equal protection scrutiny, but not significant or untrustworthy enough to qualify for fundamental or suspect status and the accompanying far more rigorous standard of review."[95]

Another feature of these decisions indicates a further limitation on the irrebuttable presumption doctrine. The Court has applied it only when a statute clearly purports[96] to distribute a benefit or burden according to some identifiable characteristics and yet distributes the benefit or burden according to a more general characteristic in order to save costs. But not every classification is a surrogate for another easily ascertainable class. More commonly the purpose class is extremely difficult to identify, and membership often would not be more accurately determined even by individual hearings. The doctrine is sensible only where "it *is* possible to specify those factors which, if proven in a hearing, would disprove a rebuttable presumption."[97]

---

89. Cleveland Bd. of Educ. v. LaFleur, 414 U.S. 632, 660 (1974) (Rehnquist, J., dissenting).

90. Although the reasoning in Weinberger v. Salfi, 422 U.S. 749 (1975), seems implicitly critical of the irrebuttable presumption doctrine, the doctrine may not be dead. One hint of its vitality is a per curiam decision endorsing the application of Cleveland Bd. of Educ. v. LaFleur, 414 U.S. 632 (1974), Turner v. Department of Employment Security, 423 U.S. 44 (1975). Another hint is the opinion of Justice Rehnquist (the most persistent critic of the doctrine) in Craig v. Boren, 429 U.S. 190, 226-27 (1976), in which he apparently endorses the doctrine in some circumstances. *See also* Usery v. Turner Elkhorn Mining Co., 428 U.S. 1, 22-24 (1976), in which the Court evaluated the irrebuttable presumption claim, though it ultimately rejected it on the merits. Some commentators have argued that the doctrine is still alive or should be resurrected, *e.g.*, Ackerman, *The Conclusive Presumption Shuffle*, 125 U. PA. L. REV. 761 (1977); Chase, *The Premature Demise of Irrebuttable Presumptions*, 47 U. COLO. L. REV. 653 (1976).

91. Turner v. Department of Employment Security, 423 U.S. 44, 46 (1975).

92. Stanley v. Illinois, 405 U.S. 645, 651 (1972).

93. Cleveland Bd. of Educ. v. LaFleur, 414 U.S. 632, 639 (1974).

94. Actually, the cases are better described as expressing an intermediate substantive due process overbreadth standard. *See* text at note 269 *infra*.

95. Simson, *The Conclusive Presumption Cases: The Search for a Newer Equal Protection Continues*, 24 CATH. U.L. REV. 217, 228 (1975).

96. Vlandis v. Kline, 412 U.S. 441, 452 (1973).

97. Weinberger v. Salfi, 422 U.S. 749, 804 (1975) (Brennan, J., dissenting) (emphasis origi-

It is often claimed[98] that the irrebuttable presumption doctrine is a disguised form of equal protection. Although this claim has some validity,[99] equal protection will often be more satisfactory for plaintiffs than the irrebuttable presumption doctrine, since even if a plaintiff using the latter doctrine wins, he will only have won the chance to rebut the presumption.[100] The irrebuttable presumption doctrine, then, is most useful to plaintiffs when the presumption would not have failed the rational basis test. But apparently the legislature can avoid such irrebuttable presumptions simply by deceit. That is, these classifications are invalidated not because they are not rationally related to any legitimate state interest (for they are), but because they are not rationally related to the goal which the legislature was foolish enough to identify, more or less explicitly, in the statute. In *Stanley*, for example, the state *could* have argued that its purpose was to discourage unmarried fathers from raising children rather than to ensure the fitness of parents, the purpose apparently inferred by the Court. Although the new purpose might still fail to justify the classification, it would present a more difficult constitutional question than that which the *Stanley* Court addressed.

---

nal). The Court in *Salfi* at least gives lip-service to this constraint: "Unlike the statutory scheme in *Vlandis* [v. Kline, 412 U.S. 441 (1973)], the Social Security Act does not purport to speak in terms of the bona fides of the parties to a marriage, but then make plainly relevant evidence of such bona fides inadmissible." 422 U.S. at 772.

98. *See, e.g.*, Cleveland Bd. of Educ. v. LaFleur, 414 U.S. 632, 651-52 (1974) (Powell, J., concurring); Note, *The Conclusive Presumption Doctrine: Equal Process or Due Protection?*, 72 MICH. L. REV. 800 (1974).

99. *But see* note 94 *supra*.

100. To be precise, plaintiffs will be worse off under an irrebuttable presumption analysis unless those who are similarly situated also have to prove, in an individualized hearing, that they deserve the benefit in question. *See* L. TRIBE, AMERICAN CONSTITUTIONAL LAW 1095-96 (1978); Bartlett, *Pregnancy and the Constitution: The Uniqueness Trap*, 62 CALIF. L. REV. 1532, 1545-47 (1974); Note, 46 U. COLO. L. REV. 79, 102 (1974). For example, in Stanley v. Illinois, 405 U.S. 645 (1972), the Court invalidated the statutory presumption that an unmarried father was unfit to raise his child after the death of the mother. The effect of the invalidation was to require the state to prove unfitness at a hearing under the standards applicable to married or divorced parents and to unwed mothers. This result would have been reached under either the due process or the equal protection clause, and the Court held that the statute violated both clauses. By contrast, if the plaintiff had been successful under an irrebuttable presumption analysis in Weinberger v. Salfi, 422 U.S. 749 (1975), she would only have gained the opportunity to rebut the presumption that a marriage within six months of application for Social Security benefits was entered into collusively; the "presumption," if you will, of non-collusiveness would be retained as to all who married more than six months before applying for benefits. On the other hand, had she been successful on an equal protection theory, plaintiff would have been entitled to the same reverse "presumption" as the longer-married and could have recovered benefits without any hearing.

The Court in *Salfi* missed this distinction. It argued that the prophylactic six-month rule is desirable in part because it "protects large numbers of claimants who satisfy the rule from the uncertainties and delays of administrative inquiry into the circumstances of their marriages." 422 U.S. at 782. This argument is fallacious, for if plaintiff were to prevail, the presumption of non-collusiveness could still be retained for persons married earlier than six months before their spouses' deaths.

In the typical irrebuttable presumption case, the Court argues as if it were simply effectuating actual legislative intent by correcting an inconsistent or obviously irrational classification. This reliance on actual intent, which also appears in some recent sex discrimination cases, suggests that the Court will invalidate some laws simply because the legislature did not fully and deliberately consider a goal, even though that goal would otherwise have justified the classification. In a later section,[101] this intriguing development will be explored more fully.

The irrebuttable presumption doctrine is not without problems, the most serious of which is the difficulty of determining whether a particular statute "purports" to test for ascertainable characteristics or instead represents a "substantive policy determination that limited resources would not be well spent in making individual determinations."[102] Nevertheless, the doctrine offers a brave new approach to equal protection problems, allowing the legislature to enact a loose classification with a provision for hearings through which to rebut the classificatory presumption or to enact a tighter classification with no provision for hearings.[103]

In a larger sense, the irrebuttable presumption doctrine is a kind of less-restrictive-means requirement which strengthens the rational basis test. The state must explore reasonable alternatives, even if the alternatives add costs. But, like the fundamental interest approach, an implicit balancing of state and personal interests appears to explain this departure from the simple rational basis test. Thus, although a state may deny nonresidents in-state rates for a year simply because of their residence and domiciliary status at the time of application,[104] it may not do so for four years.[105]

The irrebuttable presumption doctrine is a curious amalgam of equal protection rationality, substantive and procedural due process, and the least-restrictive-alternative test. At the least, the doctrine illustrates courts' occasional willingness to soften the rigid two-tiered equal protection standard through a creative balancing of interests. The fundamental interest doctrine, of course, evinces a similar willingness. Later sections draw upon balancing notions from these two doctrines to develop a more comprehensive theory of equal protection balancing. Before analyzing misfit and balancing more closely,

101. Part III.B *infra*.

102. Weinberger v. Salfi, 422 U.S. 749, 784 (1975).

103. Simson, *supra* note 95, at 232-33; *see* Skinner v. Oklahoma, 316 U.S. 535, 546 (1942) (Jackson, J., concurring).

104. Starns v. Malkerson, 326 F. Supp. 234 (D. Minn. 1970), *aff'd. mem.*, 401 U.S. 985 (1971).

105. Vlandis v. Kline, 412 U.S. 441 (1973).

however, we must discuss how to characterize properly the purpose with respect to which the classificatory means creates misfit.

III. Analyzing Purpose

Under the traditional rational basis test, the Court analyzes a classification flexibly: any classification rationally related to a conceivable purpose will be sustained.[106] The Court itself may "conceive" a possible purpose[107] and need not rely on the express language of the statute or the ingenuity of the state's attorney. And the purpose chosen by the Court need not be the "primary" purpose,[108] though it must be "legitimate" or "permissible."[109]

This flexibility poses the danger that every legislative classification could be trivially justified.[110] As Justice Rehnquist has noted,[111] the "purpose" (in the ordinary sense of that word) of a legislature in enacting a bill is to make its language part of the law.[112] This is hyperbole, but it raises a serious point: any burden imposed on one group can be tautologically justified as furthering the state's interest in burdening that group. If the congressional purpose in *United States Department of Agriculture v. Moreno*[113] was simply to deny food stamps to "hippie communes" (and other unrelated households), if the states in *Shapiro v. Thompson*[114] honestly desired to reserve their welfare funds for residents, then the means-end "fit" in each case is perfect. But these purposes do not persuade us that the classifications are permissible. Given the centrality of misfit analysis in the strengthened rational basis test developed in this Note, some limit must be placed on the purposes a state may invoke to justify its discriminations. This section describes two potential limiting devices: the requirements that a statute's purpose must be legitimate and that only actual purposes should be considered. While the propriety of the second constraint is not clear, the first is

106. *See Developments, supra* note 1, at 1078.

107. The Supreme Court has occasionally said that any "conceivable" circumstance or set of facts would support the reasonableness of the classification. McLaughlin v. Florida, 379 U.S. 184, 191 (1964); McGowan v. Maryland, 366 U.S. 420, 425 (1961); Lindsley v. Natural Carbonic Gas Co., 220 U.S. 61, 78 (1911). If this expression suggests that a court or legislature may invent facts that would substantiate a purpose, it is surely objectionable. Probably the Court only meant to restate the argument that *any* conceivable purpose may be invoked if some actual set of facts supports it. But it is vital that the question of factual support be addressed, if only cursorily, by a court. *See* Perry, *supra* note 68; Tussman & tenBroek, *supra* note 27, at 367-68.

108. McGinnis v. Royster, 410 U.S. 263, 276 (1973).

109. *See Developments, supra* note 1, at 1078; Part III.A.2 *infra*.

110. *See* Note, *supra* note 6.

111. Trimble v. Gordon, 430 U.S. 762, 782 (1977) (Rehnquist, J., dissenting).

112. *See also* Sandalow, *supra* note 20, at 1183 ("legislatures do not act irrationally").

113. 413 U.S. 528 (1973).

114. 394 U.S. 618 (1969).

consistent with the history and purpose of the equal protection clause and helps rescue the rational basis test from tautology.

A. *Legitimacy*

The requirement of a *legitimate* purpose is probably the single most important protection of misfit review. This constraint might be viewed (wrongly) as encapsulating all other equal protection constraints, if one believed that a statutory end necessarily described a statutory means. The "end" or purpose of our driving test would simply be "to put safe drivers, and only safe drivers, on the road, by means of a driving test, with the degree of precision that that driving test provides." The question whether the classification satisfies equal protection then collapses into the question whether the "end" is legitimate.

However, this view clumsily displaces the crucial questions instead of answering them. A more satisfactory analysis, and the one generally employed in this Note, distinguishes carefully between ends and means. The legitimacy of the end must still be investigated, but that inquiry need ask only whether the end is on its face impermissible, without regard to the means.[115] Only if the end is not impermissible need the closeness of the fit be examined.

There are three categories of equal protection goals: (1) goals which are legitimate, in the special sense that courts can simply accept them without further testing their rationality (that is, fit) in terms of a further goal; (2) goals which are obviously impermissible, in the sense that they can never justify a classification; and (3) all other goals, comprising those ends which must be tested both for rationality in terms of a further end and for the rationality of the classificatory means.

It bears emphasis that "legitimate" goals, as defined in this Note, are only a subset of those goals which are not impermissible. Many goals are neither impermissible nor "legitimate" in this special sense; as we will see, such goals are only tentative equal protection justifications, unlike "legitimate" goals.

The relationship between these three categories can be summarized by briefly describing how a court should test the legitimacy of statutory ends preparatory to analyzing misfit. First, it should summarily reject any impermissible goals. Second, it should examine the remaining goals to determine whether they are legitimate. If they are not, then the Court should seek a more general, legitimate end to which each end is closely related. Misfit should be evaluated only

115. For reasons that will become apparent, the inquiry will also extend to the question whether the means-end "purpose" (the desire to achieve the end by those means) itself is obviously *legitimate*. If so, closeness of "fit" is irrelevant, for the law automatically passes equal protection review.

as to the obviously legitimate ends identified by this process. In the unusual case where the desire to achieve the narrow statutory end by the statutory means can itself be described as an obviously legitimate goal, misfit evaluation is unnecessary, for the fit will be perfect and the classification can be described as "self-justifying."

The discussion of the three categories of goals that follows will clarify the procedure that courts should use in initiating misfit review and will explicate the relationship between the categories. The order of presentation departs somewhat from the actual procedure courts should follow in that obviously illegitimate goals are discussed after obviously legitimate ones.

1. *Legitimate Goals*

Any means-end analysis of legislation risks two contrasting criticisms. First, as we have seen, the "end" may always be tautologically defined so that the "fit" is perfect. Second, even if we do know how to start the analysis (if, that is, the first criticism is rejected), how do we know when to stop? Does a goal G_1, which does not tautologically explain the classification and which the classification adequately fits, satisfy equal protection requirements, or must goal G_1 "fit" a more general goal G_2? Must this process continue until "fit" is measured against the most general goal of all, the public welfare simpliciter? This kind of superrationality requirement is surely unacceptable, yet it is the logical consequence of the means-end approach.

The concept of "legitimate" goals defuses this second criticism. Misfit analysis may stop at a legitimate goal;[116] the court should examine the adequacy of fit relative to it but to no further goals. Consider, for example, the desire to subsidize farming. To say that this goal is legitimate is to reject as superfluous those justifications such as the need to encourage farming, or to stabilize farm prices, or to ameliorate the lot of farmers.[117] The state need not prove that its discrimination between farmers and others serves the general welfare, for such a discrimination is legitimate—it is the kind of "dis-

116. For an excellent discussion of the level of generality at which statutory purpose should be defined in misfit review, see P. BREST, PROCESSES OF CONSTITUTIONAL DECISIONMAKING 565-66 (1975).

117. *See* Carmichael v. Southern Coal & Coke Co., 301 U.S. 495, 512 (1937), sustaining a state unemployment compensation law against the challenge, *inter alia*, that its exemption of agricultural employers, charities, and other groups violates equal protection. "A sufficient answer," the Court reasoned, is "that the state is free to select a particular class as a subject of taxation" (here, those employing labor in the processes of industrial production and distribution). *Id.* The Court continued: "Reasons for the selections, if desired, readily suggest themselves," such as the desire to promote one industry over another. *Id.* But it is significant that the Court simply asserted these possible reasons, without analyzing their legitimacy or whether they were substantially served by the law in question.

criminatory" decision which the legislature is absolutely free to make.

However, unless "legitimate" is properly defined, this category of goals threatens to swallow much of equal protection review. In responding to the second criticism of means-end analysis, care must be taken not to succumb to the first criticism. States might attempt to defend any challenged goals as legitimate and, by defining these ends in terms of the statutory means, create the tautology that we seek to avoid. If courts accepted these characterizations uncritically, no law would violate equal protection. A law disenfranchising blacks is perfectly tailored to keep blacks from voting, but far from being legitimate, that purpose is obviously impermissible. Most legislative ends fall into a middle category, neither legitimate nor obviously illegitimate. Thus, the critical question for the courts is how to identify legitimate goals. Only if the bounds of that category are identifiable can the courts analyze misfit.

One possible definition is that a discrimination is legitimate if it can reasonably be assumed to serve the public interest. We may properly favor farming over industry because plainly satisfactory justifications consistent with the public interest are so numerous that the court need not second-guess the legislature. But this definition is not convincing. Consider, for example, legislative classifications of crimes. Some "non-criminal" conduct threatens public security and some "crimes" do not; nevertheless, a court would accept the state's classifications of crimes within the broadest limits.[118] Had Governor Carey signed the proposed New York law imposing heavier penalties for crimes against the elderly,[119] the law could not have been successfully challenged on equal protection grounds, yet the "public interest" served by such a law can be described only tautologically. Crimes against the elderly are offensive, and the legislature may express this moral judgment by law. No further justification—that the law promotes "human dignity" or protects the most helpless victims—need be adduced. Any justification for singling out crimes against the elderly on such grounds is subject to challenge: one might assert that crimes against small children aggrieve human dignity "more," or involve weaker victims. Yet we are more certain that the

118. Apparently the only substantive constitutional constraints on defining behavior as criminal are the prohibition against punishment for mere status, *see* Powell v. Texas, 392 U.S. 514 (1968); Robinson v. California, 370 U.S. 660 (1962); the eighth amendment ban on cruel and unusual punishments; and proscriptions against interfering with rights protected by the Bill of Rights (*e.g.*, free speech, privacy). None of these constraints significantly narrows the extraordinary legislative discretion in defining crimes.

119. *See* N.Y. Times, Aug. 17, 1977, at 1, col. 6. The bill would have set mandatory prison terms and would have limited plea bargaining for most serious crimes against the elderly and physically disabled. *Id.* at B4, col. 5.

original legislative judgment is valid than we are that justifcations are available.[120]

The proper definition of legitimate goals has a metaphysical ring: those legislative purposes that differ in kind, rather than degree, from any other. Thus, in determining whether a purpose is legitimate, a court must decide that the judgment is permissible and sufficiently different from potentially comparable judgments. Absent a clear societal consensus, courts must compare the burdened class to the class assertedly similarly situated. A law which grants drivers' licenses only to those who pass a driving test is sustainable if it is rationally related to keeping dangerous drivers off the road, an obviously legitimate purpose. Saying that keeping dangerous drivers off the road is obviously legitimate simply means that a dangerous driver cannot criticize the purpose on the ground that there are persons (*e.g.*, criminals) who endanger the public more than he. There is a difference in kind between pursuing criminals and excluding dangerous drivers from the highways; thus, differential treatment of these groups should not be judicially scrutinized.

As a further example, reconsider the statute challenged in *McGinnis*.[121] The denial of good-time credit for presentence time was justified by the absence of rehabilitation programs in jails. Here, neither the immediate purpose (giving credit only for postsentence time) nor the second purpose (giving credit only when rehabilitation programs were offered) is legitimate. Yet the second purpose would not satisfactorily explain the distinction if we knew that rehabilitation programs do not work or that as much rehabilitation actually takes place in the pretrial setting as in prison.[122] Our very need to ask these further questions suggests that the purpose is not legitimate. Thus, an even more general, third purpose must be invoked before misfit review is proper. That purpose, of course, is crediting incarceration time according to actual rehabilitation—a purpose, finally, which is legitimate, since it cannot be seriously suggested that the state must justify using rehabilitative progress and

120. *Cf.* L. WITTGENSTEIN, ON CERTAINTY 17e (1969):

"I *know* that I have never been on the moon." That sounds quite different in the circumstances which actually hold, to the way it would sound if a good many men had been on the moon, and some perhaps without knowing it. In *this* case one could give grounds for this knowledge. . . .

I want to say: my not having been on the moon is as sure a thing for me as any grounds I could give for it.

121. *See* text at notes 41-45 *supra*.

122. *But see* McGinnis v. Royster, 410 U.S. 263, 271 (1973): "We pass no judgment on the success or merits of the State's efforts, but note only that at state prisons a serious rehabilitative program exists." If the Court is simply stressing the difficulty of evaluating rehabilitative progress, then its unwillingness to pass judgment is reasonable, especially since the legislature needs freedom to test new schemes, the results of which cannot be predicted, *see* discussion at notes 256-61 *infra*. But if the quoted passage means that the Court would ignore convincing proof that rehabilitation could not occur in prisons, then its reasoning is unconvincing.

not, say, blameworthiness or deterrent effect as the criterion for granting good-time credit.

In sum, the "in kind" criterion requires a court to evaluate misfit only with respect to legitimate purposes.[123] If a proffered purpose is not legitimate, another usually more general purpose which *is* legitimate must be found.

No extended search for a legitimate purpose is required in the case of the self-justifying discrimination (one in which granting the burden or benefit to the statutorily defined class is itself an obviously legitimate goal). This is simply the tautological justification for misfit rejected earlier. The tautology is acceptable in a limited context, however, because it sometimes is *not* possible to explain a distinction more convincingly than simply to describe the distinction itself. Where this is the case, and where the purpose of distinguishing between recipients and nonrecipients of a benefit or burden precisely as indicated on the face of the law is obviously legitimate, the classification is self-justifying. For example, a law imposing the death penalty for murder is self-justifying. The distinctions on the face of the law between murder and other crimes, and between capital punishment and other penalties, are distinctions in kind. A court would not be perturbed if rape were punishable only by a life sentence, if assault only by a suspended sentence, or if heroin sales were not punishable at all.[124] Other constitutional constraints aside, the facial purpose of the law, punishing a murderer with death, is legitimate.

The role legitimate goals play in misfit analysis should now be clear. But we still have no simple explanation of *when* a statutory purpose is "in kind" and thus reflects a distinction which a court should not review.[125] Professor Brest has vividly delineated the

123. Evaluations of misfit must often cumulatively evaluate misfit as to several levels of purpose. That is, if means T is used to achieve end M, which itself is a means to end M_1, and only M_1 is legitimate, then we should make a rough estimate of the extent to which T serves M, discounted to the extent that M fails to serve M_1. In practice, a court need rarely engage in such subtleties; it will often be enough to judge the extent to which M_1 is served by means T.

124. Courts addressing the constitutionality of criminal prohibitions of the use of marijuana and cocaine often must decide whether the distinction is in kind or in degree. Of course, they do not describe the issue in these terms. But the success of an equal protection or due process challenge often depends on whether the court is convinced that marijuana and cocaine are no more harmful than alcohol. *See, e.g.*, State v. Erickson, 574 P.2d 1 (Alaska 1978) (cocaine more harmful than marijuana; prohibition of use is constitutional); State v. Leigh, 46 U.S.L.W. 2425 (Fla. Cir. Ct. 1978) (marijuana less harmful than alcohol or tobacco; prohibition of use is unconstitutional). If a court does not even accept the comparison between marijuana and alcohol as a matter of degree, however, the constitutional claim would probably fail. There are other minor criminal offenses, such as loitering or disturbing the peace, which clearly are not comparable to alcohol abuse. The difference is in kind, and therefore such offenses cannot as easily be challenged on equal protection grounds.

125. Useful discussions other than Professor Brest's (*see* note 116 *supra*) include Ely, *supra* note 26, at 1235-49; Note, *supra* note 6, at 139-51; Sandalow, *supra* note 7, at 656-61; *Developments, supra* note 1, at 1081; Tussman & tenBroek, *supra* note 27, at 356-61. Most of these

problem with a hypothetical regulation requiring the installation of exhaust emission devices on cars:

> [A]s one moves to higher levels of generality, the likelihood of finding rational grounds for distinction increases. But where to stop? How does one know that cars are to be compared with trucks and buses, and perhaps with airplanes and factories, but not with cigars, jackhammers, and pornography?[126]

As Brest's example illustrates, somehow we often *do* know when to stop.[127] Despite its metaphysical ring, the "in kind/in degree" distinction is one which courts naturally, albeit unconsciously, do employ. However, there may not be objective standards for when to stop requiring justifications,[128] and a court probably can only express intuitive,[129] informed and considered judgment.[130]

That this standard is amorphous is not critical, however, for a purpose which is not legitimate is not necessarily illegitimate. In

---

simply pose the problem, however. For example, one commentator notes that courts presume that favoritism toward an industry ultimately benefits the public and reasons that "denominating one industry as more worthy than another does no great violence to egalitarian ideals. . . . By contrast, a similar judgment made about the worth of individuals would require clear justification." *Developments, supra* note 1, at 1081. But the crucial questions are *why* our egalitarian ideals endorse different treatments of these two forms of favoritism, *how* different that treatment should be, and even how to *decide* what kind of favoritism a given classification embodies (*e.g.*, whether a subsidy to farmers promotes agriculture or is a judgment that certain individuals, *i.e.*, farmers, are worthier).

126. P. BREST, *supra* note 116, at 566.

127. *See* North Carolina v. Pearce, 395 U.S. 711, 722-23 (1969), where the allegedly discriminatory distinction seemed so clearly legitimate that the Court was unable even to fit the problem into an equal protection framework. The defendants had complained that a judge's imposition of a harsher sentence on retrial created an invidious classification: since convicts who do not seek new trials cannot have their sentences increased, that risk is imposed only upon those who succeed in having their original convictions set aside.

128. P. BREST, *supra* note 116, at 566.

129. Resort to intuition does not necessarily mean that a principle is weak. Professor Rawls has persuasively said that moral argument should seek a reflective equilibrium between principles and considered intuitive judgments. Thus, in determining the principles of justice, we should

> see if the principles . . . match our considered convictions of justice or extend them in an acceptable way. We can note whether applying these principles would lead us to make the same judgments about the basic structure of society which we now make intuitively and in which we have the greatest confidence; or whether, in cases where our present judgments are in doubt and given with hesitation, these principles offer a resolution which we can affirm on reflection.

J. RAWLS, A THEORY OF JUSTICE 19 (1971).

130. The intuitive approach, however, is occasionally unconvincing. For example, in a concurring opinion famous for its spirited defense of equal protection, Justice Jackson disposed of the merits of the claim with disturbing ease. He agreed with the majority that merchants could be permitted to advertise on their own trucks but not on the trucks of others, but he did not accept the majority's rationale that the first class posed a slighter traffic hazard than the second. According to Justice Jackson, the discrimination was fair "because there is a real difference between doing in self-interest and doing for hire, so that it is one thing to tolerate action from those who act on their own and it is another thing to permit the same action to be promoted for a price." Railway Express Agency, Inc. v. New York, 336 U.S. 106, 116 (1949). Unfortunately, Justice Jackson did not further define this "real difference," and a more obviously "real" (*i.e.*, legitimate) explanation would have been more persuasive.

finding that a purpose is not legitimate, a court is not invalidating the classification but only requiring that it be subject to misfit review with respect to some obviously acceptable purpose. Where a further justification would be instructive the court should therefore hesitate to conclude that a purpose is legitimate. While a driving test may serve the laudable purpose of giving licenses to those who pass driving tests, a court should rest its decision in favor of a more obviously legitimate purpose.

2. *Impermissible Goals*

What goals may not be pursued at all, whatever the means-end fit? In addressing this question, the Supreme Court has found impermissible not only objectives forbidden by other clauses of the Constitution,[131] but also the objectives of disadvantaging any suspect class[132] as well as those asserting "fundamental interests."[133] The unadorned desire to discriminate against women,[134] aliens,[135] illegitimates,[136] or "politically unpopular" groups[137] is also unlawful. However, the Court has not clarified the standard for determining "impermissibility." The Court's suggestion that discrimination for its own sake is never legitimate[138] is put in doubt by our analysis of "legitimate" discriminations. No reasoned basis for defining illegitimate purposes emerges from the Court's decisions.

This Note does not attempt to define precisely "impermissible" purposes. For several reasons, we need not do so. First, although *some* unlawful purposes invalidate the laws which they motivated,

131. *See* Chicago Police Dept. v. Mosley, 408 U.S. 92 (1972) (first amendment); *Developments, supra* note 1, at 1081.

132. *See* Griffin v. County School Bd., 377 U.S. 218 (1964); Yick Wo v. Hopkins, 118 U.S. 356 (1886).

133. *See* Shapiro v. Thompson, 394 U.S. 618 (1969).

134. *See* Stanton v. Stanton, 421 U.S. 7 (1975) (rejecting the state's interest in equipping boys, but not girls, to be family providers); Goesaert v. Cleary, 335 U.S. 464, 467 (1948) (statute denying bartenders' licenses to women other than relatives of a male owner might be invalid if its real purpose were "an unchivalrous desire of male bartenders to try to monopolize the calling").

135. *See* Examining Bd. v. Flores de Otero, 426 U.S. 572, 605 (1976) (finding "unpersuasive on its face" the state's desire to prevent the uncontrolled influx of Spanish-speaking aliens into the practice of civil engineering in Puerto Rico—a justification that "amounts to little more than an assertion that discrimination may be justified by a desire to discriminate").

136. *See* Mathews v. Lucas, 427 U.S. 495, 520 & n.3 (1976) (Stevens, J., dissenting).

137. *See* United States Dept. of Agriculture v. Moreno, 413 U.S. 528, 534 (1973) (although legislative history suggested a regulation was designed to exclude "hippie communes" from the food stamp program, equal protection "at the very least mean[s] that a bare congressional desire to harm a politically unpopular group cannot constitute a *legitimate* governmental interest"). *But see* Carrington v. Rash, 380 U.S. 89, 100-01 (1965) (Harlan, J., dissenting) (Texas may bar members of the military who moved to the state from voting in state election as long as they remain in the military, for it "could rationally decide to protect state and local politics against the influences of military voting strength").

138. *In re* Griffiths, 413 U.S. 717, 722 n.8 (1973).

all such purposes should probably not have this effect.[139] Impermissible purposes in the context of misfit analysis have a limited consequence—they are simply not accepted as justifications by the reviewing court, and the classification must be justified in terms of another (legitimate) purpose. Thus, a "liberal" interpretation of impermissible goals will not severely restrict legislative freedom. Currently both the courts and counsel for the state attempt to justify laws in terms of appealing legitimate goals.[140] The Court defended a Social Security formula benefitting women more than men as compensating women for past job discrimination[141]—a tentatively legitimate, if not "legitimate," goal—and not simply as increasing the material welfare of women more than the welfare of similarly situated men.

The imprecision of "impermissible" is not crucial for a second, contrasting reason. Even if impermissibility is conservatively interpreted, that is, if few goals are found impermissible, important constraints remain. A goal which is not unlawful might not be "legitimate" either, in which case the rationality of the means must still be tested against a more general legitimate goal.

Finally, insofar as a law stigmatizes, heavily burdens personal interests, or does not substantially further valuable state interests, it may not matter whether its purpose is unlawful, for the statute may fail under the equal protection balancing that follows misfit analysis.[142] The relationship of near illegitimacy to this balancing can be illustrated by examining one element of the balance, the significance of the state interest. An impermissible interest can be described as having *no* significance and thus can be ignored by the court. But a court more anxious to defer to a legislature can, more palatably, discount the significance of dubiously permissible interests in weighing the state's side of the equal protection balance.

Although equal protection analysis as described here and as practiced by the courts allows flexibility in assessing legitimacy, an impermissible goal should be dispositive of the equal protection claim in some cases. The courts dislike telling a state that its goal is unlawful, and they often purport to invalidate a law either because its purpose was not rationally served or because the state could not have intended so dubious a purpose. However, the most honest response would often be either to find the purpose impermissible or to

139. *See* the suggested approach in Eisenberg, *supra* note 26, at 134-46.

140. *But see* Stanton v. Stanton, 421 U.S. 7 (1976), in which the state court had upheld a gender discrimination as rationally furthering the commonly held notion that it is "the man's primary responsibility to provide a home." 421 U.S. at 10 (quoting 30 Utah 2d 315, 318, 517 P.2d 1010, 1012 (1974)).

141. *See* Califano v. Webster, 430 U.S. 313 (1977) (per curiam).

142. *See* Part VI, especially VI.B.3, *infra*. For a discussion of the requirement that courts review only actual purposes, see Part III.B *infra*.

find its validity so slight that it is weighed lightly in an equal protection balance.

A good illustration of the Supreme Court's reluctance to grapple with the problem of impermissible purposes is *Carey v. Population Services International*,[143] which held certain restrictions[144] on access to contraceptives violative of the first amendment and the due process clause.[145] The state argued that its ban on the sale of contraceptives to minors rationally furthered its goal of discouraging sexual promiscuity among minors. Members of the Court offered several responses to this argument. The first response, that the state could not have meant to "punish" fornication with pregnancy,[146] is a loaded characterization of the question.[147] The second response, that the regulation did not actually discourage promiscuity, seems more plausible.[148] However, as Justice Stevens noted in the third response, this only suggests that the statute was meant to have a symbolic impact.[149] Stevens summarily dismissed this "propaganda" interest, yet the value of symbolic impact might be more significant than he indicated.[150]

These responses notably avoid the most obvious and straightforward argument for invalidating the law—that discouraging fornication by the young is an impermissible purpose.[151] This avoidance technique may be properly conservative jurisprudence, but as ap-

143. 431 U.S. 678 (1977).

144. New York law prohibited any person from selling or distributing contraceptives to minors under the age of 16, prohibited anyone other than a licensed pharmacist from distributing contraceptives to persons over the age of 16, and prohibited any person from advertising or displaying contraceptives. 431 U.S. at 681.

145. Although *Carey* is a due process and not an equal protection case, its discussion of the legitimacy of state purposes is also applicable to equal protection cases. The Court freely quoted Eisenstadt v. Baird, 405 U.S. 438 (1972), an equal protection case, in analyzing the scope of the privacy interest.

146. 431 U.S. at 695.

147. The state's effort to encourage juveniles to be chaste is not necessarily "punitive" simply because minors may circumvent the regulation with consequent harm to themselves. One could just as well describe a prohibition on sales of alcohol as punitive because people will avoid the prohibition by illegally manufacturing impure and dangerous alcohol.

148. 431 U.S. at 695, 702 (White, J., concurring); 431 U.S. at 715 (Stevens, J., concurring).

149. 431 U.S. at 715-16 (Stevens, J., concurring).

150. Many of our criminal laws are ineffectual, and many actually encourage other forms of criminal behavior which "objectively" are as menacing as the prohibited activity. (The most obvious example is the probable effect of narcotics and gambling laws on the growth and power of organized crime.) Yet we feel it important that the criminal law express our moral judgments, however difficult it is to enforce them. True, New York might have contented itself with a criminal prohibition against sex between minors. But the improbability of a criminal prosecution for such an offense suggests the utility of the regulation on contraceptive sales—the unavailability of contraceptives to minors may be a better reminder to the young that the state condemns their sexual activity.

151. The plurality declined to reach this issue, 431 U.S. at 694 n.17. Two concurring Justices reached it and decided in favor of the end's legitimacy, 431 U.S. at 702-03 (White, J., concurring); 431 U.S. at 713 (Stevens, J., concurring).

plied here it has produced a holding supported by awkward and unconvincing reasoning.

Nevertheless, the Court was correct in invalidating the ban on sale of contraceptives to minors. As suggested, a better rationale would have been to deny the legitimacy of the end. Alternatively, the Court could simply have *discounted* the state's interest in discouraging promiscuity among the young to the extent that that interest is dubiously legitimate. Under this rationale, all of the arguments earlier criticized gain force. The plaintiffs' interests in avoiding pregnancy and disease demand a sounder justification than a questionably legitimate interest in discouraging their sexual conduct.

A final and fundamental objection can be raised to the underlying premise of legitimacy analysis. Professor Linde has argued that any "rational basis" test (and a fortiori the approach suggested in this Note) is premised on "a thoroughly instrumentalist view of law. It not only assumes that a law is always a means to an end, but it also asserts that law is constitutionally required to be a means to an end, and a rational means at that."[152] Linde believes that the instrumentalist view is fundamentally misconceived and misleading.

In part, Linde's objection is that lawmakers do not and *cannot* act rationally—an instrumentalist approach to legislation is just too complex and time-consuming.[153] This objection, however, is not dispositive. Even if the instrumentalist view is not a correct *description* of the political process, it does not follow that that view is an improper method of *justification*. If legislative decisions could be predicted perfectly by analyzing what the legislators had for breakfast, a court would nevertheless uphold those decisions if their *content* transgressed no constitutional limitation.[154] Judicial review under the equal protection clause does not assume that the legislature acts instrumentally, just as judicial review of first amendment claims does not assume that the legislature intends only to establish reasonable time, place, and manner restrictions on expressive conduct. The assumption is only that the legislature *can* pass laws satisfying the rationality standard. Rationality depends not on whether the legislature was "rationally" motivated,[155] but on the content of the law. Thus, the court should freely accept even justifications for the law which do not appear in the legislative record.

---

152. Linde, *Due Process of Lawmaking*, 55 NEB. L. REV. 197, 204 (1976). In addition, see the similar argument in Posner, *supra* note 23, at 27-28.

153. Linde, *supra* note 152, at 222-29.

154. Of course, the equal protection guarantee does prohibit laws passed from certain illicit motives. But the present discussion concerns not "hostile" discriminations, only discriminations that are simply "unfair." *See* discussion accompanying notes 19-27 *supra*.

155. Linde, *supra* note 152, at 229-32, would agree with this point.

Linde's main concern is really more basic. "The duty to defend the rationality of a decision," he writes, "depends very simply on whether the policy makers are limited to prescribed aims, or whether they are free to pursue any aim of their own choice."[156] This concern over limiting legislative aims is understandable, but the alternatives are less stark than Linde implies. He would admit that policy makers cannot pursue *any* aim they choose, *e.g.*, discrimination against blacks for its own sake. That courts forbid some aims does not mean, however, that they "prescribe" the rest, unless that term is given an unreasonably expansive meaning.

Linde's general point, that equal protection limits legislative choice, is valid, but we must not lose sight of the nature of those limitations. The legislature is *absolutely* free to choose between different goals if they are legitimate,[157] and it is absolutely free to employ any means of achieving a legitimate goal if the means is tailored closely enough to the end. That is, if the means chosen does not create too much "misfit" under the balancing test later described, then the law is valid, even if a more "rational," closer-fitting means might have been chosen. In short, no equal protection goals are ever "prescribed" by a court in the usual sense of that term.[158] Equal protection merely proscribes certain goals—those that are illegitimate and those that seek to achieve an end by an inadequately precise means.[159]

This is not to say that the effect of equal protection rationality requirements on legislative freedom is never a matter of concern. Indeed, the ultimate choice of the substantive values that are the referents of equal protection balancing will depend significantly on one's conception of the institutional relationship between courts and legislatures. The Supreme Court has ceded the legislature freedom to pursue economic ends, and it has thus formulated a weak "rationality" requirement for that field.[160] By contrast, the Court has disfavored legislative flexibility to experiment with important personal interests. Insofar as Linde is simply saying that courts should maintain a strong presumption that legislative decisions are rational, his position is defensible. More radically, however, Linde is also saying that *any* rationality requirement unacceptably interferes with the legislative branch, and this Note disputes that assertion.

---

156. *Id.* at 229.

157. *See* Part III.A.1 *supra*.

158. The fundamental interest approach can come rather close to prescribing certain goals, however. *See* Part VI.B.3 *infra*.

159. We would not want to describe "goals" in this way, for as this Note has already suggested, it is a clumsy way to analyze equal protection issues. *See* text at notes 110-15 *supra*.

160. *See* New Orleans v. Dukes, 427 U.S. 297, 303 (1976) (per curiam).

This discussion of legitimacy of purpose has attacked the contention that equal protection analysis can only ask whether an end is legitimate. Of course, equal protection misfit analysis and balancing involve value choices as fundamental as the judgment of an end's legitimacy. But the choices in misfit analysis and balancing are more constrained, and more easily adaptable to any particular view of judicial responsibility in this field.[161] Rescued from tautology, the rational basis test can be a convincing analytic tool. Following a discussion of the "actual" purpose problem, that rescue will continue.

B. *Finding the Actual Purpose*

A second possible constraint on purpose may be emerging in recent cases in which the Supreme Court has declined to consider purposes not actually contemplated by the legislature. The rational basis test has traditionally not required, and possibly even forbidden, inquiry into the legislature's "actual" purpose.[162] The Court has accepted *any* legitimate goal which might sustain the statutory classification. However, traditions change quickly in the equal protection field, and the Court has become more willing to examine actual legislative purposes, especially in sex discrimination cases.[163] As the Court declared in *Weinberger v. Weisenfeld*, "This Court need not in equal protection cases accept at face value assertions of legislative purposes, when an examination of the legislative scheme and its history demonstrates that the asserted purpose could not have been a goal of the legislation."[164]

The examination of actual legislative purpose is a significant development in equal protection jurisprudence,[165] although its implications are still hazy.[166] Where the actual purpose would not have justified the classification, but the discarded purpose *would* have, the law would be valid if reenacted after an explicit consideration of the purpose which was originally ignored.[167]

Califano v. Goldfarb,[168] for instance, strikingly illustrates this new

---

161. *See* text at note 431 *infra*.

162. *See* text at notes 106-07 *supra*.

163. This willingness to examine actual purposes is also evident in irrebuttable presumption cases. *See* text preceding note 101 *supra*.

164. 420 U.S. 636, 648 n.16 (1975).

165. Of course, this approach has no interesting consequences when a purpose which is discarded as non-actual was not rationally furthered by the classification, *see, e.g.*, Carey v. Population Servs. Intl., 431 U.S. 678, 694-95 (1977) (plurality opinion); Jimenez v. Weinberger, 417 U.S. 628, 634-35 (1974). Even if such a classification *had* been analyzed according to the discarded purpose, that purpose would not have justified it.

166. One helpful analysis is L. TRIBE, *supra* note 100, at 1085-88 (1978).

167. *See id.* at 1086 n.29.

168. 430 U.S. 199 (1977).

"activism,"[169] for both Justice Brennan's plurality opinion[170] and Justice Stevens' concurring opinion[171] plainly intimate that the statute might have been upheld had its actual purpose been different.[172] (Indeed, the dissenters and the plurality essentially divided over what this purpose was.)[173]

This way of handling legislative motive is not entirely new, for it resembles the treatment which numerous scholars[174] have advocated for racially motivated laws. However, the usual "improper purpose" discussion concerns motives which are rationally furthered by the law but which are illegitimate (*e.g.*, the intention to disadvantage blacks), while cases here involve legitimate motives which are not substantially furthered by the legislative means. The arguments which would favor a judicial "remand" to the legislature of racially motivated laws might not apply as forcefully to laws motivated (but insufficiently justified) by a permissible purpose. Laws in the latter category are not invalid because of judicial concern about stigma; they are invalid only because they are unfair. Since the analogy to impermissibly motivated statutes is imprecise, some other rationale must explain why a law which would have been fair if designed to achieve one goal must be invalidated because it was actually designed to serve another permissible goal which it served "imperfectly."

The rationale might be as follows. Judicial deference to legislative classifications producing unequal treatment rests on the belief that the populace may validly consent to any adequately general dis-

169. This might seem *less* activist than the traditional approach, since it allows a court to avoid the hard decision to invalidate a law as unfair with respect to *any* conceivable purpose. However, if an unconsidered purpose would have justified the law, this approach has a more severe consequence than traditional equal protection doctrine (which would permit the classification). If an unconsidered purpose would not have justified the law, the two approaches produce the same result.

This approach is only more conservative if it has the broader effect of validating *any* law once it has been reconsidered by the legislature, if its purpose, considered abstractly, seems insufficient as a justification. In *Goldfarb*, for example, a frank congressional declaration that widows should be treated differently than widowers because their "need" for benefits is qualitatively different (the "unconsidered" purpose), and not because of "administrative convenience" (the actual purpose), would automatically validate the law, whatever the misfit. But under this approach, it is not lack of misfit, but the very process of reconsidering the statute's fairness, that justifies judicial abstinence. Thus, the original "actual" purpose would also justify the law, regardless of misfit, so long as its fairness had actually been considered. *See* Sandalow, *supra* note 20, at 1185-90. The Court's treatment of purpose in *Goldfarb*, however, indicates that the Court is not willing to employ this broader and more conservative approach.

170. 430 U.S. at 214.

171. 430 U.S. at 222 & n.9, 223.

172. In Califano v. Webster, 430 U.S. 313 (1977) (per curiam), the Supreme Court upheld a Social Security benefit scheme that gave favored treatment to women, reasoning that (unlike the scheme in *Goldfarb*) the different treatment was a deliberate congressional attempt to compensate women for economic disabilities.

173. 430 U.S. at 229-35.

174. *See, e.g.*, Brest, *supra* note 26; Eisenberg, *supra* note 26, at 116.

tribution of benefits and burdens.[175] If, however, a court identifies as the purpose of a statute a goal which the legislature did not consider, the usual reasons for deference are lacking, since that goal could not have been consented to. The argument is not that such a goal was not the *primary* end sought by the legislature, but rather that it did not enter into the general calculus of interests at all, and hence does not deserve the usual judicial respect.

This activist rationale for confining a court's attention to actual legislative purposes invites the objection noted earlier[176] that a court in an equal protection case should not be a surrogate legislature, protecting any loser in the legislative process. But when actual purpose review is considered in its context this objection can be overcome, for even if courts should not ensure that all relevant interests are considered when a law is passed, perhaps courts should intervene when *none* of the relevant interests have been considered. In *Goldfarb*, for example, if the Court is correct that Congress actually intended to benefit only persons who were dependent on their deceased spouses, then Congress did not intend to favor widows over widowers to compensate for past discrimination against women. If this purpose *had* been considered, the statute might have been different, for the affected interest groups could have advanced the arguments relevant to this purpose. An explicit consideration of competing interests might have caused widows to be treated differently than widowers throughout the Social Security Act. Thus, a court's refusal to justify a law according to an unconsidered purpose might at least compel the legislature to take a pro forma look at the fairness of the law.[177]

---

175. The notion that people consent to differential treatment is at best a legal fiction. It is, nonetheless, a useful fiction, for it explains why courts closely scrutinize classifications that burden "suspect" groups historically underrepresented in the political process and that infringe certain "fundamental interests" in equal access to that process, such as the right to vote. *See* Kramer v. Union Free School Dist., 395 U.S. 621, 626-28 (1969). *See generally Developments, supra* note 1, at 1087-103, 1120-32. Of course, not all "fundamental interests" can be explained by this fiction, *e.g.*, the right of privacy (Eisenstadt v. Baird, 405 U.S. 438 (1972)), or the right to procreate (Skinner v. Oklahoma, 316 U.S. 535 (1942)).

176. *See* text at notes 19-21 *supra*.

177. Judicial deference to legislative decisions is often also predicated on the belief that the legislature is more competent to make judgments in most policy areas. *See* Carmichael v. Southern Coal & Coke Co., 301 U.S. 495, 510 (1937). But if this competence is not exercised because the legislature has not even considered the purpose, then this rationale for deference is inapposite.

Another possible argument for this activist approach to "purpose" is that it discourages the perpetuation of "archaic and overbroad generalizations." For example, Justice Stevens interpreted the challenged statute in *Goldfarb* as follows:

> It is fair to infer that habit, rather than analysis or actual reflection, made it seem acceptable to equate the terms "widow" and "dependent surviving spouse." That kind of automatic reflex is far different from either a legislative decision to favor females in order to compensate for past wrongs, or a legislative decision that the administrative savings exceed the cost of extending benefits to nondependent widows.

This argument has some cogency, though the Court's aversion to facing the uncertain merits of some unconsidered purposes, more than any particular rationale, may have inspired this doctrine. Whatever the Court's reasons for selectively confining itself to actual purpose, the doctrine itself is subject to several fundamental criticisms. First, most laws, especially state laws, have *no* recorded legislative history; it would be anomalous to strike these laws down. Perhaps the doctrine only requires ignoring a purpose where the legislative record clearly shows that the purpose was not considered, but it is unsafe to infer from a silent record that a purpose was not considered.

Second, even if these evidentiary obstacles can be overcome, the doctrine is theoretically unsound. It assumes that equal protection guarantees a legislative process guided by rationality and fairness. But the actual decision-making process should not be the central subject of the equal protection clause. Even if legislators do not act rationally, the Court may impose a limited rationality "requirement"—not to ensure that the *process* is rational, but to ensure that the laws it produces are.[178] If we agree that (aside from suspect categories) equal protection primarily ensures that laws *operate* fairly, it is difficult to justify the doctrine that a reviewing court should ignore unconsidered purposes. Thus, while actual purpose has been a limiting principle in some equal protection cases under the Court's current approach, the principal constraint on purpose should be the requirement, described above,[179] that misfit be evaluated with respect to a "legitimate" goal.

IV. MISFIT ANALYSIS

The analysis of misfit has been remarkably scanted by courts and commentators. Its first serious treatment, Tussman and tenBroek's

---

I am therefore persuaded that this discrimination against a group of males is merely the accidental byproduct of a traditional way of thinking about females.
430 U.S. 199, 222-23.

This explanation of the activist approach is incomplete, however. It does not necessarily follow that simply because a generalization is "archaic," "traditional," or "based on habit" it is too broad to satisfy equal protection. Surely equal protection does not demand the invalidation of every statute that is not periodically reconsidered and reenacted. Nevertheless, Stevens' approach has some advantages, for determining whether a classification is overbroad involves sensitive and complex judgments that are simplified somewhat by the knowledge that the underlying generalization is "archaic." The deference courts must accord legislatures (sometimes even described as a "presumption of rationality") is less compelling when the legislative assumption is of the *type* that is often adopted without examination. It may then be reasonable to conclude that the assumption was *not* examined in this particular case and that it is thus more likely to be overbroad. Nevertheless, the tenuousness of these inferences suggests that the "archaic generalization" argument should be used cautiously.

178. *See* text at notes 152-55 *supra*.

179. *See* Part III.A.1 *supra*.

influential 1949 article,[180] is almost the only one.[181] This section essays an examination of misfit which goes beyond the vague notion that a law is unfair which is "too" overinclusive or underinclusive. The discussion initially describes the minimum rationality constraint; the kinds of differences equal protection minimally demands to justify different treatment; and the three elements of misfit analysis—harm, cost, and personal interest. This section next distinguishes between misfit in due process and equal protection and between overinclusion and underinclusion. It also investigates the special analytical problems posed by laws with multiple goals. The relationship between misfit analysis and equal protection balancing, alluded to throughout the following discussion, is considered in detail in the section on balancing.[182]

A. *Minimum Rationality Constraint*

With some limited exceptions,[183] it is improper to burden one individual and not another unless the burdened individual is "relevantly" different. An individual is relevantly different only if he belongs to a class which has (1) a greater index of the social harm or mischief which the burden is designed to control, (2) a lower index of the personal interest in not being burdened, or (3) a lower index of administrative cost in applying the burden.[184] More succinctly, unless one class, on the average, is more dangerous than another class, less interested in avoiding the burden, or less costly to identify or to apply the law to, then burdening the first class and not the second is a denial of equal protection.

Consider, for example, the hypothetical law which reduced air pollution by denying drivers' licenses to left-handed people. This classification would violate the minimum rationality constraint, absent evidence that the left-handed drive "dirtier" cars, that they have a weaker interest in driving, or that the administrative cost of applying the law to the right-handed is much higher. If the individuals in each class posed the same average danger, it would violate the minimum rationality constraint to burden one class simply because that class was responsible for a larger proportion of the total mischief than the other. This is why it violates the constraint to keep only left-handed drivers off the road to reduce pollution, even though

180. *See* Tussman & tenBroek, note 27 *supra*.

181. *See also* P. BREST, *supra* note 116, at 558-75.

182. Part VI *infra*.

183. *See* Part V *infra*.

184. This Note singles out these three factors because each represents an element of "social utility" (broadly conceived) which may vary among individuals and which thus may help explain and justify differential treatment. By contrast, social interests as to which all persons are equal, such as a total cost constraint in a program in which the cost per individual is constant, are not relevant in evaluating misfit. *See* Part V.B *infra* (discussion of "neutral goals").

most drivers are right-handed. Minimum rationality prevents the state from attacking the larger part of a "mischievous" class unless the members of that part are also more mischievous.

This constraint is both intuitively satisfying and rationally sound. If the state has decided that there is net social utility in denying licenses to left-handed drivers to reduce pollution, then it must have concluded that the benefit of the incremental reduction in pollution outweighs the dual costs of the expense of administering the law and the infringement of people's interest in driving. But since, by hypothesis, the incremental harm, administrative cost, and interest in driving are precisely the same for the right-handed and left-handed, the social utility in keeping each right-handed or left-handed person off the road is precisely the same. Burdening only one class is thus irrational on the legislature's own terms.[185]

The minimum rationality constraint, importantly, draws attention to the *kind* of misfit equal protection prohibits. As we shall see,[186] the simple overinclusion/underinclusion terminology can be seriously misleading. The crucial inquiry is not *how much* misfit is created in any absolute sense, but rather *how different* the groups are that the law treats differently.

Of course, the minimum rationality constraint only begins misfit analysis, for satisfaction of the requirement is neither necessary nor sufficient to satisfy the equal protection guarantee. The constraint is not a necessary condition because it contains a number of exceptions, discussed subsequently.[187] It is not a sufficient condition because after it is satisfied, the law still must be subjected to a balancing test.[188] The constraint does identify *which* elements—harm, administrative cost, and personal interest—are to be balanced.

Although the minimum rationality constraint is not a radical concept, its role has often been overlooked because of the misconception that misfit analysis need evaluate only the absolute amount of overinclusion and underinclusion. Tussman and tenBroek's graphic representation of misfit is probably a principal source of the misconception. After criticizing that representation, this Note proposes a different model, one which illuminates the role of the minimum rationality constraint and which allows due process and equal protection misfit to be easily distinguished.

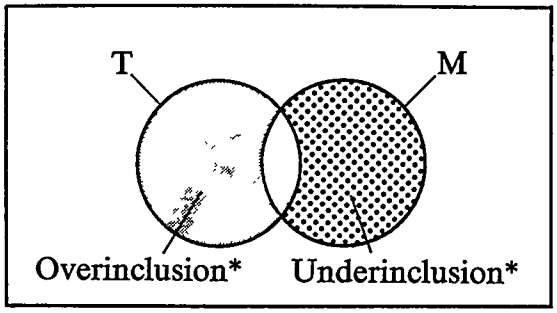
185. There are complications, of course. Diminishing returns, a total cost constraint, or a "satisficing" approach in which the government is content with a certain level of mischief-reduction (or need-alleviation), all indicate that it may be rational to go only "part of the way." Nevertheless, such a partial solution, since it violates the minimum rationality constraint, is only justifiable as an *exception* to misfit analysis. *See* Part V *infra*.

186. Part VI.B.1 *infra*.

187. Part V *infra*.

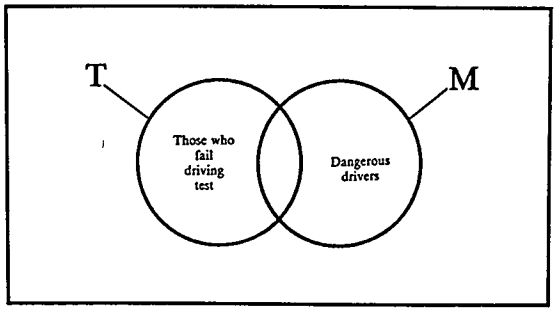
188. *See* Part VI *infra*.

Tussman and tenBroek represented the relationship between classes M and T by a Venn diagram:



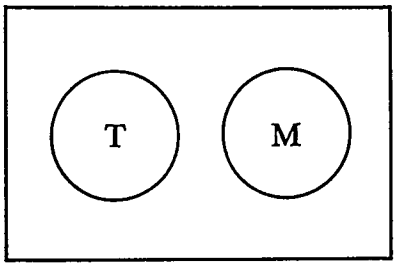
*Overinclusion is represented by solid shading and underinclusion by the dotted area.

For example, Tussman and tenBroek would represent the driving-test hypothetical as follows:

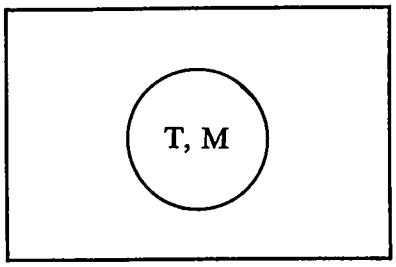


Their model helpfully stresses the following basic relationships:

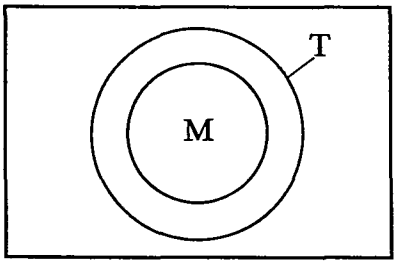
No intersection of the classes:



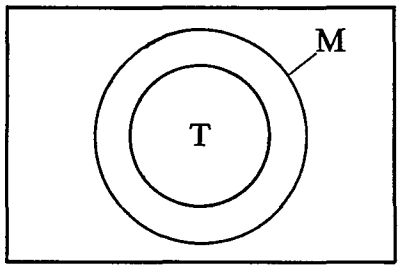
Complete congruence:



Overinclusion without underinclusion:

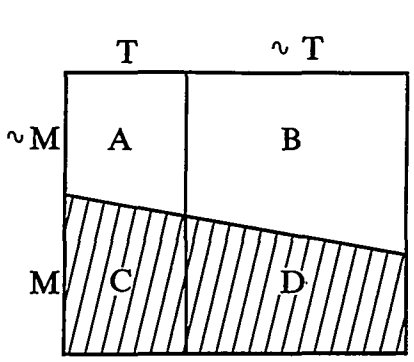


Underinclusion without overinclusion:

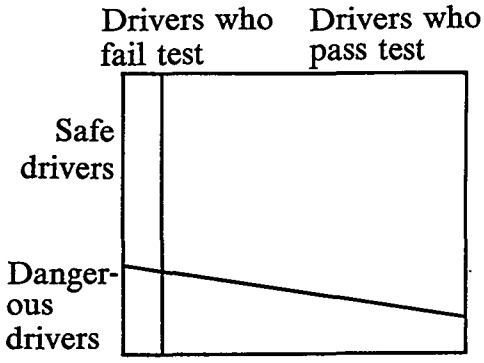


However, their model is deficient and misleading since it fails to identify the more rational cases of partial intersection short of complete congruence and to account for the properly excluded (∿M∩∿T) (represented by the field surrounding the circles). That is, the model suggests that misfit simply measures the absolute amount of overinclusion and underinclusion. It fails to emphasize that a classification is at least minimally rational, regardless of the absolute amount of misfit, if the burdened (T) are likelier than the nonburdened (∿T) to pose the relevant harm (to intersect with M). In the above example, there is no clear way to deduce from the graph whether those who fail the driving test are more likely to be dangerous drivers than those who pass. Any representation which fails to provide that information is seriously incomplete.

This Note proposes a somewhat different model:



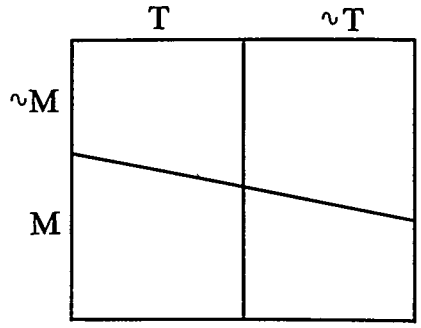
MODEL



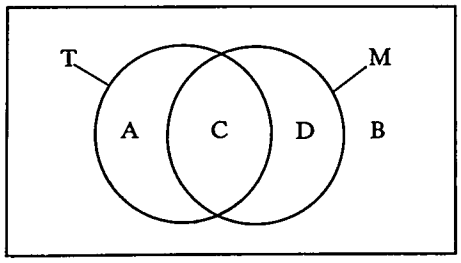
EXAMPLE

This model contains all the elements of the Venn diagram, but displays them more meaningfully.[189] The lower trapezoid (C plus D) is M; the rest of the rectangle is ∿M. The left rectangle (A plus C) is T; the right rectangle is ∿T. The relative size of each region (A, B, C, and D) is significant, which is not true of the Venn diagram. The key equal protection inquiry[190] concerns the slope of the dividing line between M and ∿M. The steeper this line (assuming a negative slope), the greater the difference in the harm posed by members of T (those burdened or benefitted) than by members of ∿T (those not burdened or not benefitted)—and the more "rational" the classification.

The minimum rationality constraint simply requires that this dividing line have some negative slope, that is, some "downhill" tilt:

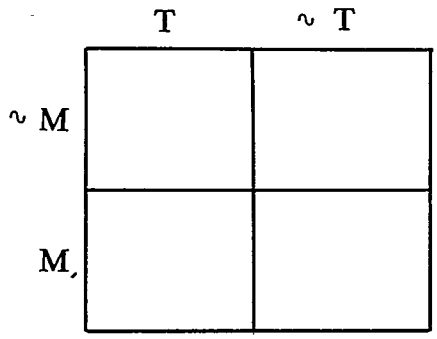


189. Tussman and tenBroek's model, with analogously labeled regions, is as follows:

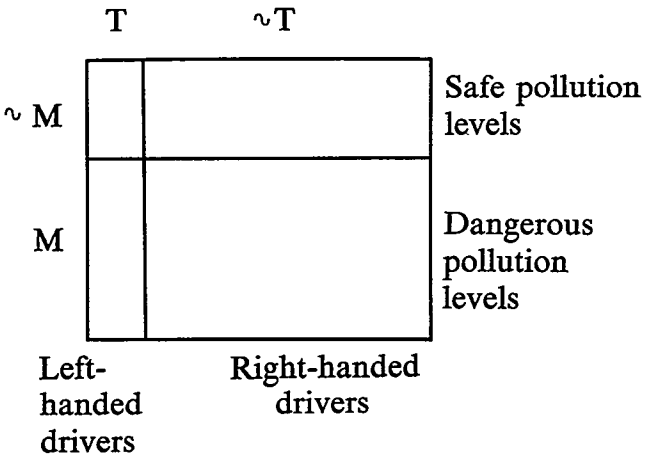


190. This model only describes the state interest (M) and does not include personal interest or administrative cost. A more complete (four-dimensional!) model could be constructed, but as a practical matter, it would not greatly aid the analysis: the state interest is often the most important element of analysis, and it tends to vary more between classes T and ∿T than do the other two elements.

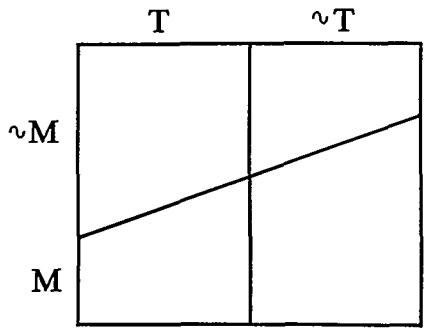
Thus, the negative tilt in the driving test example graphed above indicates that the classification is minimally rational. If the line has a slope of zero, then the law is not rational, for those not burdened (∼T) deserve to be burdened just as much as those who are burdened (T):



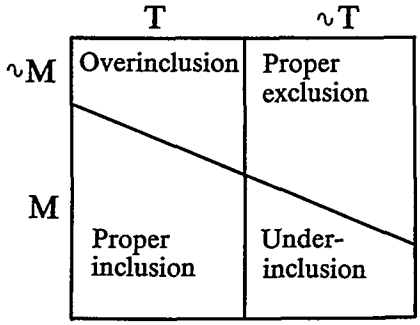
For example, the hypothetical pollution control law which precluded the left-handed from driving would be indicated thus:



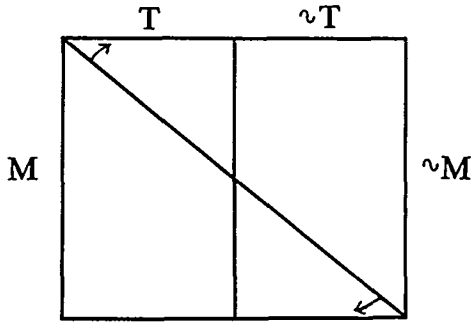
If the line has a positive slope, then the law is actually *irrational*, for those not burdened deserve to be burdened even more than those who are burdened:



The model indicates overinclusion and underinclusion as follows:

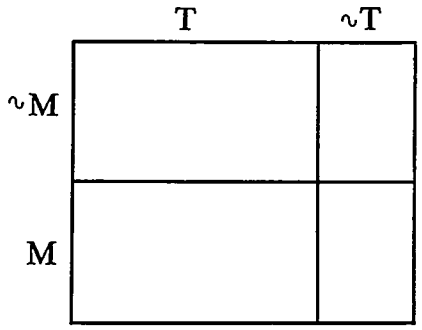


A classification which barely satisfies the minimum rationality constraint will have a large amount of absolute misfit. But as the negative slope of the dividing line between M and ∿M increases, the amount of misfit decreases:



Perfect fit is achieved when the dividing line becomes perfectly vertical, coinciding with the dividing line between T and ∿T.

These diagrams illustrate why the part-of-a-harm rationale, without more, does not justify misfit. Even if a classification attacks the larger part of the harm, *e.g.*:



the classification might not be minimally rational, since (as in the diagram above) the group burdened might not pose the harm more acutely.

Of course, although this model represents the important equal protection relationships more clearly than the older model, it is mainly a theoretical tool. Courts will rarely have clear proof of the precise areas of each region or of the exact slope of the dividing line. But if the model demonstrates the proper *focus* of misfit analysis, it will have more than served its purpose.

The remainder of this section describes the three central but often loosely analyzed elements of misfit analysis—the harm the state seeks to eliminate, the administrative costs of imposing the burden, and the personal interest in avoiding the burden. Some consequences of the minimum rationality constraint for different kinds of misfit will also be discussed. The substantive question how misfit is weighed in the equal protection balance once the minimum rationality constraint is satisfied is deferred.[191]

B. *Elements of Misfit Analysis*

1. *Harm*

The harm which the state seeks to eliminate was extensively discussed in the section on purpose,[192] where criteria were proposed for gauging the minimal level of generality at which the purpose, and thus the harm, must be expressed if judicial review is to be effective. A few additional comments may clarify the relationship of harm to misfit review. A harm may be a discrete quality, such as having

191. *See* Part VI *infra*.

192. *See* Part III.A.1 *supra*.

committed a crime (one either has committed the crime or has not), or a continuous one, such as being a dangerous driver (one can be more or less likely to cause accidents[193] or to terrorize other motorists). In the latter case, we ask not simply how many individuals in each class threaten the harm, but also how much each threatens. Thus, the conclusion that one class has a higher average index of harm than another can mean either that the first contains a greater proportion of (discretely) harmful individuals or that the mean harm, obtained by averaging each individual's (continuous) harm, is greater.[194]

An intermediate characterization of harm is also possible: the legislature may believe that only individuals who pose a threshold level of harm should be burdened, even though there may be some utility in burdening the less mischievous. In our driving example, for instance, an "intermediate" characterization was implicit: only those persons reaching a threshold of "dangerousness" need be forbidden to drive. This intermediate characterization can, however,

193. Even though an accident is a discrete event, and though a given person either will or will not cause one, at the time a law is passed persons are only more or less likely to cause accidents; we are not omniscient.

194. Self-justifying classifications are analytically related to the distinctions between discrete and continuous harm. Although it *is* possible to classify "discretely" harmful persons with perfect rationality by burdening all and only the harmful, it is not possible so to burden the "continuously" harmful, for the very notion of precisely identifying the "harmful" makes no sense in this context. Thus, we can *always* ask whether a classification by continuous harms is rational; the classification will never be self-justifying.

Line-drawing problems present an interesting analogy to the point that "continuous" classifications are never self-justifying. The legislature must decide "where to draw the line" between the classified (T) from the unclassified ($\sim$T) only when the trait by which members of T are identified is a continuous rather than a discrete one. (The mischief itself may or may not be continuous.) For example, the problem arises when the state establishes a mandatory retirement age, but not when it discriminates on the basis of sex.

Whenever we have a trait which is continuously and positively correlated with a mischief, it will be "rational" to classify according to any degree of the trait, since the resulting classification will always burden a class that more acutely poses the harm. (Imagine a round target so shaded that its color is more intense at the bull's-eye and gradually less intense in all directions toward the circumference. Every circle, regardless of size, whose center is the bull's-eye defines a "rational" classification according to most intense color.) The difficulty arises in determining when a quantitative difference becomes a qualitative one. Often a result is clearly permissible at one extreme and prohibited at the other, and it becomes difficult for a court to justify the point at which it imposes the prohibition.

For example, the Supreme Court has forbidden one-year residency requirements for voting, Dunn v. Blumstein, 405 U.S. 330 (1972), but has permitted residency requirements of 50 days, Marston v. Lewis, 410 U.S. 679 (1973). Further litigation may fix the limit at, say, 60 days, but the precise point at which the line is drawn cannot be justified by *any* test. The balance of interests would be only negligibly different were the limit 61 days, and the most we can say is that a line must be drawn somewhere. *See* Louisville Gas & Elec. Co. v. Coleman, 277 U.S. 32, 41 (1928) (Holmes, J., dissenting).

The continuity problem arises in similar fashion when we are evaluating the efficacy of a burden or benefit in reducing a mischief or need. A burden (or benefit), like a trait or mischief, can be discrete or continuous. For example, deciding whether to send a convicted felon to prison or to a drug rehabilitation program is not a line-drawing problem, since the choices are not (in any apparent way) part of a single continuum. By contrast, deciding whether to sentence him to two or to five years *is* a line-drawing problem.

be misleading insofar as it suggests that any sub-threshold harm is nonexistent, for as we shall see, this sub-threshold harm can be important. Unfortunately, many courts do not see misfit in continuous terms, perhaps because the continuous interpretation is not easily captured by traditional overinclusion/underinclusion terminology.[195]

Jefferson v. Hackney[196] illustrates the tendency to interpret harm and need as intermediate rather than continuous in order to simplify equal protection review. The Texas Constitution limits the state's welfare budget. The Texas legislature, after constructing a need standard for various categories of welfare recipients, applied reduction factors to keep the total payments within the budget: the aged received 100 per cent of their "need," the disabled and blind 95 per cent, and recipients of Aid to Families with Dependent Children (AFDC) only 75 per cent.[197] Members of the last class sued, alleging, *inter alia*, that imposing a different reduction factor upon different recipient groups violated equal protection. The Supreme Court upheld the state's procedure, reasoning:

> [T]he State may have concluded that the aged and infirm are the least able of the categorical grant recipients to bear the hardships of an inadequate standard of living . . . [and] that the young are more adaptable than the sick and elderly, especially because the latter have less hope of improving their situation in the years remaining to them.[198]

In effect, the Court assumed that the state had in mind a "severe need" standard independent of the actual need standard upon which payment levels were determined. The state's unequal reduction[199] factors belie any motive of maintaining a rough equality between groups with respect to the original actual need computations. Rather, the state's method suggests a weighting of the importance of needs *within* the original need computations, suggesting "severe need" thresholds.

The result in *Jefferson v. Hackney* is probably correct, if only because the aged and disabled suffer such distinct hardships that their needs may be "discrete." Indeed, a welfare program which served *only* their needs (and excluded AFDC recipients entirely) would probably be constitutional, since determining "need" involves not just an objective computation of food and housing costs, but also a judgment about relative "worth," about how much sympathy is aroused and about which groups seem helpless. The case neverthe-

---

195. *See* Tussman & tenBroek, note 27 *supra*.

196. 406 U.S. 535 (1972).

197. At the time suit was filed, AFDC recipients were receiving only 50% of their standard of need. 406 U.S. at 537 n.3.

198. 406 U.S. at 549.

199. Notice, for example, that the aged's share is not reduced at all.

less shows how an "intermediate" characterization of the harm vitiates the difficulties of review posed by "continuous" characterizations. The Court essentially ends its analysis with the assertion that the old and disabled are less able to withstand hardship. It does not ask *how much* less able they are.

Since even a discrete characterization was possible in *Jefferson v. Hackney*, the intermediate approach is not greatly troubling. But if a discrete characterization were not possible, perhaps the intermediate approach should be unacceptable. Suppose, for example, that Texas applied different reduction factors based on the sex of the AFDC children. Since the discrete discrimination (welfare funds only for boys or only for girls) is *not* acceptable here, the intermediate approach may also be unacceptable. A general assertion that when the going gets rough, boys can be forced to make a greater sacrifice is simply impermissible. Slight differentials might be sustainable if correlated with proved differences in need, but this would be a straightforward application of the "continuous" approach. In short, the intermediate approach should be used cautiously where the discrete is unavailable.

2. *Administrative Cost*

Administrative cost is the second major factor in evaluating misfit. This cost has two components: the costs of identifying to whom the burden shall apply, and the cost of applying the burden. Courts typically speak of administrative "convenience" (which is essentially the reduction of identification cost and administrative discretion) as merely one of several goals a legislature might or might not pursue. Despite the courts' usual analysis, administrative cost is central in equal protection analysis.

Identification cost is a basic element of the minimum rationality constraint and of general equal protection balancing because almost any classification can be made more precise at *some* cost. In other words, harm and identification cost are closely interrelated,[200] for if the state has satisfied the minimum rationality constraint, then the

200. If class A has a greater average harm than class B and if the classes can be distinguished at little administrative cost, then we could describe a classification burdening A as rational *either* because A has a greater index of harm than does B *or* because it costs more (per class member) to pick out harmful members of B than to pick out such members from A. Simply by burdening A and not B we have made a cheap and reasonably efficient selection, and it will be more expensive to find a subclass of B which is as harmful (on the average) as the class A.

The "higher cost of discriminating" explanation of a law's rationality refers, of course, to average, not aggregate, cost. For example, if X represents the class of jaywalkers and Y represents the class of jaywalkers minus redheads, then it is irrational to prosecute only class Y rather than the larger class X. Although prosecuting Y rather than X saves money in the aggregate, the average administrative cost of prosecuting a member of either group is presumably the same. Assuming there is no difference in "harm" or "personal interest" indices between X and Y, the prosecution violates the minimum rationality constraint. A law

unburdened will usually pose a lower average harm; to identify those underincluded from that class will generally cost more than it has already cost to separate the burdened from the unburdened.[201]

For example, consider the lament of a dangerous driver who fails our driving test but complains that many who passed it were also dangerous. He claims that he is no more dangerous than many who are on the road and that therefore even minimum rationality has not been achieved. He is wrong, for to *find* the dangerous drivers who passed the first test, the state would have to pay for a second, better test. In terms of the constraint, the dangerous drivers who slipped through the first time have a higher "index" of administrative cost, and treating them differently than the plaintiff is minimally (though not conclusively) rational.[202]

By misunderstanding this interrelationship between harm and cost, the Supreme Court has reasoned imprecisely in several sex discrimination cases. The Court has held that discrimination by gender must fall if its "only" justification is administrative convenience,[203] but that it can stand if the differential treatment is "rational" in terms of the statutory purpose, that is, if the class selected by the statutory classification is more acutely harmful or needy than the class not selected.[204] But the Court's distinction is illusory, since classifications in these cases are minimally rational in terms of *both* administrative cost and harm.

In *Schlesinger v. Ballard*[205] the Court upheld a sex-based classification which set a longer period of service for female than male naval officers before mandatory discharge for want of promotion. The Court reasoned that the discrimination was rationally related to the goal of equitable career advancement since it compensated for lesser

---

mandating such discriminatory prosecution should only stand if one of the misfit exceptions is applicable, *see* Part V *infra*.

201. In the usual case, this relationship holds because the state will ordinarily try to get at the most harm at the least cost. But occasionally the state will have made an expensive and inaccurate "pick," and it may be *both* cheaper and more accurate to redraw the classification. Often, though not invariably, it will also be the case that the greater the class difference in average harm, the greater the difference in average identification cost. Of course, there is nothing *necessary* about these relationships.

202. There is admittedly a certain awkwardness in characterizing our dangerous driver as "more costly" to identify. What is readily meant is that he is a member of a *class* more expensive to identify than the class now identified by the law. But the identification cost he creates is not as unique to him as the mischief he creates.

There is a second kind of administrative cost besides identification cost—the cost of applying the benefit or burden. This kind of cost is analogous to mischief in that individuals create unique "amounts" of it. *See* text following note 223 *infra*.

203. The Court has made similarly broad assertions in many irrebuttable presumption cases about the constitutional insignificance of administrative cost. *See, e.g.*, Vlandis v. Kline, 412 U.S. 441, 451 (1973); Stanley v. Illinois, 405 U.S. 645, 656 (1972).

204. *See* cases discussed in text at notes 205-11 *infra*.

205. 419 U.S. 498 (1975).

opportunity for promotion.[206] *Reed v. Reed*[207] and *Frontiero v. Richardson*,[208] in which sex-based classifications had been invalidated, were distinguished by the Court as involving classifications based on administrative convenience "alone"[209] and on "archaic and overbroad generalizations."[210] By contrast, the Court characterized the classification in *Schlesinger* as one of "complete rationality."[211]

In distinguishing the three cases, the Court failed to recognize that the classifications in *Reed* and *Frontiero* might be at least somewhat rational—women might be generally less competent administrators than men (*Reed*),[212] and wives of servicemen may be more likely to be dependent upon their spouses than are husbands of servicewomen (*Frontiero*).[213] Furthermore, in all three cases the classification was overbroad and could have been improved at some administrative expense—a case-by-case test of competence in *Reed* and of dependence in *Frontiero*[214] would be more "precise," and in *Schlesinger* the Navy could evaluate each woman's opportunities for promotion.[215] Yet the Court was not prepared to maintain that, by not evaluating opportunities individually, the Navy indefensibly relied on "administrative convenience."

This criticism of the Court's attempt to distinguish *Schlesinger* from *Reed* and *Frontiero* is not meant to suggest that all three classifications were equally justifiable. While these cases do not prove that a classification cannot stand if its only justification is "administrative convenience," they do support a more qualified proposition: if the burdened pose only a slightly or questionably greater harm than those not burdened (or if the benefitted are only slightly more needy than the unbenefitted), then the classification might violate equal protection, even though a more "rational" classification

206. Apparently women could not compile records of seagoing service comparable to those of men because they were precluded by law from most sea duty. 419 U.S. at 508.

207. 404 U.S. 71 (1971).

208. 411 U.S. 677 (1973).

209. 419 U.S. at 510.

210. 419 U.S. at 507-08.

211. 419 U.S. at 509.

212. The opinion in the state court had so suggested. Reed v. Reed, 93 Idaho 511, 514, 465 P.2d 635, 638 (1970).

213. This observation is not, of course, meant to imply that this minimal rationality, even if factually established, should validate the classifications.

214. That is, the test would require servicemen to prove that their spouses were dependent. The inequality could also be rectified by requiring neither sex to prove dependency, but this would decrease the overall classificatory accuracy (assuming the purpose is to aid all and only dependents).

215. However, if a plausible argument can be made that *all* women in the Navy deserve compensation for past discrimination, then this redefined purpose clearly does *not* require a case-by-case evaluation.

would be more costly.[216] Thus, the Court may have described the classification in *Schlesinger* as "completely rational" because it was much more accurate than the classifications in *Frontiero* and *Reed*—virtually all servicewomen lacked career opportunities available to servicemen, but a significant percentage of women are not dependent on their husbands in the service or are competent administrators. And the Court may have scoffed at the "cost" rationale in *Reed* and *Frontiero* because there was such a slight gender-based differential in "harm" that the administrative savings were illusory (*e.g.*, the cost saved by waiving proof of wives' dependency might have been less than the cost of benefits paid to nondependent wives).[217] In short, identification cost or "administrative convenience" is just one factor to be balanced in an equal protection case; it cannot alone determine whether a classification satisfies the guarantee.

The relationship between identification cost and misfit is such that attempts to increase classificatory accuracy often increase identification costs. Identification cost will be lowest when a legislative rule alone defines class membership. When such a rule is facially overbroad, the overbreadth can be cured at essentially no cost.[218] For example, a law punishing both criminals and all redheads can simply be rewritten. But when greater accuracy requires more than rewriting a general rule—when, in other words, it is costly to determine who falls within the new rule—then identification cost can be a serious concern.[219]

Consider, for example, *Massachusetts Board of Retirement v. Murgia*,[220] which tested a statute requiring that all uniformed state police officers retire at age fifty. The Court upheld the statute, finding that the rule reasonably expressed the general relationship between advancing age and decreasing physical ability to be a good

216. *See* Mathews v. Lucas, 427 U.S. 495, 509 (1976): "[P]resumptions in aid of administrative functions, though they may approximate, rather than precisely mirror, the results that case-by-case adjudication would show, are permissible . . . so long as that lack of precise equivalence does not exceed the bounds of substantiality tolerated by the applicable level of scrutiny."

217. The plurality in *Frontiero* noted that the government had offered no concrete evidence that the differential treatment in fact saved money—a demonstration it said was required to satisfy the demands of strict scrutiny. 411 U.S. 677, 689 (1973). *See* Mathews v. Lucas, 427 U.S. 495, 509-10 (1976), holding that such a strict demonstration is *not* necessary under the more relaxed standards of review. *See also* Califano v. Goldfarb, 430 U.S. 199, 219-20 (1977) (Stevens, J., concurring) (arguing that because the presumption of dependency caused the payment of substantial additional benefits, administrative convenience could not have been the reason for the discrimination).

218. *See* Note, *Less Drastic Means and the First Amendment*, 78 YALE L.J. 464, 468-70 (1969) (making an analogous point about first amendment overbreadth).

219. *See id.* at 470 n.31.

220. 427 U.S. 307 (1976).

policeman. Although, as Justice Marshall noted in dissent,[221] Massachusetts could easily have provided that officers be individually tested—indeed, it did annually test officers who were 40 to 50 years old—the majority sustained the rule as rational, apparently because the ease of improving accuracy was irrelevant in the absence of a fundamental interest or suspect class. The additional cost of tests may also have influenced the Court.

Murgia illustrates that the crucial inquiry with respect to identification cost is usually not the cost of redefining a general rule, but the cost of determining whether an individual falls within the terms of the redefined rule. In arguing for "individualized" testing, Justice Marshall is asking for a redefined rule—required retirement of those who fail a medical examination—and the real cost difference lies in the greater cost of *proving* a person's status under the new rule than under the old. It simply costs more to conduct a physical exam than to check birth dates.

Moreover, attempts to increase the accuracy of a rule may, depending on the rule's complexity and the discretion with which it is to be applied, decrease its predictability, thus frustrating the interest in predictability shared by the state and by persons subject to the rule.[222] An age limit is simple to administer and permits few disputes about proof. A medical test involves complex judgments about what constitutes good physical condition and is easily abused by doctors anxious to keep particular officers on (or off) the force. This loss of predictability is part of a classification's identification cost.

Finally, "individualized hearings" are not a costly but perfect solution to classificatory inaccuracy. Often, "individualized hearings" are simply a sensible device for administering a rule which requires more complex determinations of fact than statutory presumptions can easily accommodate.[223] Nothing about a hearing as such guaran-

221. 427 U.S. at 325-27.

222. *See* Weinberger v. Salfi, 422 U.S. 749, 782-83 (1975).

223. For example, the Court in Vlandis v. Kline, 412 U.S. 441 (1973), struck down a general presumption that certain broad categories of applicants for in-state tuition were not bona fide residents, since some such applicants *did* "possess many of the indicia of Connecticut residency, such as year-round Connecticut homes, Connecticut drivers' licenses, car registrations, voter registrations, etc. . . ." 412 U.S. at 448. The Court concluded that Connecticut should have granted every individual applicant the chance to prove bona fide residence. But this conclusion only asserts that the rule should be made more precise (at least as to those disfavored by the presumption), not that every case should be evaluated equitably on its own facts. Presumably an applicant who lacked a year-round home in a state, a driver's license, or an automobile or voter registration could still be denied in-state rates, if these indicia are accurate enough. This would be true whether the proof is presented to an individual hearing officer or simply mailed to the registrar of the university.

Of course, individualized hearings may serve important functions other than improving

tees perfect accuracy, for accuracy will often be limited either by the kinds of proof accepted or by the need to rely on the subjective opinion of the hearing officer. Whether individualized hearings are required to improve the accuracy of a classification depends upon a number of factors, including the ease with which proof requirements can be satisfied and the desirability of allowing hearing officers to exercise discretion.

Administrative cost includes, in addition to identification cost, the cost of applying a benefit or burden. It is not obvious how the courts should deal with differential application cost. May prosecutors refrain from enforcing laws against Orthodox Jews because of the high cost of providing Kosher diets in prison? May welfare departments cut costs "rationally" by excluding recipients with the highest benefits? Surely not. Having defined a harm or a need, the state should save costs by proportional diminution of burdens or benefits; the actual cost of applying the benefit or burden seems to bear at best an accidental, and at worst an inverse, relationship to desert. To be sure, classifying those harmful individuals who can most easily be identified is also arbitrary in that two equally mischievous individuals might be treated differently simply because one possessed, and the other lacked, the classificatory trait. But this arbitrariness in identification costs is a necessary evil of legislation; application cost does not seem so indispensable.

Nevertheless, when a difference in application cost results in significant differences in cost-effectiveness, application cost should be considered in deciding whether the minimum rationality constraint has been satisfied. Professor Brest, in an informative discussion of the problem, maintains that either a difference in cost efficiency between individuals (which he calls criterion 1) or a difference in the harm posed by different individuals (which he calls criterion 2) will support a classification's rationality.

> Consider, for example, the decision to require the installation of exhaust emission control devices on vehicles. Suppose that vehicle *Q* produces 1,000 units of pollutants, and vehicle *R* produces 500 units; but for technical reasons, installation of the control device will reduce *Q's* output by only 200 units, while the same device will reduce *R's* output by 400 units. Requiring *R* but not *Q* to install the device would satisfy criterion *1*: the choice would be premised on efficiency—attaining the greatest reduction in pollution per dollar. Requiring *Q* but not *R* to install the device would satisfy criterion *2*: the choice here would be premised on the arguable fairness of imposing

---

classificatory accuracy, such as requiring the state to engage in a personal dialogue with the plaintiff before acting. *See* P. BREST, *supra* note 116, at 691-92; Tribe, *Structural Due Process*, 10 HARV. CIV. RIGHTS-CIV. LIB. L. REV. 269 (1975).

the costs of the regulation on those who impose the greatest pollution costs on society.[224]

Brest's analysis is impeccable and as a theoretical matter is irreconcilable with the hypotheticals posed earlier: it clearly is cost-efficient to concentrate resources on catching "low-cost" prisoners who do not keep Kosher and on helping only the less expensive welfare recipients.[225] If we have divergent intuitions about pollution devices and Kosher diets, it must be because we have implicitly balanced in each case—*e.g.*, the extra cost of Kosher diets hardly justifies not prosecuting a suspected felon. Nevertheless, this discussion does indicate that application cost should be an element of the equal protection balance.

3. *Personal Interest*

Finally, classes may differ with respect to their personal interest in avoiding a burden. This differential is rarer than that of cost or harm, for we commonly assume that people have an equivalent interest in freedom from restriction and that restrictions will burden that interest equivalently. The guilty value liberty as much as the innocent. But such differential interests can exist. Classifications limiting the rights and privileges of children may rest partly on the sense that the young do not expect equal treatment and are less concerned when burdens are imposed. Similarly, some restrictions on the rights of aliens seem to be partially justified by their lesser "interest" in full participation in the political community.[226]

Differential personal interests can often be characterized as differential needs which the state has an interest in recognizing. Thus, we allow police to exceed speed limits because they have a special "interest" in chasing suspects, which is simply to say that the state is pursuing both the goal of crime control and of highway safety.[227]

Personal interest is more than merely an element of the minimum rationality constraint. Courts frequently weigh differential harm or cost according to the personal interest burdened without explicitly discussing the *differential* interest in avoiding the burden. Even

224. P. BREST, *supra* note 116, at 560.

225. This assumes that meeting the full need of one person is "equivalent" to meeting the full need of another. *But see* Jefferson v. Hackney, 406 U.S. 535 (1972), discussed in text at notes 196-98 *supra*.

226. Of course, this justification is unconvincing as applied to resident aliens who do all that is possible to become citizens. But as applied to nonresident aliens and aliens uninterested in citizenship, the argument has some force. *See generally* Rosberg, *Aliens and Equal Protection: Why Not the Right To Vote?*, 75 MICH. L. REV. 1092, 1110-11 (1977).

227. In a more general sense, we can say that the government also has an interest in respecting individual liberty. *See* P. BREST, *supra* note 116, at 988 n.1.

though two classes have nominally the same personal interest, the nature and strength of that interest has equal protection importance simply when one class is burdened and the other is not. Dangerous and safe drivers may have the same interest in driving, but what is crucial in evaluating misfit is the importance of that interest; less misfit will be tolerated as the significance of that interest increases. The fundamental interest branch of strict scrutiny, the "adequate opportunity to present claims" limitation on the indigent's rights in the criminal appellate process,[228] the apparent limitation of irrebuttable presumption analysis to important personal liberties,[229] and the special protection which the new substantive due process affords privacy rights[230] and family values,[231] all indicate that misfit is more "closely scrutinized" as the significance of the affected personal interest grows. A fuller description of how the courts should integrate personal interest into equal protection balancing appears later.[232]

C. *Types of Misfit*

1. *Due Process vs. Equal Protection*

As some commentators have noted, the analysis in substantive due process cases resembles that in equal protection cases.[233] The irrebuttable presumption doctrine suggests that procedural due process is also related to equal protection. This section attempts to expound the relationships between equal protection and the two forms of due process by identifying the relationship of misfit to each.

The intimacy of substantive due process and equal protection is immediately apparent from the fact that the language of the tests applied is nearly identical. Under either doctrine, if a fundamental personal interest is infringed the state must show that its classification is necessary to achieve a compelling state interest; otherwise,[234] the classification need only be rationally related to a legitimate state interest. This similarity is deceptive, however. Substantive due process straightforwardly balances state and personal interests to determine what means the state may employ against anyone (and everyone) to achieve a goal. Equal protection, by contrast, balances only incidentally, to determine whether a given amount of inequality

228. *See* Ross v. Moffitt, 417 U.S. 600 (1974).

229. *See* text at notes 90-93 *supra*.

230. *See* Carey v. Population Servs. Intl., 431 U.S. 678 (1977).

231. *See* Moore v. City of East Cleveland, 431 U.S. 494 (1977).

232. Part VI *infra*.

233. *See, e.g.*, Gunther, *supra* note 2, at 41-43; Linde, *supra* note 152, at 203-05; Perry, *supra* note 68, at 385.

234. Of course, equal protection also applies the strict scrutiny test to suspect categories.

is justifiable—the premise, always, is that an "equal" deprivation would be valid.[235]

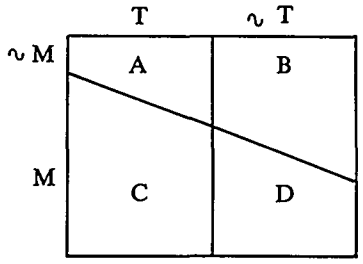
Both doctrines require that classes be compared, and to this extent both involve "misfit" analysis. But the classes relevant to each doctrine are different. Substantive due process speaks only to overbreadth[236] and therefore compares only two classes, the properly included and the overincluded.[237] Equal protection speaks to differential harm between the burdened and the unburdened and therefore compares four classes.[238] Put another way, it compares two broad classes—the burdened and unburdened—and asks whether the former class threatens more harm than the latter.

A plaintiff can often make both substantive due process and equal protection arguments. His due process claim is: "the state may not treat me the same way it treats him, since we are different,"[239] and his equal protection claim is: "the state must treat me

235. The details of each kind of balance are described in Parts VI.B and VI.C *infra*.

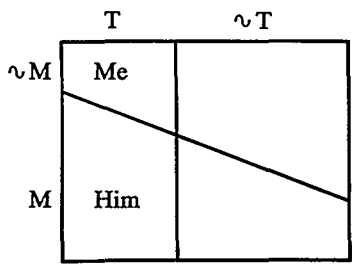
236. As we will see, overinclusion must be specially defined for benefits—it must mean overinclusion with respect to the "burden" of excluding persons from a benefit. *See* text accompanying note 273 *infra*.

237. *See* note 37 *supra*. In terms of the model, substantive due process compares only A to C:



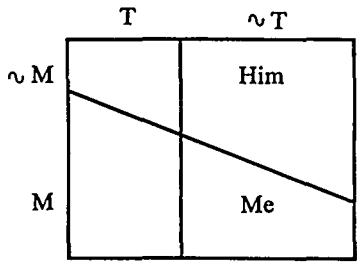
238. Equal protection compares the ratio A/C to the ratio B/D.

239. In terms of the model, the substantive due process claim emphasizes the *difference* in terms of harm or need. (T represents those who receive a burden or a benefit.)



(Burden)

or



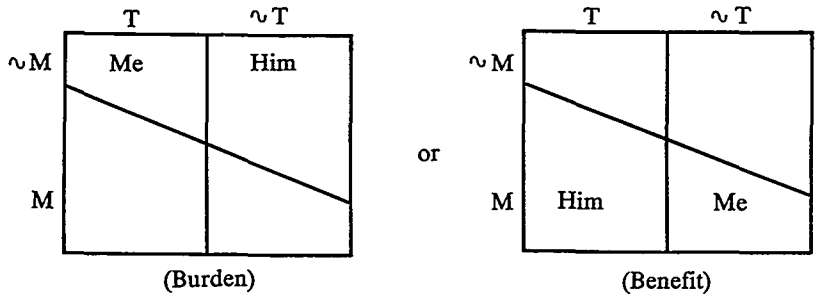
(Benefit)

the same way it treats him, since we are no different."[240] The essence of a due process claim is that equal treatment is unjustified; the essence of an equal protection claim is that dissimilar treatment is unjustified.

Since different kinds of misfit are involved, a classification that creates excessive due process misfit does not necessarily create excessive equal protection misfit, and vice versa. A few examples will illuminate these relationships. A situation presenting excessive due process misfit but not equal protection misfit is suggested by the "zebra killer" incident in San Francisco.[241] Knowing only that a suspected killer was a black man of medium build, the police made sweeping detentions of blacks. Detaining blacks and not whites was not "irrational"— *i.e.*, this is not a case of equal protection misfit—but the broad interference with the freedom of blacks may well have been too high a price to pay for the chance that one of them could be identified as the killer.

The case of a misfit which violates equal protection but not due process is even more common: almost every recent decision that a law violated equal protection dealt with a law that would have easily passed due process scrutiny. Consider, for example, gender-based discriminations in benefit programs.[242] Although a presumption of dependence for one sex and not the other might violate equal protection, surely the state may employ a presumption for both sexes or for neither. In fact, unless an equal protection violation is based upon a

240. In terms of the model, the equal protection claim emphasizes the *similarity* in terms of harm or need. (T again represents those receiving a burden or a benefit.)



241. N.Y. Times, April 26, 1974, at 1, col. 6. The police questioned over 600 people during a period of about one week, looking for a black man, in his 20s or 30s, of medium build, and 5 feet, 9 inches to 6 feet tall. Blacks complained that almost all young black men except the very tall and very fat were liable to be stopped and questioned. *Id.* at 15, cols. 1-3.

242. *See, e.g.*, Califano v. Goldfarb, 430 U.S. 199 (1977); Weinberger v. Weisenfeld, 420 U.S. 636 (1975); Frontiero v. Richardson, 411 U.S. 677 (1973).

suspect classification, it would be unwise for a court to rest on equal protection a decision which could have been decided on substantive due process, since the state might in good faith respond to the decision by imposing the invalidated restriction uniformly.

Finally, consider the old substantive due process case, *Lochner v. New York*,[243] in which the Court discussed both kinds of misfit. The Court found "irrational" a statute forbidding employment in a bakery for more than sixty hours a week or ten hours a day. First, the Court asserted that other occupations were just as "unhealthy" as the trade of a baker but were not regulated.[244] This is an equal protection argument, though the Court did not so denominate it. Second, the Court in effect argued that the law was overbroad: that an occupation is not "perfectly healthy" does not justify a sweeping infringement of freedom of contract.[245] This is a due process argument.

Of course, *Lochner* is not good law, for we now would reject both the equal protection argument (bakeries probably differ in kind from law offices, carpentry shops, etc.) and the due process argument (courts no longer value contract rights so dearly and they are properly quicker to accept legislative evaluations of the significance of social evils). Nonetheless, the case illustrates the differences between equal protection and substantive due process misfit.

The independence of due process and equal protection misfit suggests a problem: should a classification that barely satisfies each type of rationality standard ever be invalidated because of its cumulative "irrationality"? Yes, since each kind of misfit aggravates the general unfairness of a law.[246] A detailed discussion of *Craig v.*

243. 198 U.S. 45 (1905).

244. 198 U.S. at 59.

245. 198 U.S. at 59-60.

246. Doctrinally, this might seem incorrect, for we would not invalidate a law which only barely satisfied any other independent constitutional provisions, such as the first amendment and the privilege against self-incrimination. However, a similar kind of misfit analysis is now conducted under the due process and equal protection clauses, and that analysis relies more on general reasoning about the requirements of equality and fairness than on textual exegesis of the Constitution. It would therefore be odd to ignore cumulative harms for such doctrinal reasons. *Cf.* United States Dept. of Agriculture v. Murry, 413 U.S. 508, 517, 518, 519 (1973) (Marshall, J., concurring) (defending the irrebuttable presumption doctrine) (citations omitted):

> It is a corollary of [the Equal Protection] requirement that, in order to determine whether persons are indeed similarly situated, "such procedural protections as the particular situation demands" must be provided . . . . [W]here the private interests affected are very important and the governmental interest can be promoted without much difficulty by a well-designed hearing procedure, the Due Process Clause requires the Government to act on an individualized basis, with general propositions serving only as rebuttable presumptions or other burden-shifting devices.
>
> . . . .

Boren[247] demonstrates the value of the cumulative approach.

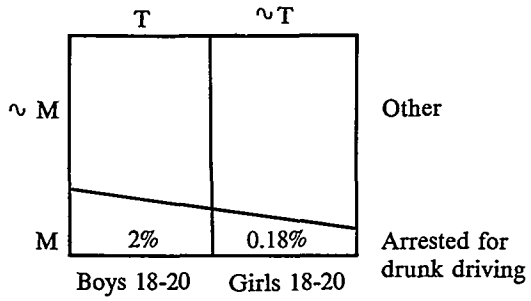
In *Craig*, the Supreme Court examined an Oklahoma statute that prohibited the sale of 3.2% beer to males under 21 and females under 18. The state defended the law by arguing that young males were more likely than young females to drive while drunk. A survey showed that 2% of males aged 18-20, but only .18% of females of that age, were arrested for driving under the influence of alcohol.[248] The Court nevertheless invalidated the classification: "if maleness is to serve as a proxy for drinking and driving, a correlation of 2% must be considered an unduly tenuous 'fit.' "[249] Justice Rehnquist, dissenting, deplored this use of the statistics, arguing that the only equal protection issue was whether the 2%/.18% differential justified the disparate treatment, and concluding that it did.[250] He maintained that the Court's reference to the 2% correlation was relevant not to equal protection, but to due process—"whether there are enough persons in the category who drive while drunk to justify a bar against purchases by all members of the group."[251] If the Court's reasoning were correct, he continued, it would entail the absurd consequence that the state could not validly bar 18-20 year olds, male or female, from purchasing beer, since only 1% (not even 2%) actually were arrested for driving while drunk. Justice Rehnquist balanced the state's interest in preventing drunk driving by young adults

---

This analysis, of course, combines elements traditionally invoked in what are usually treated as distinct classes of cases, involving due process and equal protection. But the elements of fairness should not be so rigidly cabined.

247. 429 U.S. 190 (1976).

248. In terms of the model, we have:



The percentages 2% and 0.18% refer to the percentages of T and of ∿T, respectively.

249. 429 U.S. at 201-02.

250. 429 U.S. at 224, 226.

251. 429 U.S. at 226.

against the injury to personal interests caused by any over- or under-inclusiveness and concluded that due process was not violated.[252]

Justice Rehnquist's dissent illuminates the distinction between differential harm (equal protection) and overbreadth (due process) misfit.[253] His contention that as a due process issue a 1% or 2% fit is adequate seems correct, given the relative insignificance of the injured personal interests.[254] But *Craig* might be a case in which a due process "fit" is satisfactory but is so weak that the classification must fail because of the cumulative misfit. If the elevenfold differential between men and women were maintained[255] but the due process misfit were reduced—*e.g.*, if 80% of young males were arrested for driving under the influence and 7.2% of young females—the Court would uphold the classification. However, since the degree of equal protection misfit is unchanged, the difference in result must be due to the weakness of the due process fit in the actual case. Of course, the Court in *Craig* was strongly influenced by the fact that the line was drawn on the basis of sex, and the Court was remarkably skeptical of the state's statistical arguments. Nevertheless, the hypothetical suggests that the due process misfit evaluation was necessary to the *Craig* result.

Craig illustrates another point: fewer justifications will be available to explain overbreadth in a case presenting both equal protection and due process claims than in a case presenting only a due process claim. For the equal protection claim, overbreadth misfit will be evaluated with respect to an *equal protection purpose*, that is, a purpose invoked to justify *differential* treatment and only indirectly invoked to justify the overinclusion. For the due process claim, the overbreadth will be evaluated with respect to a purpose that is directly invoked to justify the law. Another look at *Craig* shows why this distinction is important. More frequent drunk driving by males was cited as the equal protection justification for the law. While the purpose of reducing drunk driving perhaps adequately explained the differential treatment, the Court in effect invalidated the law because

252. 429 U.S. at 226-27.

253. *See* 429 U.S. at 225-26.

254. Note that a contrary result might forbid any town from remaining dry, if the rate of arrests for drunkenness at all ages were as low as 2%.

255. Of course, one may question whether maintaining the same C/A:D/B ratio is equivalent to keeping the equal protection differential "constant." Some weight might also be given to the absolute size of the percentages. Thus, if 1% of all boys and .5% of all girls drove while drunk, and if in another situation 100% of all boys and 50% of all girls caused or posed a certain mischief, we might find the latter misfit intolerable and not the former, even though the differential was "constant."

that purpose inadequately explained the due process overinclusion. Yet, had the state not been restricted to the kinds of arguments that would also justify gender-based discriminations, it could have marshalled many additional arguments to justify a uniform age limit for both sexes. For example, the state could have argued a uniform age limit is sensible because young people can only handle liquor when they reach a certain level of maturity; because peer pressures among the young foster excessive drinking; or because beer-drinking young people are likely to form boisterous public crowds, litter streets, and annoy bystanders. Since none of these is likely to justify constitutionally a sex differential—are girls sufficiently more mature than boys, less subject to peer pressure, or less likely to form public crowds?—they were unavailable to the state because they would have negated the equal protection defense.

In short, we should not readily assume (with Justice Rehnquist) that if a law's overbreadth does not violate substantive due process, that overbreadth has no equal protection significance. As we have just seen, an equal protection claim may impel the state to adduce a justification that defeats that claim but which creates serious due process problems, even where a good due process justification would have been available but for the equal protection claim. Moreover, Justice Rehnquist's conclusion ignores the possibility of a cumulative approach to due process and equal protection misfit.

A final substantive due process issue should be noted. When we describe a law as overbroad, we usually mean that it burdens the innocent as well as the guilty or excludes the needy as well as the undeserving from a benefit. But courts have occasionally asked whether the burden imposed actually reduces the harm. Even if overbreadth passes the traditional test—even if the class the legislature chooses to burden is sufficiently harmful to justify treating the entire class uniformly—the law may still fail if the burden the state has imposed on the class barely reduces the perceived harm. In *Lochner*, for example, the Court made not only the traditional overbreadth argument that there were too few unhealthy workers in bakeries to justify any health regulation. It also made this "effectiveness" argument: since New York health inspections already protected the health of bakery workers "so far as possible" the regulation of working hours did not in fact (or intention) reduce health problems in bakeries.[256]

The Supreme Court has not confined the argument that a burden is ineffectual to old and discredited substantive due process cases. In *Craig*, for example, the Court concluded that forbidding young

256. 198 U.S. 45, 61-62 (1905).

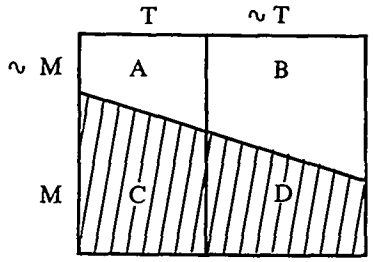
males to buy beer did not rationally serve the interest in reducing drunk driving. The Court reasoned that the statute did *not* forbid young males from drinking the beer purchased by young female friends or by older persons[257] and that the statistics were generally "lacking in controls necessary for appraisal of the actual effectiveness of the male 3.2% beer prohibition."[258]

The effectiveness argument is perhaps the argument implicit in the test: "Does the means significantly promote the end?" If "means" only describes how much of the total harm is *addressed* by the law[259] (in *Craig*, the proportion of all drunk driving by young adults that the regulation purported to reach), then the "promotion" test is incorrect. That males make up only 61%, or even only 2%, of all drunk drivers is irrelevant—equal protection only demands that males be sufficiently more likely than females to drive while drunk.[260] However, if the "promotion" test simply restates the effec-

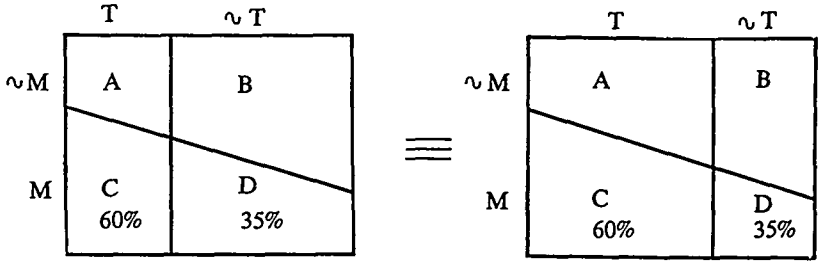
257. 429 U.S. 190, 204 (1976).

258. 429 U.S. at 202 n.14. The concurring opinions of Justices Powell and Stevens make the same argument. 429 U.S. at 211, 213 n.6.

259. In terms of the model, the shaded portions (regions C and D) represent the total harm, and C/(C+D) represents the portion of the harm addressed by the law:



260. In graphic terms, the location of the vertical line dividing T and ∿T is entirely irrelevant to equal protection analysis, so long as the proportions C/A and D/B remain constant.



The percentages represent the proportion of the vertical column that the lower segment constitutes. For example, C is 60% of C+A.

tiveness argument, then it is conceptually valid under both equal protection and due process. Had Oklahoma defended its regulation in *Craig* by asserting that the law reduced the incidence of rape, the Court's reaction is easy to imagine. Young males, of course, may pose the danger almost uniquely, and thus as an abstract matter the group singled out for special treatment is a proper subject for regulation. But the burden imposed has no apparent relevance to the asserted harm.

Although the effectiveness argument is conceptually sound, courts should use it sparingly. Where the burden chosen by the legislature *might* reasonably reduce the harm, the court should not ordinarily analyze its actual effectiveness, since that is a consummately legislative task. If a harm exists and if the class selected is a reasonably good surrogate for that harm, then the method and vigor of the attack on the harm should usually be left to legislative choice. This approach respects legislative flexibility where respect is most warranted.

The Court in *Craig* paid lip-service to this conception of institutional roles. After noting the shortcomings of the statistical sample and asserting that neither the judiciary nor state officials should be expected to be adept in statistical techniques, the Court concluded that "this merely illustrates that proving broad sociological propositions by statistics is a dubious business, and one that inevitably is in tension with the normative philosophy that underlies the Equal Protection Clause."[261] Underlying this statement is the important truth that ultimately the *Court* cannot test the effectiveness of a regulation or the seriousness of the harm which assertedly necessitates it. However, courts should not ignore evidence that the legislature's factual assumptions are *valid*, and the Court is incorrect if its statement suggests otherwise. On the other hand, the statement is quite right if it is only insisting that the Court, and not the statistics, ultimately decide *normative* equal protection questions. Even if it accepts the state's statistical proof, as it normally should, the Court must judge the significance of the facts proved. What personal interests are burdened by the law, and how much constitutional protection do they merit? How significant and legitimate are the state interests asserted? How much differential in harm will defeat an equal protection claim? How much overbreadth violates due process? What classificatory traits have stigmatizing effects which the Court should recognize? All these questions are independent of the validity of statistical proof, and all are questions that only a Court should answer.

The Supreme Court's standard for what *procedural* process is

261. 429 U.S. at 204.

"due," like the substantive due process standard, bears a noticeable similarity to equal protection tests:

> [I]dentification of the specific dictates of [procedural] due process generally requires consideration of three distinct factors: first, the private interest that will be affected by the official action; second, the risk of an erroneous deprivation of such interest through the procedures used, and the probable value, if any, of additional or substitute procedural safeguards; and finally, the Government's interest, including the function involved and the fiscal and administrative burdens that the additional or substitute procedural requirement would entail.[262]

These factors determine whether a hearing is required, either before or after the deprivation of liberty or property,[263] and how thorough that hearing should be.[264] The substantive deprivation itself, of course, is not the subject of review.

Procedural due process addresses not what the government may do, but how it may do it. Because purpose is not directly considered, misfit analysis might seem entirely inapposite. But the distinction between "what" and "how" is analogous to the substantive due process/equal protection distinction between ends and means. In substantive due process and equal protection the courts are often less concerned with the state's ends than with whether the means fit them. Thus, conceptually, procedural due process might be seen as a lower tier of means-end misfit analysis in which the substantive "means" are treated as the procedural "end." Misfit analysis in the lower tier compares the procedural means to the procedural end, evaluating the extent to which the procedures used accurately select the class described in the statute. This contrasts with misfit analysis in the higher tier (of substantive due process/equal protection), which evaluates the extent to which the statutory class accurately reflects the class that *would* be picked out if the statute were to achieve its goal perfectly.[265]

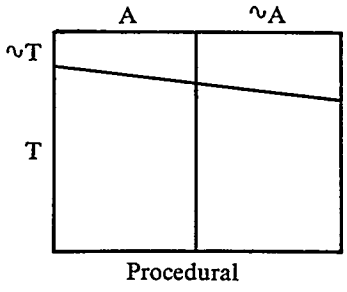
---

262. Mathews v. Eldridge, 424 U.S. 319, 335 (1976).

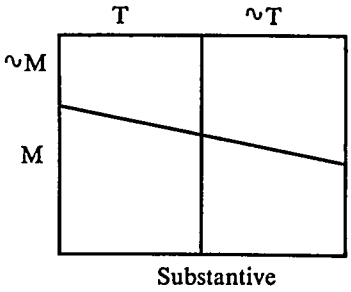
263. *See* L. TRIBE, *supra* note 100, at 543-50 (1978).

264. *See* Goss v. Lopez, 419 U.S. 565, 579 (1975).

265. In terms of the model, we may define a third class A, comprising those actually burdened or benefitted by the government. We then simply have two misfit evaluations, one procedural, the other substantive.



Procedural



Substantive

Procedural due process, like substantive due process and equal protection, promotes more than accuracy.[266] Besides protecting classificatory accuracy, equal protection, through the doctrines of suspect categories and of illicit motivation, proscribes stigmatic and hostile treatment.[267] Procedural due process allows individuals to present their claims to state officials and requires those officials to explain their actions, functions which have value even if they do not assure fewer erroneous decisions.[268] Nevertheless, both procedural and substantive misfit are important elements of due process and equal protection analysis and can, as we have seen, be integrated into a generalized two-tiered test of classificatory fairness.

Finally, we should briefly note the role misfit plays in the irrebuttable presumption doctrine. Irrebuttable presumptions are usually substantive due process cases of a special kind, *not* disguised equal protection cases. When the Court held that Illinois could not take away Stanley's motherless children without letting him rebut the presumption that he was an unfit parent,[269] it did so not because fathers are generally as fit as mothers to be a family's only parent (they may well not be), but because *many* fathers will be fit parents. Whatever the talents of widows, widowers should have the chance to exercise the important personal interest in child-rearing, since a significant number of men (even if a minority) can exercise it well. Thus, to the extent that misfit analysis is relevant to the irrebuttable presumption doctrine, substantive due process overbreadth misfit is the most apposite kind.

Many irrebuttable presumption cases do resemble equal protection cases. A claim that the state irrebuttably presumed that Stanley was unlike the average widow sounds like an equal protection argument. As suggested above, however, widows' competence is probably irrelevant to the result in *Stanley*. The result would have been the same had the state automatically taken children away from a surviving parent of *either* sex. That irrebuttable presumptions are actually due process cases is confirmed by the difficulty of imagining an irrebuttable presumption describable *only* in equal protection terms.

Procedural misfit, like substantive, can take the form of both overinclusion (*e.g.*, a girl is mistakenly "burdened" and not allowed to buy beer, in *Craig*) and underinclusion (*e.g.*, some 20-year-old boys pass for 21).

266. *See generally* the thoughtful analysis in L. TRIBE, *supra* note 100, at 501-06.

267. *See generally* Part I.A *supra*.

268. *See* note 223 *supra*.

269. Stanley v. Illinois, 405 U.S. 645 (1972).

This section concludes with a hypothetical case that illustrates the characteristics of the various forms of due process and equal protection. Assume that a city police department has a policy of dismissing from the force officers whom it discovers are practicing homosexuals, and that the plaintiff is dismissed for that reason. The constitutional arguments available to him will reflect concern with both legitimacy of purpose and degree of misfit,[270] and can be summarized as follows:

(1) Procedural due process

A hearing is required to determine whether the plaintiff is a practicing homosexual, for existing departmental procedures do not reliably determine that fact.

(2) Irrebuttable presumption

The departmental policy is obviously designed to ensure public role-models who share the community's values and, more specifically, who will not encourage homosexuality. Although the plaintiff is a practicing homosexual, the department cannot automatically assume that he is a poor role-model. A hearing is required to determine whether his private life actually reflects upon his ability to uphold community values in public.

(3) Substantive due process

The departmental policy has an unlawful purpose, namely, prohibiting persons from exercising personal-privacy interests and prohibiting them from expressing those privacy interests in public.

Even if the end served is not illegitimate, the departmental policy is overbroad, for not all practicing homosexuals are poor role-models. There is a reasonable, less restrictive alternative: dismissing only those homosexual policemen who publicize their preferences.

The plaintiff admits being a practicing homosexual. But because he does not publicize this fact, the departmental policy should be invalidated for its overbroad interference with his privacy.

(4) Equal protection

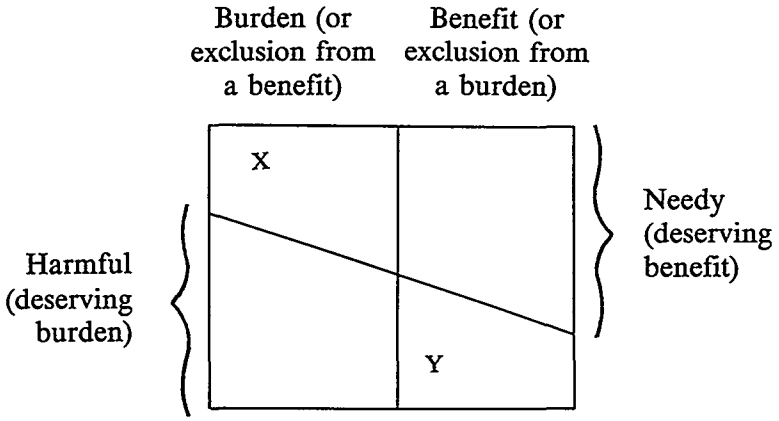
The departmental policy has an illegitimate purpose. But even if the end is not illegitimate, it is irrationally served here, for a similar proportion of nonhomosexuals as homosexuals are poor role-models. Furthermore, it is easy enough to identify those homosexuals who *are* a problem, by prohibiting all homosexuals from publicizing their sexual preferences. Given the availability of this convenient, less restrictive alternative, the doubtful difference in harm posed by homosexuals compared to nonhomosexuals, and the law's substantial interference with the plaintiff's privacy, this law should be invalidated.

270. Some of the arguments refer to the availability of less restrictive alternatives, an issue discussed in Part VI.B *infra*.

2. *Overinclusion and Underinclusion; Burdens and Benefits*

Courts and commentators occasionally suggest that equal protection should treat overinclusion differently from underinclusion,[271] and benefits differently from burdens.[272] In light of the parallels between due process and equal protection misfit, these suggestions should be examined with respect to the due process clause as well. This section contends that no analytic distinction generally need be made between overinclusion and underinclusion or benefits and burdens when evaluating misfit. The primary exception arises because of the due process "deprivation" requirement, but the irrebuttable presumption doctrine helps restore the symmetry.

The following composite diagram may be helpful:



As this diagram suggests, we can interdefine burdens and benefits and overinclusion and underinclusion. Region X represents overinclusion with respect to a burden. But if we describe exclusion from a benefit as a burden, then X also represents underinclusion with respect to a benefit. For analogous reasons, Y represents either underinclusion with respect to a burden or overinclusion with respect to a benefit.[273] With these relationships in mind, we can consider in

271. *See* Tussman & tenBroek, *supra* note 27, at 348-53; *Developments, supra* note 1, at 1084-87; note 284 *infra*.

272. *See* Maher v. Roe, 432 U.S. 464 (1977), and Nashville Gas Co. v. Satty, 434 U.S. 136 (1977), *discussed in* note 288 *infra*.

273. Under these definitions, the labels "harmful" and "needy" are imprecise, and the parenthetical labels "deserving burden" and "deserving benefit" are more accurate.

more detail first the relationship of overinclusion and underinclusion to misfit analysis and then the role of burdens and benefits.

One commentator has recommended that courts approach overinclusion and underinclusion (with respect to a burden) with different degrees of tolerance. Overinclusion is more offensive because it burdens the "innocent" while underinclusion merely fails to burden the "guilty," and because the need-for-experimentation rationale that partly justifies underinclusion[274] is inapposite to overinclusion.[275] The commentator also notes that burden-overinclusion, while more objectionable when it does arise, is less common than underinclusion, "since the greater numbers who are adversely affected by the action will bring political pressure to bear on the lawmakers."[276]

This reasoning does not dictate any approach to underinclusion and overinclusion with respect to a *benefit*.[277] In the case of benefits, overinclusion may seem less objectionable than underinclusion, since the latter fails to benefit the deserving, and the former benefits the undeserving.[278] On the other hand, this "preference" for overinclusion is undercut by the experimentation rationale. Just as the advantages of legislative flexibility in dealing with the "harmful" may justify burden-underinclusion, so the wish to deal flexibly with the "needy" may justify benefit-underinclusion; in each case the state has decided to extend its efforts to only *some* of the "deserving." And again, *under*inclusion with respect to a benefit might be less likely to occur than overinclusion, since the deserving who are denied benefits may mobilize political pressure.[279]

While there may be some good reasons for different judicial tolerances toward overinclusion and underinclusion (and for differentiating further between benefits and burdens), this Note suggests a somewhat different analysis, the elements of which have already been introduced. Under this analysis, the proper judicial approach

---

274. *See* text at note 339 *infra*.

275. *Developments, supra* note 1, at 1086.

276. *Id.* at 1086-87. *See* Railway Express Agency v. New York, 336 U.S. 106, 112-13 (1949) (Jackson, J., concurring).

277. *See Developments, supra* note 1, at 1086 n.47.

278. Which of these is actually less objectionable is a question which can only be answered intuitively. While it seems clearly preferable to fail to burden a guilty person rather than to burden an innocent person, the correct preference is less obvious for benefits. It seems "better," though debatably so, to benefit a non-needy person (we are only wasting our money) than to fail to benefit a needy person (we are allowing the need to continue unabated). In any event, these broad generalizations are of only limited value, for as will be suggested in the text at notes 280-87 *infra*, overinclusion and underinclusion can for the most part be similarly treated under the equal protection and due process clauses.

279. Of course, as a practical matter, these people may lack political power.

is essentially as follows. If a court has evaluated a classification in terms of the minimum rationality constraint and the equal protection balance, then underinclusion (with respect to either a burden or denial of a benefit) is of no further judicial concern. But the Court's concern with overinclusion goes beyond this equal protection analysis. Judicial review of overinclusion is complete only when the classification has also been tested for due process overbreadth. Benefits and burdens are treated similarly under equal protection except insofar as they affect the significance of the personal interest.[280] But under due process, a special problem is posed by benefits in which the plaintiff does not have a property interest. Overinclusion analysis cannot, strictly speaking, apply to the denial of such benefits, for those who are overinclusively denied the benefit have no due process claim to it. Although such claims may be reviewed under the irrebuttable presumption doctrine, they are reviewed more deferentially than claims of overinclusion with respect either to burdens or to the denial of "property benefits."

This proposed analysis follows directly from the earlier discussion of due process and equal protection misfit. Simply to evaluate the "amount" of overinclusion or underinclusion is meaningless in equal protection analysis; the size of these classes is relevant only in comparison with the "properly fitted" classes of the properly included and properly excluded.[281] In equal protection analysis, the differential misfit approach compares the proportion of the included who are harmful to the proportion of the excluded who are harmful and validates the law if the comparison is both minimally rational (*i.e.*, if there is *some* positive differential) and "on balance" fair (*i.e.*, if the differential is sufficient in light of the governmental and personal interests involved).[282] In contrast, substantive due process analysis asks whether the class of overincluded is "on balance" too large compared to the class of properly included.[283]

Thus, insofar as underinclusion and overinclusion are both probative of differential misfit, courts should scrutinize them with equal tolerance. But since only overinclusion, not underinclusion, is relevant to substantive due process, we might characterize judicial review of misfit as less tolerant of overinclusion. As long as a classification creates a sufficient differential in harm, underinclusion needs no justification. If drivers under age 18 are sufficiently more likely to be drunk than older drivers, then they can be prohibited from purchasing alcohol, whether they are 15%, 5%, or only 1% of all

280. *See* text at notes 288-96 *infra*.

281. *See* note 260 *supra* and accompanying text.

282. The mechanics of equal protection balancing are discussed in Part VI *infra*.

283. *See* Part VI.C *infra*.

drunk drivers. This feature of equal protection may have impelled courts to declare that the state may go "one step at a time" and may attack only "part of a harm."[284] While this declaration is entirely sound, it should not be confused with the incorrect view that the state may attack only part of the harm *regardless* of the relative acuteness of the harms it has and has not addressed.

It might appear, thus, that courts err in discussing underinclusion in any substantive due process cases[285] and in many equal protection cases.[286] But there is a rational motive for this judicial practice, though it may not be the actual motive. Underinclusion *is* indirectly relevant to both types of misfit, for it helps define the significance of the state interest actually served by the law. For example, if only 5% of all drunk drivers are under 18, then this underinclusiveness suggests that the state interest in "reducing drunk driving" should be discounted by 95% in any equal protection or substantive due process balance.[287] Of course, a mathematical approach is impractical. Nevertheless, the Court may validly consider, with

284. As the Court has rather broadly put it: "[M]ere underinclusiveness is not fatal to the validity of a law under the equal protection component of the Fifth Amendment, even if the law disadvantages an individual or identifiable members of a group." Nixon v. Administrator of Gen. Servs., 433 U.S. 425, 471 n.33 (1977) (citations omitted). *See* notes 59-65 *supra* and accompanying text.

285. For an example of this judicial practice, see Moore v. City of East Cleveland, 431 U.S. 494 (1977). In arguing that the state interests asserted for a zoning ordinance were at best "marginally" served, the plurality specifically pointed to the underinclusiveness of the ordinance (as well as its overinclusiveness). For example, the ordinance was underinclusive in remedying traffic and parking congestion since some who deserved the burden were not burdened: "the ordinance permits any family consisting only of husband, wife, and unmarried children to live together, even if the family contains a half dozen licensed drivers, each with his or her own car." 431 U.S. at 500 (opinion of Powell, J.). The mischief of overcrowding was also remedied underinclusively, since "[t]he ordinance would permit a grandmother to live with a single dependent son and children, even if his school-age children number a dozen. . . ." 431 U.S. at 500.

286. For example, in Jimenez v. Weinberger, 417 U.S. 628 (1974), the Court pointed to both the overinclusiveness and underinclusiveness of a statutory provision denying benefits to certain classes of illegitimate children of disabled claimants. 417 U.S. at 637. This Note suggests that underinclusiveness is, strictly speaking, irrelevant. Here, the fact that the burden (of being denied a benefit) was conferred underinclusively—*i.e.*, that some who did not deserve the benefit obtained it—is irrelevant to the due process overbreadth claim that many "innocent" persons who deserved the benefit failed to obtain it. And it is not directly relevant to the equal protection claim that the excluded and included classes of illegitimates were equally deserving, and "the potential for spurious claims [was] exactly the same as to both subclasses," 417 U.S. at 636, since the number of undeserving recipients is only of equal protection consequence if it affects the extent of the differential in need between the excluded and the included. If that differential has already been evaluated, then the number of undeserving recipients is irrelevant. (What this Note describes as "underinclusion" with respect to the denial of a benefit the Court describes as "overinclusion" with respect to a benefit. Though awkward, this terminology is necessary for a consistent interpretation of underinclusion for both benefits and burdens.)

287. In Moore v. City of East Cleveland, 431 U.S. 494 (1977), for example, the fact that the congestion and overcrowding problems were underinclusively remedied suggests that in any equal protection balance the state's interest in reducing these problems should be discounted. *See* note 285 *supra*.

whatever precision is possible, the extent to which a social interest is actually being served.

Finally, the treatment of burdens and benefits under equal protection and due process must be explained. Equal protection presents few problems here, for benefits differ from burdens only to the extent that they implicate different personal interests, or implicate them to a different degree, in particular cases.[288] As a general

288. A few comments about the relationship between personal interests and burdens or benefits are in order. One interesting question is whether an interest should be evaluated differently if the burden or benefit represents a change from the status quo. Is it worse to deprive a group of persons of welfare benefits they have been receiving than to exclude them from a program at its inception? Is it worse to impose a new criminal penalty on group A than to exempt all but A from an existing penalty? The answer in each case should be yes, because of the significant difference in personal expectations. Because of past reliance and the interruption of living patterns, statutes that impose a new burden or withdraw an old benefit should be subject to special scrutiny.

This argument should not be taken too far, however. It certainly does not follow that because persons are only excluded from a new benefit, their interests in receiving the benefit are inconsequential. However, the Supreme Court in Maher v. Roe, 432 U.S. 464 (1977), seemed to reason in that fashion. In upholding a state welfare program's exclusion of nontherapeutic abortions from medical coverage (an exclusion best described as an "exclusion of a new benefit," since recipients had not been "receiving" it continuously), the Court reasoned in part:

> The Connecticut regulation places no obstacles—absolute or otherwise—in the pregnant woman's path to an abortion. An indigent woman who desires an abortion suffers no disadvantage as a consequence of Connecticut's decision to fund childbirth; she continues as before to be dependent on private sources for the services she desires.

432 U.S. at 474. But this argument proves too much; we would not say that women suffer "no disadvantage" if a state welfare program were initiated but its benefits were limited to men. After all, it is only *relative* disadvantage that the equal protection clause addresses.

The above discussion does not necessarily imply that benefits and burdens generally require different treatment. Denial of even an existing benefit, *e.g.*, a government transfer payment, might seem less onerous than imposition of a burden, *e.g.*, a criminal sanction. Whether a benefit is conferred at all is a discretionary matter, while our intuition about burdens seems to be quite the contrary; we do not feel "privileged" whenever the government *fails* to punish our conduct. But to give judicial recognition to this vague intuition would threaten to resurrect the discredited "right/privilege" distinction. *See generally* Van Alstyne, *The Demise of the Right-Privilege Distinction in Constitutional Law*, 81 HARV. L. REV. 1439 (1968). If there is a difference between the proper valuation of benefits and burdens, it is that the personal interests and opportunities threatened by unequal treatment will often be greater for burdens than for benefits. Generally the burdens imposed by the criminal process are qualitatively more severe than other affirmative burdens (such as taxes) and than the government's failure to confer a benefit. *See Maher*, 432 U.S. at 475:

> There is a basic difference between direct state interference with a protected activity and state encouragement of an alternative activity consonant with legislative policy. Constitutional concerns are greatest when the State attempts to impose its will by force of law; the State's power to encourage actions deemed to be in the public interest is necessarily far broader.

In Nashville Gas Co. v. Satty, 434 U.S. 136 (1977), the Court held that the employer's practice of denying accumulated seniority to a woman who had, as the employer required, taken a leave of absence during pregnancy violated Title VII, 42 U.S.C. §§ 2000e to 2000e-17 (Supp. V 1975), but that the employer's practice of denying such a woman sick pay did not, per se, violate that law. The Court emphasized the distinction between benefits and burdens, 434 U.S. at 142. Justice Stevens' concurrence convincingly argues that that distinction does not provide a meaningful test of discrimination. 434 U.S. at 154 n.4. The test this Note has suggested could nevertheless achieve the same result that the majority reached. Denying accumulated seniority under the facts of *Satty* is less tolerable than denying sick pay—but not because denial of seniority is a "burden" and denial of sick pay is not (how is a woman "bur-

matter, since equal protection applies to *all* burdens and benefits, we can conveniently interdefine them and apply a uniform "underinclusion" analysis to ordinary burdens and to the "burden" of exclusion from a benefit.

It is more difficult to apply a uniform "overinclusion" analysis under substantive due process. Overinclusion probably applies to all affirmative government burdens, such as taxes or criminal penalties, for the due process clause explicitly protects against *deprivations* of life, liberty, or property, and "liberty" has always been expansively interpreted.[289] But we cannot simply define exclusion from any government benefit as a burden in order to make the overinclusion analysis complete. A plaintiff who has *never received* a government benefit he deserves has no obvious substantive due process claim, for he must have been *deprived* of a property interest, not simply prevented from obtaining it.[290]

Of course, substantive due process does protect against arbitrary or overbroad exclusions from benefits when the plaintiff already *has* a property interest in the benefit. Thus, for example, even the property interest of a welfare recipient should be sufficient to raise a substantive due process claim if benefits were arbitrarily terminated or reduced for all recipients.[291] And, of course, even where a plaintiff has no property interest, he may have a viable equal protection claim which would remedy some of the law's unfairness (aside from its overbreadth). Nevertheless, misfit analysis contains a serious asymmetry: people who have been denied participation in a benefit program in which they do not have a due process "property" interest cannot complain about overbroad criteria for their exclusion. For

---

dened" by agreeing to take a job with limited seniority rights?). It is because the first denial is a more severe intrusion on the woman's employment interest and is less proportionate to any rational interests of the employer.

Of course, there is a serious definitional problem with the majority's approach: excuse from any burden can be called a "benefit," and exclusion from any benefit can be called a "burden." The distinction is at best intuitive—any government action restricting our freedom or pleasure from what it would "naturally" be is a burden, and any action increasing it is a benefit. But the concept of a "natural state" is itself vague. Fortunately, the Court can avoid these conceptual problems simply by inquiring whether a personal interest is significantly infringed.

289. *See* Meyer v. Nebraska, 262 U.S. 390, 399 (1923). *But see* Meachum v. Fano, 427 U.S. 215, 228 (1976).

290. One possible argument is that all citizens have a *liberty* interest in obtaining deserved government benefits, under the broad liberty definition of Meyer v. Nebraska, 262 U.S. 390 (1923). *See* Board of Regents v. Roth, 408 U.S. 564, 588-89 (1972) (Marshall, J., dissenting) (discussing procedural, not substantive, due process). However, such an argument seems to read "liberty" too broadly, since the argument renders the fifth and fourteenth amendment language "of property" superfluous. Justice Marshall's suggestion that a property interest can be anticipatory—that any applicant for a government job has a property interest in that job, 408 U.S. at 588—has not been endorsed by other members of the Court.

291. *Cf.* Goldberg v. Kelly, 397 U.S. 254 (1970) (requiring *procedural* due process protections before recipients are deprived of their welfare benefits).

example, although the policemen in *Kelley v. Johnson*[292] ultimately lost their due process argument that policemen may not be fired for having long hair, they could at least get into court. But a police-force applicant would have no property interest in the denied benefit of government employment.

Those who treasure conceptual completeness may bemoan this state of affairs, but they may call upon the irrebuttable presumption doctrine. That doctrine was earlier described as essentially a substantive due process doctrine under which overbreadth is partially remedied by affording plaintiff a chance to rebut the presumption of "desert" of a burden or "nondesert" of a benefit. The presumption of nondesert of a benefit is subject to this doctrine even if the plaintiff has no property interest in the benefit. For example, Mrs. Salfi claimed that the Social Security Administration could not irrebuttably presume from her husband's death soon after their marriage that she married him only to obtain benefits.[293] Although the Court rejected her claim, it did consider its merits, and the opinion does not foreclose the possibility that a less rational law would fail. While the Court emphasized that hers was merely a noncontractual claim to funds from the public treasury,[294] that emphasis goes only to the significance of the personal interest at stake, not to the cognizability of an irrebuttable presumption claim to benefits in which plaintiff lacked a property interest.[295]

Of course, the irrebuttable presumption doctrine only partially restores symmetry to misfit analysis, since that doctrine applies only when a statute clearly purports to test for a quality with respect to which a plaintiff could offer more convincing proof at a hearing.[296] Nevertheless, it moves the law closer to the ideal of providing a remedy for all forms of classificatory unfairness, whether premised upon overinclusion or underinclusion, burden or exclusion from a benefit.

D. *Multiple Goals*

The problem of multiple legislative goals, each importing a distinct mischief or need, deserves separate comment. Legislatures obviously may and do pursue many objectives in a single law, and therefore a court should not judge classifications according only to their primary purpose.[297] Of course, even if a law has multiple pur-

292. 425 U.S. 238 (1976).

293. Weinberger v. Salfi, 422 U.S. 749, 753 (1975).

294. Weinberger v. Salfi, 422 U.S. 749, 772 (1975).

295. *But see* Note, *The Conclusive Presumption Doctrine: Equal Process or Due Protection?*, 72 MICH. L. REV. 800, 816 (1974) (criticizing the Supreme Court for ignoring the "property" or "liberty" requirement in irrebuttable presumption cases).

296. *See* text at notes 96-97 *supra*.

297. McGinnis v. Royster, 410 U.S. 263, 276-77 (1973). *See* Sandalow, *supra* note 7, at 660; Note, *supra* note 6, at 132-37.

poses, any given classification within it will often be justifiable in terms of only one of these purposes. Because of the peculiar problems it presents for misfit analysis, we will discuss the special case of a single classification that serves a set of goals. For the sake of simplicity and without loss of generality we will discuss a classification that serves *two* purposes.

A classification can serve two goals in two different ways. First, the law might burden a group that poses two mischiefs (or benefit a group that has two needs). For example, recidivist felons might be more severely punished than first-time felons and recividist misdemeanants. Second, the law might excuse from a burden a group that would otherwise deserve it because the group also deserves a benefit or, conversely, excuse from a benefit a needy group because that group also poses a harm. For example, if an ordinance requires drivers to stop at red lights, but exempts police and firemen, the combined purpose is reducing the risk of accidents to the extent that this does not interfere with emergency fire and police services.

The analysis of multiple purpose classifications might seem straightforward. In each of the cases just described, we could simply analyze misfit relative to the combined purpose class (the class of those possessing both mischiefs, or both needs, or the need despite the mischief, or the mischief despite the need). But why may we burden those who possess both harms, yet not burden those who possess only one? Justifications for doing so would require more than this simple misfit evaluation.

In our traffic light hypothetical, for example, in determining the propriety of burdening only the combined purpose class of those who both lack a special need for haste and pose the harm of being a traffic hazard, we cannot simply ask whether the exemption for police and firemen fits these combined goals. That is, it is not sufficient to conclude that enough of those who are burdened possess both harms—that enough drivers would jeopardize highway safety and would have no legitimate interest in speed—and thus that the misfit is tolerable. We must also ask *why* the state may exempt from this particular burden those who need to travel without interruption rather than excusing them from some other burden or providing them with some special benefit, such as faster vehicles or special roads.

That additional inquiry is hardly troubling in this example, since the harm and the need addressed by the law are obviously interrelated—imposing on emergency vehicles the burden of stopping at red lights directly affects their ability to move rapidly and without interruption. The kind of interrelationship necessary to satisfy minimum rationality will be discussed below. But possession of com-

pound mischiefs alone is not a sufficient reason for treating a class differently from a class which possesses only one mischief. If it were sufficient, then the "multiple goals" rationale could be subtly used to evade the minimum rationality constraint. An example should make this point clearer.

Levy v. Louisiana[298] invalidated a statute allowing legitimate but not illegitimate children to sue for the wrongful death of their mothers. The Supreme Court reasoned that both classes of children were similarly situated with respect to the loss and the "wrong" they suffered at a mother's death.[299] But the statute may have had a second purpose—to encourage the formalization of relationships between parents.[300] One commentator[301] has suggested that although neither purpose considered alone is adequately served by the classification, it *is* rationally related to the following combined purpose: compensating children for the wrongful death of their mothers to the extent that such compensation does not encourage relationships that have never been legally formalized.

This suggestion is intriguing, but it does not explain *why* the state may pursue its interest in marriage through the law of wrongful death. The situation would be different if a law denied illegitimate children all the rights of legitimate children *or* if there were a pressing need to formalize relationships that produced illegitimate offspring likely to sue for wrongful death. The absurdity of the latter possibility perhaps accounts for the Court's casual treatment of the interest in marriages.

This discussion of the multiple purposes in *Levy* suggests that though it may be rational to pursue goal A by burdening class T* and rational to pursue goal B by burdening class T**, it is not necessarily rational to pursue goals A *and* B by burdening only those persons common to T* and T**. That invalid syllogism circumvents the minimum rationality constraint and would, for example, justify enforcing trespassing laws only against illegitimate children, or stimulating music and shipbuilding by granting tax deductions only to musicians who buy yachts.[302] In effect, the syllogism rationalizes

298. 391 U.S. 68 (1968).

299. 391 U.S. at 72.

300. 391 U.S. at 80 (Harlan, J., dissenting).

301. Note, *supra* note 6, at 131, 136-37.

302. While the state might decide to subsidize musicians or to spur the yacht industry, the objectionable feature of this tax deduction is that it does neither fairly. However, the state could give a deduction to all musicians and a separate deduction to all yacht purchasers. Although musicians who buy yachts now gain the greatest benefit, this is fair since all who deserve either kind of benefit receive it.

A helpful analogy is the problem of how to "count" black women when assessing the success of affirmative action programs designed to increase the numbers of blacks *and* of women. If the program is really designed to achieve both goals, then it is improper to double-count black women toward both the black and the female hiring goals, since under this procedure a

differential treatment of persons similarly situated with respect to *one* harm or need by reference not to that harm or need, but to another harm. Yet differential treatment with respect to the second harm is justifiable only with reference to the first harm, and we thus have a vicious circle.[303]

Despite cases like *Levy*, classifications by compound mischiefs can usually be justified. Four such justifications will now be discussed.[304]

The first such category of classifications consists of those cases in which the presence of one mischief is associated with a higher degree of another. For example, while some have questioned the propriety of selectively prosecuting for tax violations those suspected of other kinds of criminal activity, it might be contended that such suspects are more likely to have cheated on their taxes than law-abiding citizens, and that devoting investigative resources to the former class is more efficient. It would be a different matter if the tax laws were *less* effective than ordinary sanctions, or if criminals were *less* likely to have cheated.

In a sense, this first category does not involve *multiple* goals at all. Vigorous use of the tax laws against criminals could be explained not as serving the twin goals of (1) raising revenue and (2) pursuing criminals, but as serving the single goal of raising revenue efficiently, *i.e.*, by concentrating on the class of criminals filing tax

---

company could hire no black men or white women, but only black women. It is also improper either to count all black women once as blacks, or to count them once as women, since the result could again fall short of the goal: if all black women were counted as "black," then the company would hire more white women and fewer black men than was intended. The correct procedure, rather, is to establish separate hiring goals for black men, black women, and white women. Of course, if black women were found to have suffered more severe discrimination than either black men or white women, then "double-counting" would be more rational. But if the program has the broader two-fold purpose noted above, the purpose is disserved if only black women are hired.

303. In *Levy*, distinguishing between legitimately and illegitimately conceived wrongful-death plaintiffs is justified in terms of the state's interest in formalizing marriage relationships. However, distinguishing between penalizing illegitimate wrongful-death complainants and failing to penalize most other illegitimates (or their parents) is justifiable, if at all, only in terms of an interest such as the interest in discouraging wrongful-death suits. But we need an explanation for burdening only illegitimates to serve that interest—the circle is complete.

304. This Note does not disapprove of all multiple-goal justifications. In particular, it does not subscribe to the disturbing "divide and conquer" technique which the Supreme Court has occasionally used to invalidate classification. Under this technique, the Court simply evaluates misfit with respect to each of several goals in turn. Because when it is evaluating one goal it is ignoring the other goals, the Court inevitably finds the classification irrational with respect to each goal. For example, in Smith v. Cahoon, 283 U.S. 553 (1931), the Court invalidated a highway bond requirement that applied to all commercial carriers except agricultural or seafood carriers. The Court found the law irrational because the exempted carriers posed the same danger to public safety as other carriers. Other plausible purposes, such as a desire to subsidize the exempted industries or to encourage their use of the roads (*see* note 306 *infra*), were simply ignored by the Court. *See* Ely, *supra* note 26, at 1225-26. This technique was also used in Eisenstadt v. Baird, 405 U.S. 438 (1972), a use effectively criticized in Note, *supra* note 6, at 124-28.

returns. In other words, the second goal need not be referred to at all if the burdened class more acutely poses the first harm. Almost *every* law could be seen as having "multiple goals" in some trivial sense—*e.g.*, we forbid the young from driving not because they are likely to be worse drivers, but because we are pursuing the goals of keeping unsafe drivers off the road and of restricting some privileges to the mature. Obviously, the first explanation is more realistic and more convincing.[305]

In the second category of cases, multiple goals can be justified because the pursuit of one goal inevitably affects the pursuit of another. Such goals are interdependent rather than independent. When goals are interdependent, the first goal will be furthered to a *different* degree when the second goal is not pursued at all, than if the second goal is pursued fully, without exemptions. Some examples should make this distinction clearer.

Cases presenting interdependent goals usually involve a goal which inhibits, rather than aids, the pursuit of another. An example of such a tradeoff has already been discussed: permitting firemen and police to ignore traffice lights is rational because the state can only pursue its interest in fire and police protection at the expense of some loss in traffic safety. This tradeoff rationale is entirely distinct from the first justification for pursuing multiple goals (that a high degree of one harm implies a higher [or lower] degree of another), since the validity of the traffic light exemption does *not* depend on the possibility that firemen and policemen are less hazardous than other motorists who ignore traffic signals.[306]

---

305. At first glance the "multiple goals" problem appears to be trivial, for if we define M as M* and T as M**, then the question whether the state may pursue two goals (M* and M**) simultaneously seems identical to the question whether the classification contains too much misfit (between M and T). However, T refers to those who are in fact burdened; it should be compared to the class that *should* be burdened, which in the present context is the intersection of the two goals, M* and M**, and not simply M*. If indeed there is a legitimate state interest in simply burdening or benefitting members of T, then this state interest will justify the classification, and there is no need to engage in subtler "multiple goals" analysis. For example, our driving test hypothetical could be forced into the "multiple goals" mold by asserting that the state is pursuing both the goal of driver safety (M or M*) and the goal of burdening persons who fail driving tests (T or M**). But the second "goal" is artificial; if it *were* legitimate, then there would be no need even to refer to the first goal.

306. Smith v. Cahoon, 283 U.S. 553 (1931), might be another example of interdependent goals. The Supreme Court invalidated a highway bond requirement that exempted only carriers of agricultural goods and certain seafoods, perhaps because it assumed that the law's only purpose was to improve traffic safety or perhaps because it considered highway bonds an inappropriate means through which to favor agriculture. *See* Note, *supra* note 6, at 134; Ely, *supra* note 26, at 1225-26. Indeed, under the interdependence analysis suggested in the text, such a compound purpose would be insufficient to justify the law. But if the state has a defensible interest in encouraging agricultural carriers to use the roads rather than other means of transportation, *e.g.*, to prevent products from perishing, then the classification might be justifiable, for the pursuit of this goal inevitably inhibits the goal of highway safety. In other words, the state should be permitted to make this tradeoff, since the two goals are interdependent.

This Note has also presented examples of purported tradeoffs which in fact do *not* involve interdependence. A law prohibiting only illegitimate children from maintaining wrongful death suits is described in "tradeoff" terms if the purpose is said to be to compensate children for the loss of a parent "to the extent that" this does not condone nonformalized relationships between the parents. However, these goals are independent, and the tradeoff is artificial, since the need for compensation is unrelated to any need for formal marriage. More precisely, it would not affect the state's formalization interest if *all* children could sue for wrongful death, or if *none* could. By contrast, strict enforcement of traffic laws would inhibit the goal of efficient fire and police protection more than lax enforcement would.[307]

If multiple goals are interdependent, then the state need not defend its decision to pursue both goals simultaneously or to pursue one goal at the expense of the other. Such decisions about the proper tradeoff between legitimate goals are classically the responsibility of the legislature.[308] They require a balancing of incommensurables, a battle between competing values, and a court should therefore accede to the legislature's solution. Of course, a court still must evaluate the degree of fit between the combined-purpose class and the class actually burdened. For example, if a traffic signal exemption for firemen is justified by their need for haste, then others with such a need—ambulances, or private motorists in a medical emergency—might assert that they are similarly situated with respect to the tradeoff between traffic safety and emergency needs. But the question whether the tradeoff itself is unwise or unfair (*i.e.*, whether the need for safety outweighs the need for any exemptions) is for the legislature.

The third category in which classification by compound mischiefs is permissible consists of those instances in which one can infer, from the combination of mischiefs, a single need or harm properly served by the law. Of course, this combined purpose must be legitimate and the classification must satisfy the minimum rationality constraint. In this third category, the intersection of M* and M** defines a class that is different in kind from both M* and M**. For

307. Although it is less common for the pursuit of one goal inevitably to further another goal than for it to inhibit the second, some examples might be noted. If both a chronic food shortage (M*) and serious unemployment (M**) exist, the state might decide to offer public subsidies to agricultural workers. Neither goal is being pursued fully, since non-agricultural employment still exists and other means of aiding agriculture (besides a worker subsidy) are available. But one justification for this choice of multiple goals is that remedying unemployment here inevitably helps remedy the food shortage.

By contrast, a prohibition of jaywalking by illegitimates exemplifies independent, rather than interdependent, "harms." (Again, these goals are at best mutually supportive, not mutually inhibiting.) Discouraging illegitimacy has nothing to do with discouraging jaywalking.

308. *See* Ely, *supra* note 26, at 1240 n.110.

example, an affirmative action program giving blacks preference in admission to medical school might be characterized as serving the twin goals of compensating for past discrimination against blacks generally and making health care more accessible to blacks. If these two goals are interdependent, then the state is justified in combining them (assuming that they are *legitimate* goals); it is no objection that the law benefits no black victims of discrimination other than would-be doctors, or that it benefits no white medical students who would treat blacks.[309] But an even more straightforward justification than "interdependence" is the argument that encouraging blacks to become doctors is itself a legitimate goal. To succeed, this argument must show that blacks have a sufficiently unique social experience and that doctors' social roles sufficiently differ from those of lawyers, businessmen, and others, that the goal differs in kind from the separate goals.

This form of argument *can* be abused— *e.g.*, one could argue that the classification in *Levy* is justified by the state's special desire to discourage illegitimate children from filing wrongful death claims. But this potential danger is not unique to classifications involving multiple goals and can be avoided simply by enforcing the requirement that the goal be legitimate.

Finally, classifying by compound mischiefs may be rational when it is part of a general policy of attacking one of the mischiefs in many or most of its manifestations. Thus, as suggested earlier,[310] the classification in *Levy* might have been rational were preventing illegitimates from bringing wrongful death suits just one part of a general scheme to deny illegitimates the rights generally accorded legitimate children. Similarly, another possible justification for using the tax laws against criminals is that there appears to be a general policy of burdening criminals in any way possible. The more general the policy, the less the "misfit" between felons who fill out tax forms and felons who do not, or (in the case of a law excluding felons from certain occupations) between felons who would have pursued a now-prohibited line of work and those who still may pursue their chosen occupation.[311] Of course, under this fourth justification for pursuing multiple goals, the "general policy" in question

309. These two objections may be relevant to whether the law would satisfy equal protection. Although they are irrelevant to the propriety of pursuing the two goals together, these objections are quite relevant to the ultimate question whether T and M are sufficiently close fitting, for the latter inquiry will also consider the availability of less restrictive alternatives.

310. *See* text after note 298 *supra*.

311. It may be objected that a preference for a "more general" policy simply resurrects the impermissible "part of the harm" approach, *see* text at notes 59-65 *supra*. Strictly speaking, that objection is correct. However, this Note defines "general policy" to mean an attempt to get at a harm by whatever means are available. On this view, a "general policy" approach satisfies the minimum rationality constraint (particularly its cost element). That is, there may be no feasible way to address the "harm" of nonfiling felons who have illegal income, and it is

must still be legitimate. (It is unlikely that the purpose of either of the above hypotheticals is legitimate.)

To conclude, classifications according to multiple legislative goals can ordinarily be justified under one of at least four arguments: that possession of one harm implies a higher degree of another; that pursuit of one goal inevitably aids or inhibits pursuit of a second; that a single legitimate purpose can be inferred from the several purposes; or that the classification represents part of a general policy of attacking one of the harms. Especially when the personal interest is less than fundamental and the benefitted or burdened class is not a suspect one, courts should construe these available justifications liberally. But a bare compound mischief rationale is unsatisfactory; it is the classic argument that proves too much. It would justify conferring any benefit on any class on the ground that to do so would encourage persons to join the class or would reward its members. Similarly, imposing any burden would discourage persons from joining or remaining in the class or would penalize its members.[312] Such reasoning would defeat the purpose of the minimum rationality constraint.

A final cautionary note: satisfying one of the four justifications for pursuing multiple goals does not by itself satisfy the equal protection clause. This section has been exclusively concerned with the circumstances in which a state may burden those posing two harms without also burdening those posing only one of those harms. If the state *may* burden this "compound mischief" class exclusively, then that class should be designated as "M" and treated like any other single-purpose class—that is, it must satisfy the minimum rationality constraint, and it must fit class T adequately. The present section, in other words, simply defines the permissible contours of M in the context of multiple legislative purposes.

V. Exceptions to Misfit Analysis

Misfit analysis cannot be applied to those legislative goals that inevitably describe the purpose class M only ambiguously.[313] Such

thus reasonable to pursue the filers. By contrast, there are many feasible (and more efficient) alternative means of subsidizing musicians besides tax deductions for purchases of yachts.

312. *See* Examining Bd. of Engineers v. Flores de Otero, 426 U.S. 572 (1976), in which the Court invalidated a Puerto Rican statute making American citizenship a prerequisite for a civil engineer's license. The Court rejected the justification that the law would raise the prevailing standard of living, reasoning, "To uphold the statute on the basis of broad economic justification of this kind would permit any State to bar the employment of aliens in any lawful occupations." 426 U.S. at 605-06.

313. If a goal is "legitimate," then it is *unnecessary*, rather than impossible, to analyze that goal's misfit with respect to a more general goal. Thus, although they are also "exceptions" to

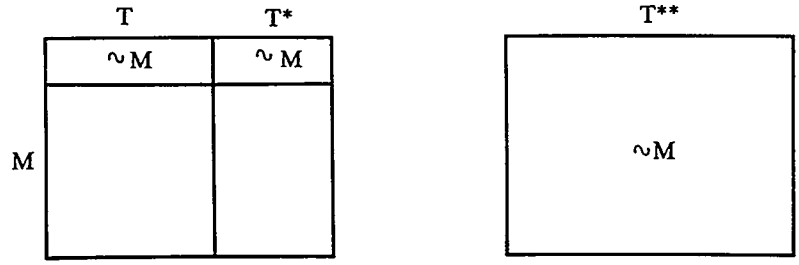
goals are arbitrary in that different legislatures might reach different yet equally supportable conclusions as to which classification *best* served a given objective. Of course, even nonarbitrary goals may be equally well served by a number of different classifications that create the same amount of misfit but impose the statutory burden or confer the statutory benefit on different persons. But with respect to arbitrary goals it is impossible to evaluate the legislature's claim that the means chosen furthers the goal as well as or better than any other means. For example, any number of different zoning plans "enhance the scenic beauty of a community" to significant, though incommensurable, degrees.

Such inevitably arbitrary classifications raise a number of problems. What kinds of classifications *are* "inevitably arbitrary"? Can the amount of "misfit" created by such classifications be sensibly measured? Do any considerations prescribe *which* of several equally arbitrary alternatives the state may choose?

A taxonomy of "inevitably arbitrary" equal protection goals divides them into (1) inherently imprecise goals, *e.g.*, matters of taste, and (2) "neutral" goals, which are served as well by classifying one group as another, *e.g.*, draft lotteries or budget reductions.[314] By fleshing out each of these categories, we may determine the appro-

misfit analysis, legitimate goals are exceptions of a very different kind from the goals discussed in this Part.

Misfit analysis is unnecessary for legitimate goals because, by definition, such goals limit the size of M. If a goal (say, encouraging farming) is legitimate, then no plaintiff *outside* M can make a cognizable equal protection claim (an oilman cannot complain that the state is not encouraging oil production). Suppose, for example, there were a subsidy to small farms (T). Large farm members (T*) could complain, if we assume that encouraging small farms is not a legitimate goal, while encouraging all farms is. But oilmen (T**) who argued that the "real" purpose was to aid the economy could not complain, for "encouraging farming" is legitimate. Thus any misfit must occur with respect to classes T and T*. The purpose class, M, simply does not extend to T**:



(The diagram realistically assumes that the law is overinclusive, *i.e.*, that some farmers who receive the subsidy do not need encouragement.)

314. While these two categories are fairly comprehensive, they are by no means mutually exclusive.

priate equal protection standard for each.[315] Such a determination is important, for if some constraints are not placed upon these categories of goals, they would undermine equal protection review.[316]

---

315. The analysis developed in the text departs significantly from Professor Ely's discussion of nonrational goals, *supra* note 26, at 1228-49. Ely explored this problem in the context of a general analysis of legislative motivation in constitutional law. Essentially he argues that a court should not look at motivation when the "disadvantageous distinction model" applies, *i.e.,* when the simple fact that the law creates a distinction is enough to trigger the government's burden of justification. *Id.* at 1228-30. But where that model is inapplicable—in situations requiring discretionary, random, or partly random choice (in other words, where the choice is nonrational)—proof of unconstitutional motivation would be required. Examples of random choice situations are the selection of jurors or draftees and the setting of voting district boundary lines. Discretionary choice situations are of two types:

> The first is the situation where the Court is prepared to credit as acceptable, *along with other relatively precise goals*, one goal—such as the promotion of "good taste"—whose relation to various choices cannot be evaluated by a calculus of "rationality" and "irrationality." The second is the situation where the Court is unprepared to restrict an area's class of acceptable goals to anything more precise than the promotion of the general welfare.

Id. at 1237 (emphasis original). Aesthetic distinctions in a dress code are an example of the first situation. Decisions as to which industry to favor under the tax code, and how to vary punishment among crimes, exemplify goals that promote the "general welfare."

The analysis in the main text clearly draws from Ely's categories, but it alters them in crucial respects. Ely's "random" choices will be discussed in this Note's category two, his "discretionary" "nonrational goal" will be discussed in category one, and his "discretionary" "general welfare" choices have been discussed in the section on legitimate goals, Part III.A.1 *supra*. More important than differences in categorization is the fact that Ely's only constraint on nonrational choices is the prohibition against unconstitutional motives. This Note's analysis, by contrast, attempts to indicate additional constraints appropriate to each category of nonrational goals.

For further criticism of Ely's approach, see Brest, *supra* note 26, at 135-46.

316. Another possible category of nonrational goals consists of those goals purporting to benefit or burden anyone a given process designates as deserving. We determine whether an individual has been "properly" treated not by reference to his actual characteristics, but only by reference to the fairness of the process leading to his selection. For example, a discretionary decision to prosecute is fair unless the method of selection is "deliberately based upon an unjustifiable standard such as race, religion, or other arbitrary classification." Oyler v. Boles, 368 U.S. 448, 456 (1962). The defendant's characteristics are irrelevant, except for his membership in the large class of bona fide criminal suspects. Other examples of discretionary process goals might be the judge's choice of a criminal sentence within the statutory range, a municipality's decision as to the location of a new public facility, and a school board's choice of curriculum. *See* Ely, *supra* note 26, at 1245. Of course, such choices also may involve "aesthetic" goals, thus complicating judicial review.

Since discretionary-process goals cannot be evaluated in terms of the class which the process ultimately burdens, a strict misfit analysis is impossible. On the other hand, principles of administrative law do apply, for discretion may only be exercised in pursuit of goals which further the general purposes for which discretion was delegated. For instance, prosecutorial discretion exercised against farmers, or mothers, or tennis players would presumably be subject to review; we may assume that the legislature intended that the prosecutor have the flexibility of selective enforcement based on such factors as the seriousness of the offense, public concern over a particular type of crime, and the like, but that strengthening the industrial sector (at the expense of agriculture) and discouraging births were not understood to be relevant factors. A court's judgment will necessarily be impressionistic and deferential, but at least something like a weak "misfit" approach is appropriate. *See, e.g.*, Hampton v. Mow Sun Wong, 426 U.S. 88 (1976); discussion of it in L. TRIBE, *supra* note 100, at 286. In short, although a discretionary choice can be justified by any of a broad range of goals, it remains subject to judicial review when the goals are apparently irrelevant to the general purposes of the scheme or when the choice is relevant to none of the goals.

A curious objection may be raised to this analysis. Since the legislature itself exercises the

A. *Inherently Imprecise Goals*

The objectives of some classifications are inherently imprecise. If retribution is a proper aim of punishment, then the state may "rationally" prescribe longer sentences for rapists than armed robbers, or vice versa.[317] This is not because we accept such classifications as legitimate or because we know what "immorality" is but cannot measure it.[318] Rather, our concept of "immorality" is so imprecise that we do not even know whether a legislature's classification is "correct."

Because of this imprecision, a legislature may be able to define social mischiefs as to which almost any specific regulation would be "rational." For example, in *McGowan v. Maryland*,[319] the Supreme Court upheld Sunday closing laws against claims that their numerous exceptions violated equal protection.[320] Among the items which

---

discretion to choose between acceptable goals, *i.e.*, goals furthering the general welfare, is not any resulting piece of legislation beyond judicial review? Is not the only question whether the process was fair? If the law serves a goal which furthers the general welfare and was not prompted by an illicit motive, is it not valid?

This objection is not frivolous, but it can be answered. First, as a general matter existing political institutions are not examples of "pure procedural justice" in which an outcome is automatically accepted as fair simply because it is the result of a properly followed procedure. *See* J. RAWLS, A THEORY OF JUSTICE 86-89, 198 (1971). The equal protection clause would be largely gratuitous if we entertained such a conception of our political processes. Second, the rationale for accepting limited discretion does not extend to all situations. We allow a prosecutor to choose freely among several goals because we think flexibility is a value in itself and because his freedom is constrained by the terms and purposes of the statutes. By contrast, the constraint of "pursuing the general welfare" is hardly a constraint at all. And the desirability of allowing a rule to be flexibly enforced in accordance with one of its purposes has no counterpart in the general legislative process.

317. *See* Ely, *supra* note 26, at 1243-44; Williams v. Oklahoma, 358 U.S. 576, 586 (1958) (nothing in the Constitution "require[s] a State to fix or impose any particular penalty for any crime it may define or to impose the same or 'proportionate' sentences for separate and independent crimes").

318. This problem of uncertainty born of ignorance could be discussed as an independent exception to misfit analysis. Sometimes, in other words, misfit analysis is impossible not because the subject matter lacks precise standards, but because we are ignorant of the dimensions of the class actually picked out by those standards (even though those standards are theoretically clear). For example, the government may surely ban cancer-causing drugs, but in any given case it may be unclear whether, or to what degree, a drug is in fact carcinogenic. In a word, the problem to be remedied is of uncertain scope. (For a discussion of the distinct question whether the remedy chosen will actually be effective in alleviating the evil, see text at notes 256-61 *supra*.)

Of course, many classifications are in some sense grounded upon ignorance; if a legislature knew precisely how to pick out M, it would not create the presumptive classification T. Since M will generally have a rough and ambiguous boundary, courts should be tolerant. The legislature can usually be trusted not to shoot in the dark against a culprit whom no one believes is there. The problem of uncertainty born of ignorance is inherent in legislation and seems to be of controllable dimensions.

319. 366 U.S. 420 (1961).

320. The equal protection claim was unanimously and summarily rejected by the Court, 366 U.S. at 425-28. The more difficult issue was whether Sunday closing laws violated the establishment clause of the first amendment, 366 U.S. at 429-53.

could be sold despite the general prohibition were tobacco, confectionaries, milk, bread, fruits, gasoline, and newspapers and periodicals. In one county, beaches, bathhouses, dancing saloons, and amusement parks could remain open and sell their usual merchandise. In the Court's words,

> a legislature could reasonably find that the Sunday sale of the exempted commodities was necessary . . . for the enhancement of the recreational atmosphere of the day—that a family which takes a Sunday ride into the country will need gasoline for the automobile and may find pleasant a soft drink or fresh fruit; that those who go to the beach may wish ice cream or some other item normally sold there; that some people will prefer alcoholic beverages or games of chance to add to their relaxation. . . .[321]

If "enhancing the recreational atmosphere of the day" is the goal, a judgment as to rationality is no judgment at all.[322] Prohibiting rather than allowing the sale of gasoline would also serve that goal by allowing gas station attendants a day off and encouraging families to stay home and engage in more private forms of recreation. Similarly, many of the items plaintiffs were actually convicted for selling—a loose-leaf binder, a stapler, and a toy submarine—could rationally have been exempted.

The inherently imprecise category of arbitrary goals is difficult to constrain. Courts might refuse to accept such goals as justifications for statutory classifications. But surely a court cannot simply dismiss all goals involving judgments of aesthetics,[323] morality,[324] "good taste,"[325] and so forth, because it cannot determine whether they have been rationally served. A later section[326] will discuss one plausible constraint, namely, that infringement of significant personal interests must be justified by correspondingly significant state interests. Thus, a zoning law seriously restricting the associational rights of college students[327] might demand a more persuasive justification than that aggregations of young people are offensive, even

---

321. 366 U.S. at 426.

322. The Court does purport to use the rationality standard in stating, after the quoted paragraph, that the "record is barren of any indication that this apparently reasonable basis does not exist . . . ." 366 U.S. at 426. But one cannot take this argument seriously; the record is "barren" because recreational preferences are not subject to proof. How could plaintiffs demonstrate that families taking a Sunday drive do *not* find soft drinks and fresh fruit "pleasant"?

323. *See, e.g.*, New Orleans v. Dukes, 427 U.S. 297 (1976).

324. *See, e.g.*, Labine v. Vincent, 401 U.S. 532 (1971). As the dissent notes, the majority opinion barely discusses the "rationality" of the classification in promoting family life.

325. *See* Ely, *supra* note 26, at 1239-40, 1242-45; *cf.* Kelley v. Johnson, 425 U.S. 238 (1976) (upholding regulation of policeman's hair length and style, since similarity in appearance of police officers is desirable to facilitate public recognition or to inculcate esprit de corps within the police force).

326. Part VI *infra*.

327. *But see* Village of Belle Terre v. Boraas, 416 U.S. 1 (1974).

though the justification cannot by its own terms be refuted. Students could not prove themselves less "offensive" than groups which *are* allowed to live together.

A second method of constraining this category of goals is to find the goal impermissible. While this method will usually be unavailable, in a significant number of cases the goal might be impermissible in the sense that it must be rejected by the courts as a justification. Inculcating anti-homosexuality or enforcing a grooming standard for employees[328] are examples of arbitrary goals within the inherently imprecise category that might be illegitimate.

These two constraints apply to all goals, of course, but they apply with special force to inherently imprecise ones. Because this category of goals is difficult to evaluate under a rationality standard, the significance and permissibility of the goal should be meticulously scrutinized. Impermissible purposes are too easily redefined in more appealing, but imprecise, terms. Enforcing uniformity in hair length becomes encouraging esprit de corps.[329] Thwarting young people's privacy right of access to contraceptives becomes expressing state disapproval of sexual promiscuity by the young.[330] To be sure, the state cannot be forbidden from pursuing certain goals simply because they are imprecise; such reasoning might necessitate the invalidation of most of our criminal laws and many of our zoning ordinances. Nevertheless, courts should be sensitive in reviewing these goals to ensure that their imprecision does not conceal their unfairness.

B. *Neutral Goals*

The second category of arbitrary goals consists of those goals whose purpose is achieved by burdening or benefitting a limited class of persons indistinguishable from a larger[331] class. Burdening one person within the broad class serves the goal as well as burdening another. Because it is a matter of indifference which individuals are burdened, goals in this category are "neutral." Draft lotteries and jury selection procedures fall within this category, since the implicit purpose is to draft a certain number of able-bodied men or to impanel a certain number of adult, unbiased jurors. Bona fide social experiments are also within this category, for their purpose is to

328. *But see* Kelley v. Johnson, 425 U.S. 238 (1976).

329. Kelley v. Johnson, 425 U.S. 238, 248 (1976).

330. *See* Carey v. Population Servs. Intl., 431 U.S. 678 (1977), where the Court refused to uphold such an interest.

331. A goal is within this category even if the class ultimately burdened is drawn from a subclass of humanity, so long as it is a matter of indifference *which* members of the subclass are chosen. A draft lottery is still random even though women, children, and the aged are excluded from the "venire."

burden or benefit an experimental group identical in all relevant[332] respects to a control group. Less obviously within this category are purposes such as reducing total administrative or program costs; excluding one person from the program presumably reduces its cost as much as excluding another.[333]

When the state invokes a neutral goal, the demand for rationality is often inapposite. But the opinion in *Reed v. Reed*[334] suggests that the Supreme Court has not always recognized this. The statute in *Reed* established a preference for men rather than women in administering an estate when competing claimants were otherwise equally qualified. Although the Court conceded that "the objective of reducing the workload on probate courts by eliminating one class of contestants is not without legitimacy," it reasoned that "[t]o give a mandatory preference to members of either sex over members of the other . . . is to make the very kind of arbitrary legislative choice forbidden by the Equal Protection Clause."[335] The Court failed to realize that if this purpose is legitimate, it *demands* an arbitrary choice; the point of the scheme is to avoid the expense and time of making rational distinctions between claimants.[336] The result in *Reed* is surely correct but requires a more sophisticated analysis than the argument that neutral goals are invalid because they are arbitrary.

Neutral goals present two important questions. First, when may the state pursue neutral goals? That is, when may it simply exclude some admittedly similarly situated persons from a benefit, rather than either proportionally diminish the benefit for all members of the class or seek rational distinctions among class members? Second, assuming the state may pursue a neutral goal, what criterion may be used to limit the class of persons which the law will burden

---

332. That is, the groups are identical with respect to those factors (other than the experimental variable) which might affect the experimental result. I am assuming that the experiment itself artificially imposes a variable condition. If it merely examines the behavior of two groups, only one of which "naturally" (*i.e.*, without government intervention) possesses this condition, then the "experiment" is merely a social study and creates no special equal protection problems.

333. *See* Shapiro v. Thompson, 394 U.S. 618, 633 (1969) (dictum) (a state cannot reduce expenditures for education by banning indigents from attending school). In some circumstances, however, it would be permissible to exclude only the more expensive recipients. *See* text at notes 223-24 *supra*, discussing application cost.

334. 404 U.S. 71 (1971).

335. 404 U.S. at 76.

336. The Court also discussed as a purpose the goal of avoiding intrafamily controversy, but it denied that the choice could "lawfully be mandated solely on the basis of sex." 404 U.S. at 77. But again, since that value is served only by avoiding a factual examination of the competence of spouses competing to be administrators, only an arbitrary choice furthers the goal. Thus, the mere fact that the classification was arbitrary cannot invalidate it; either the goal is never permissible, or there is something obnoxious about *this* kind of arbitrary choice (*i.e.*, one based upon gender). The latter rationale is probably the more accurate, since the scheme would not be unconstitutional if the final candidates simply flipped a coin.

or benefit? More concretely, the first question asks when, for example, government benefits may be eliminated for half the recipients instead of either being diminished by half for all the recipients[337] or "rationally" allocated to the more deserving half of recipients. The second issue is whether a random method must be used to reduce the class. These two problems will be discussed in turn.

The question whether the state can pursue a neutral goal can be answered only after considering the applicability of the minimum rationality constraint to this class of goals. Classifications furthering neutral goals neither satisfy nor fail to satisfy the constraint,[338] for they extend a benefit to persons who deserve it no more and no less than persons not receiving it (since the two classes are, by hypothesis, similarly situated). This characteristic of neutral goals might seem to suggest that they should be evaluated like any classification that barely satisfies the minimum rationality constraint— *i.e.*, as the following section of this Note outlines, a court would evaluate the fairness of the law in terms of the closeness of fit and the comparative significance of state and personal interests, and it would investigate reasonable alternatives. But if this suggestion were correct, the state could seldom justify a neutral goal. Since the "closeness of fit" for neutral goals is always minimal, the state would suffer a built-in disadvantage in any equal protection balance and would always lose whenever the personal interest was significant or alternatives were available.

The better approach is to recognize that neutral goals often present a new kind of state interest, a special interest that can *only* be achieved by pursuing a goal "neutrally" rather than "rationally". This interest should be weighed along with the state's traditional, underlying interest in alleviating the harm. The extra weight may compensate for the loss of closeness of fit on the state's side of the equal protection balance.

Bona fide social experiments[339] are the best example of a special

337. Of course, the example is oversimplified. Because of administrative overhead, neither excluding half the recipients nor reducing all benefit levels by one half would reduce costs by exactly half.

338. In terms of the model presented in text preceding note 189 *supra*, the dividing line between M and ∿M has a slope of zero, and the law is neither "rational" nor "irrational".

339. This is not to say that anything a legislature denominates a social experiment is one, nor that bona fide experiments are entirely immune from equal protection review. For reasons already noted, the broad view of experiments—that the state may always attack "part of a harm"—is unacceptable. However, bona fide social experiments can be appropriately defined without extending the "experimentation" rationale to every case of underinclusion. If an experiment is bona fide, then the state's independent interest in gaining knowledge can properly be weighed in the equal protection balance.

Even minimal requirements to ensure the legitimacy of social experiments should keep most underinclusive classifications exposed to judicial review. Such requirements might include a reasonable time limit; clearly designated control and experimental groups; a significant potential not simply of deriving knowledge for its own sake, but also of applying the experi-

state interest in pursuing a goal neutrally. Although one purpose of

---

mental treatment to all within the relevant class; and, perhaps most important, a limitation on the differences in treatment between the control and experimental groups. In some areas, especially criminal justice, we might even require that the experimental treatment not, in the broadest sense, be "worse" than the treatment of the control group. *See* Morris, *Impediments to Penal Reform*, 33 U. CHI. L. REV. 627, 647-48 (1966) (suggesting a "less severity" principle as an ethical safeguard on experiments in the criminal justice field). The courts may justifiably acknowledge that the burden of a harsher experimental treatment is usually a greater personal deprivation than the burden of *excluding* many persons from an experimental *benefit*. *See* note 288 *supra*.

Aguayo v. Richardson, 473 F.2d 1090 (2d Cir. 1973), *cert. denied*, 414 U.S. 1146 (1974), appears to be the only case examining whether a social experiment violates the equal protection clause. The case involved a New York work-relief plan which imposed on welfare recipients in selected counties certain conditions inconsistent with applicable federal standards. (The Department of Health, Education, and Welfare approved the plan by waiving normal standards.) The court found that the differential treatment did not violate equal protection.

Whatever the propriety of the result, the court's reasoning is disappointing. Chief Judge Friendly stated that determining how to impose the welfare system is as "legitimate" as any purpose can be and that social and economic experimentation is an acceptable means to that end. Citing a Brandeis dissent to the effect that due process should not prevent state experimentation, he concluded that, by analogy, equal protection should not prevent experimentation by a state on less than a statewide basis. 473 F.2d at 1109. This analogy, of course, is imperfect, for to say that a policy can be imposed across a state is not to say that partial imposition comports with equal protection. Capron, *Social Experimentation and the Law*, in ETHICAL AND LEGAL ISSUES OF SOCIAL EXPERIMENTATION 127, 156-57 (A. Rivlin & P. Timpane eds. 1975).

The court continued by rejecting Justice Jackson's argument against underinclusion (*see* note 276 *supra*) as

> inapposite to the selection, on a random but rational basis, of certain areas of the state to try out a program for the very purpose of determining whether it, or some variation of it should be made applicable to all. The Equal Protection clause does not place a state in a vise where its only choices in dealing with the problems of welfare are to do nothing or plunge into statewide action.

473 F.2d at 1109-10. What is curious about the court's reasoning is that it sounds as if it were discussing and legitimizing *any* case of underinclusion. The very special nature of the instant government action—to wit, a carefully controlled bona fide social experiment—is only hinted at ("random but rational") by the court. Thus, the opinion offers no guidelines as to the permissible limits of social experiments.

Social experiments should be subject to firm constraints. But forbidding less constrained forms of experimentation will not put the legislature to the uncomfortable "all-or-nothing" choice between an imperfect present and an uncertain future, *see Developments, supra* note 1, at 1084-86. Governments can and do draw upon a vital source of "experimental" data other than their own projects—namely, the experience of other governments: "It is one of the happy incidents of the federal system that a single courageous State may, if its citizens choose, serve as a laboratory; and try novel social and economic experiments without risk to the rest of the country." New State Ice Co. v. Liebmann, 285 U.S. 262, 311 (1932) (Brandeis, J., dissenting).

Of course, as Justice Brandeis' opinion assumes, interstate differences cannot be challenged under the equal protection clause (unless those differences are dictated by a federal program). Similarly, local governments (to whom the state has delegated legislative powers) may differ in their treatment of their own citizens; it is only intrastate variations imposed by a single sovereign, the state, that must be justified. The Supreme Court, however, seems not to have clearly understood the proposition that one sovereignty may treat its citizens differently from another sovereignty but must justify differential treatment of its own citizens. *See* Salsburg v. Maryland, 346 U.S. 545 (1954), and Missouri v. Lewis, 101 U.S. 22 (1879), each of which upholds without serious examination intrastate variations imposed not by local sovereignties, but by the state.

The Court in *Salsburg* argued that equal protection demands equality between persons, not places, 346 U.S. at 551—a claim which fails to address the question whether a given discrimination between persons in different places is justifiable. More sophisticated is the argument that since the state could grant each county home rule under which local option might create

an experiment is the usual state interest in attacking a social harm, another substantive purpose is to study the extent of social problems and the efficacy of various solutions. That study is impossible if all who pose the harm, or those who pose it most acutely, must be similarly burdened, for most experiments need a similarly situated control group. Of course, not every bona fide social experiment satisfies the equal protection guarantee. For example, a court might not allow an experiment which forbad a randomly chosen group to consume alcohol, even though a general prohibition is permissible; the experiment's possible informational benefits pale in comparison with the unfairness of imposing this significant burden on only a few drinkers. Insofar as the state is simply pursuing its traditional interest, *i.e.*, attempting to reduce the evils of drinking, it should do so either "rationally," by penalizing the most problematic drinkers (*e.g.*, drunk drivers), or "proportionally," by burdening all drinkers similarly.

There are examples other than experiments of an independent state interest in pursuing a goal neutrally. Consider a draft lottery or a jury selection procedure. Like social experiments, these classifications may evidence a substantive purpose other than numerically diminishing a larger class—namely, placing a cross section of the community in the jury box or on the battlefield. Although "rational" distinctions would be possible—the draft could be limited (more than it has been) to the physically fittest, and the jury to the brightest and most conscientious—such distinctions would less effectively promote citizenship than the neutral, cross-sectional scheme. Here, as in the case of experiments, the independent interest deserves special weight in the equal protection balance, since it is an interest which the alternative of rational classification will not further.[340]

---

variation, the state can impose variant conditions directly. 346 U.S. at 552. Even this argument is troublesome without some showing of rational differences between the counties, as would be the analogous argument for different federal regulations for different states.

340. Note, however, that the independent state interests in draft and jury classifications sometimes *can* be served by a proportional diminution of the benefit or burden. Although the "cross-sectional" interest is not furthered by rational classification, it *is* served by a proportional diminution; if more persons enter military service or jury duty for a shorter time, the inculcation of citizenship values might be even more effective. Social experiments, by contrast, demand not simply a nonrational classification, but also a nonproportional approach, since in either case there can be no control group.

The question naturally arises in this kind of case whether the state *must* burden proportionally when it *can* do so without sacrificing an independent state interest. There is no simple answer. Since proportional diminution is fairer than random or semi-random exclusion, it will often be a reasonable, less-restrictive alternative, especially since its use does not compromise the independent state interest. Of course, proportional diminution might entail considerably more administrative expense—compare the costs of training 100,000 soldiers for two years of duty with the costs of training 200,000 for one year. Whether proportional diminution is required will thus depend on the details of the equal protection balance in any given case.

A comparable problem arises with respect to neutral goals that do *not* serve an independ-

However, not all neutral goals serve special interests of this sort. Attempts to reduce a program's costs and decisions that only part of a harm is worth alleviating, for instance, both reflect neutral goals. If it costs about as much to benefit or burden one individual as another,[341] and if the program's budget must be cut by a certain percentage, then the state might simply exclude that percentage of the class from the program. But there is no *independent* state interest in excluding rather than rationally classifying here. Although it might well be cheaper to exclude beneficiaries randomly than to search for and apply a "rational" criterion (recall the procedure at issue in *Reed*), cost savings are not the kind of independent state interest previously described. Rather, they are already an element of the minimum rationality constraint and equal protection balance and will thus be considered by a court in any event.

Classifications with neutral goals which do not further independent substantive state interests are especially vulnerable to equal protection criticism. If the classification imposes a *discrete* burden that cannot be proportionally expanded or diminished, the state will nonetheless have to show that a rational distribution of the burden or benefit would be difficult or costly, especially if the personal interest is of any significance, or if the classifying trait is stigmatizing. In *Reed*, for example, the benefit of being appointed an administrator is "discrete"—only one person can effectively hold the position—yet the classification is invalid because the possible cost saved by automatically preferring men does not justify stigmatizing women and imposing a financial burden on the losing claimant. By contrast, a court would rarely overturn a city's decision as to where to begin installing a sewer line (a discrete benefit, since the city cannot begin everywhere), for the decision probably stigmatizes no one and burdens no significant personal interest.

---

ent state interest and whose burdens or benefits can be proportionally diminished. *See* discussion in text at notes 342-44 *infra*.

341. If individual *A* costs more to benefit than does individual *B*, but *A* also has higher needs, is it more "rational" to exclude *A* from the program or to exclude *B*? In theory, the answer would seem to depend on the cost-efficiency of benefitting each. Thus, if *A* has 5 "units" of need and costs $10 to benefit, while *B* has 3 "units" of need and costs $9 to benefit, the state should exclude *B*, despite his lower total cost (unless of course there is only $9 left to spend in the entire program). However, there are obvious practical and theoretical problems with this approach. A meaningful utilitarian calculus is often impossible because of the difficulties of measurement; a "unit" of need is only a social scientist's dream. Even the *concept* of need is disturbingly slippery. For example, if the state has $100 to spend, it might prefer benefitting 11 *B*s to benefitting 10 *A*s, even though fewer "units" of need were alleviated (only 33 units, rather than 50), on the ground that it is more important to benefit more *people* than to alleviate a greater "amount" of need.

If the benefits can be proportionally diminished, these complexities are multiplied, for one naturally asks, Proportional to what? To cost, or to need? If to need, how is need defined? For an interesting example of these problems, see Jefferson v. Hackney, 406 U.S. 535 (1972), discussed in text at notes 196-99 *supra*.

If the classification does not further an independent state interest, but its benefit is such that it *can* be proportionally expanded or diminished,[342] then equal protection demands that it be conferred upon all persons whom the legislature has determined deserve it. Arbitrary exclusion, even by lot, is impermissible.[343] By the state's own admission, these persons are similarly situated with respect to desert and must therefore be treated similarly.[344]

To summarize, some persons may be excluded from a burden or benefit even though those included are similarly situated, provided that the "neutral" classification furthers a substantive purpose (other than the mere numerical diminution of a larger class) which would *not* be served by rationally excluding some members of the class. If no such purpose exists, then the classification will be vulnerable to the criticism that a rational criterion of exclusion should have been

---

342. The proportional-diminution problem is more general than first appears, for it also arises whenever a program is instituted or expanded. That is, funds for programs are often separately authorized and appropriated, and thus can either be insufficient or more than originally planned, necessitating a proportional diminution *or* expansion. (The term "diminution" has been used for the sake of simplicity.) Moreover, conceptually identical "diminution" problems occur when, instead of entirely extinguishing the burden or benefit for part of a class (*i.e.*, "excluding" them), the state merely diminishes the burden or benefit for that part. *See* discussion of Jefferson v. Hackney in text at notes 196-99 *supra*.

343. *See* Townsend v. Swank, 404 U.S. 282, 291 (1971) (dictum) ("a State's interest in preserving the fiscal integrity of its welfare system by economically allocating limited AFDC resources may not be protected by the device of adopting eligibility requirements restricting the class of children made eligible by federal standards. That interest may be protected by the State's 'undisputed power to set the level of benefits. . . .' ") (citation omitted).

See also dictum in Ohio Bureau of Employment Servs. v. Hodory, 431 U.S. 471, 493 (1977), implying that if "rational" exclusion were not possible, random exclusion might raise an equal protection concern. The issue was whether workers whose unemployment was caused by labor disputes other than lockouts could be denied unemployment compensation benefits.

> It is clear that protection of the fiscal integrity of the fund is a legitimate concern of the State. We need not consider whether it would be "rational" for the State to protect the fund through a random means, such as elimination from coverage of all persons with an odd number of letters in their surnames. Here, the limitation of liability tracks the reasons found rational above [namely, increasing employer contributions only when *he* locks employees out, and allowing pressure on the employer to settle only in that situation], and the need for such limitations unquestionably provides the legitimate state interest required by the equal protection equation.

344. If the "random exclusion" approach were fully accepted, it would simply resurrect the "part of a harm" rationale for misfit criticized earlier and would violate the minimum rationality constraint. True, the problem as now stated is slightly different, since here a budgetary constraint is explicit, and the exclusion method is random. But the situation is not significantly different from that in which a legislature simply does not "choose" to attack all of a harm and uses a fairly neutral trait to limit the burdened class.

In a sense, of course, limiting the class rather than diminishing the burden does comport with the principle of equal treatment for the similarly situated, for when the class is limited by lot, all are "treated equally" with respect to their chance to win the lottery. *See* Breit & Elzinga, *Antitrust Penalties and Attitudes Toward Risk: An Economic Analysis*, 86 HARV. L. REV. 693, 707-08 (1973). But by similar reasoning, any law might be upheld as fair on the ground that all groups have a chance to win in the legislative halls. The proper approach is to define the fairness of treatment "rationally," that is, in terms of *desert* of the benefit or burden.

used. In addition, if the burden or benefit is not discrete, it must be proportionally diminished.

Assuming that a neutral goal *can* be pursued and that the burden or benefit is discrete, we confront our second problem: what criterion may be used to limit the class of persons? Since the purpose in neutral goal cases is served by benefitting or burdening a limited class of persons indistinguishable from a larger class, a random selection process is the most acceptable method. The most straightforward random selection method would be a lottery that selected in a single drawing recipients of the particular burden or benefit.

A "natural" lottery might also be used which allocated burdens and benefits according to some adventitious characteristic randomly distributed across the entire population[345]—perhaps hair color or month of birth.[346] A natural lottery is troublesome, however, for the public might improperly associate any personal trait, even a randomly distributed one, with desert of the burden or benefit, and the "winners" of the lottery might be stigmatized because of their selection. Even if the trait did not promote a suspicion that the legislature intended to stigmatize an unrepresented minority, a natural lottery might have this undesirable effect,[347] especially were the trait objectively observable and socially significant.[348]

Although ideally a "pure" lottery should be imposed on all neutral legislative classifications involving discrete burdens or benefits, the cost would often be prohibitive. A rough balancing approach should be applied—departures from the ideal become more justifiable as the burden or benefit becomes less significant—but the courts should be alert to stigmatizing effects, particularly where the deprivations imposed are great. Thus, filling the draft rolls with redheads should not be allowed. On the other hand, a court would sensibly ignore a complaint by a person with a high ZIP code number that the Postal Service first sorts letters with low ZIP code numbers instead of sorting by lot.

---

345. Sandalow, *supra* note 7, at 669 n.49.

346. It is difficult to find a natural characteristic that is distributed across the population entirely randomly. Many natural traits, such as hair or eye color, height, and race, are influenced, if not determined, by characteristics of the biological parents. Others, such as left-handedness or date of birth, are controllable to some extent, at least in the long term, and could therefore only be used on a short-term basis.

347. The danger of legislative oppression of minorities is a significant reason for preferring a "non-natural" lottery, but stigma is an independent reason. Arguably an actual random selection of one of a large number of "natural" adventitious characteristics—"redheads," "women," "crooked noses," "ear-wigglers," and so forth—should still be considered suspect, especially when the burden or benefit to be conferred is significant. *See* text following note 348 *infra*.

348. The trait is likely to have social significance when it is employed in a number of different classifications, for it then poses the risk of imposing cumulative disabilities. *See* San Antonio Independent School Dist. v. Rodriguez, 411 U.S. 1, 28 (1973).

VI. BALANCING

A. *Introduction*

Although only a few justices describe their analysis in such terms,[349] the Supreme Court does "balance" in equal protection cases. For example, the Court balances at the margin where fundamental interests are only indirectly burdened,[350] in analyzing irrebuttable presumption doctrine cases, and when it selects one of the three different tests employed in different equal protection cases. However, this section neither demonstrates that the Court balances, nor justifies the basic propriety of doing so. Rather, it explores the relationship of misfit analysis to equal protection balancing and identifies some problems that a balancing test must address.

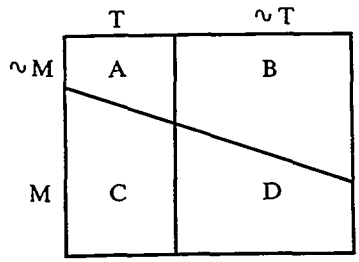
As described by the commentators, the structure of a balancing test is basically the same whether the balance involves equal protection, substantive due process, or the first amendment.[351] Traditionally, four elements are balanced, two on the individual's side and two on the government's. On the individual's side are the personal interest (I) that is burdened by the regulation and the extent to which that interest is burdened (B); on the state's side are the governmental interest served by the regulation (G) and the extent to which that interest is served, *i.e.*, the means-end "fit" (F).[352] Thus, a regulation fails the balancing test if the individual's interest, represented by I ×

349. *See* the dissenting opinions of Justice Marshall discussed in note 66 *supra*; the concurring opinion of Justice White in Vlandis v. Kline, 412 U.S. 441, 456 (1973); the concurrence of Justice Douglas to Justice Marshall's dissent in San Antonio Independent School Dist. v. Rodriguez, 411 U.S. 1, 70 (1973); and the concurrence of Justice Brennan to Justice Marshall's dissent in Dandridge v. Williams, 397 U.S. 471, 508 (1970). It is unclear from his mystical declaration that "[t]here is only one Equal Protection Clause" whether Justice Stevens subscribes to a flexible balancing approach. Craig v. Boren, 429 U.S. 190, 211 (1976). And see 429 U.S. at 210 n.1, where Justice Powell concedes that "[t]here are valid reasons for dissatisfaction with the 'two-tier' approach."

350. *See* text at notes 79-80 *supra*.

351. *See, e.g.*, P. BREST, *supra* note 116, at 988-90; Perry, *supra* note 68, at 388; Simson, *supra* note 5, at 678-81; Note, *supra* note 218, at 466-68.

352. In terms of the traditional model, "fit" is generally considered to be represented by either the ratio C/(C+D) or the ratio (B+C)/(A+B+C+D). Equal protection "fit," by contrast, is indicated by the ratio A/C: B/D.



B, "outweighs" the governmental, represented by G × F, that is, if I × B>G × F. Focussing on fit, F, we see that misfit is more tolerable (F can be less) the greater the governmental interest, G, and the lesser the individual interest, I, or the extent to which it is burdened, B.[353]

Despite the surface plausibility of this model, it obscures several important elements of equal protection balancing. These elements may be isolated by examining the three tests the Court currently employs. In reviewing a classification, the Court asks whether it is

(1) Rationally related to a legitimate state interest;

(2) Substantially related to the achievement of important governmental objectives; or

(3) Necessary to achieve a compelling state interest.

The third test applies to fundamental interests and suspect categories;[354] the second to gender-based classifications[355] (and perhaps to laws based on illegitimacy);[356] the first, purportedly, to all other classifications.[357]

These three tests diverge from the simplified balancing model in several ways. To be sure, the progression from the first test to the third parallels the model in requiring closer fit and a greater government interest as the personal interest becomes more fundamental. But the three levels of review differ in more than the *closeness* of the fit they demand. The first test asks for rationality, that is, at least a minimal relevant difference between those who are and are not burdened by the law. The second test asks for "substantiality," which presumably *includes* rationality. But what must be substantial—the differential in harm between those who are and are not burdened, or the "accuracy" of the classification in selecting *all* and only those who are harmful? If the former, then substantiality is an equal protection standard; if the latter, then substantiality is essentially a substantive due process standard. These standards are of course quite different. The state interest is not "substantially served" in the due

353. P. BREST, *supra* note 116, at 988.

354. *See Developments, supra* note 1, at 1087-104.

355. Craig v. Boren, 429 U.S. 190, 197 (1976).

356. The intermediate test has not been explicitly applied to classifications based on illegitimacy. However, the Supreme Court has stressed that such classifications, while not judged under the strict scrutiny standard, are not judged under a "toothless" standard, Mathews v. Lucas, 427 U.S. 495, 510 (1976). A classification based on illegitimacy is examined to see whether it broadly discriminates between legitimates and illegitimates or is instead carefully tuned to alternative considerations. 427 U.S. at 513. Where "[t]he reach of the statute extends well beyond the asserted purposes," the statute may be invalidated. Trimble v. Gordon, 430 U.S. 762, 772-73 (1977).

357. *See Developments, supra* note 1, at 1077-87.

process sense if a law allows girls but not boys to buy beer when very few boys are drunk drivers. On the other hand, the interest is not "substantially" served in the equal protection sense if, although quite a few boys are drunk drivers, a similar proportion of girls are also drunk drivers. In short, the balancing model (and the test applied in *Craig*) is ambiguous as to the *kind* of fit that must be "close."

A second problem with the model is suggested by the third standard of review (that where fundamental interests are involved, the means must be "necessary" to the end). "Necessity" denotes an absence of alternatives; it does *not* imply closeness of fit.[358] As we will see, there may be no alternatives to a loosely fitting law. Yet the "strict scrutiny" test simply assumes that "necessity" implies greater precision, and the traditional balancing model simply ignores the availability of alternatives.[359]

With these two[360] criticisms of the model in mind, we can more carefully analyze the relevance of "balancing" to the fairness of a law.[361] The next two sections discuss separately equal protection and substantive due process balancing. A separate discussion is crucial, for both the traditional balancing model (comparing I and B to G and F) and the tripartite equal protection standard applied by the courts are misleadingly phrased and conceal the distinct functions of equal protection and substantive due process analysis.

B. *Equal Protection Balancing*

The first requirement of equal protection is that the minimum rationality constraint be satisfied.[362] A law which satisfies this test should be subject to a second "balancing" test. This balancing process assesses the relative importance of the various elements of the

358. *See* P. BREST, *supra* note 116, at 990-93.

359. The distinction between closeness of fit and availability of alternatives is obscured in the formulation given by Simson, *supra* note 5, at 679-80, in which he characterizes the means-end fit as "necessary, significant, insignificant, or non-existent." Since a "necessary" relation may nevertheless create a considerable misfit, Simson's test might too easily validate overinclusive or underinclusive laws.

360. A third problem with the model is its potentially misleading treatment of the interplay between the factors on either side of the balance. The model is usually understood to mean that even a "perfectly rational" law may be invalid if the government interest is insubstantial in comparison with the significance of the personal interest (and the extent to which that interest is burdened). Analogously, a law may be valid even though the personal interest is "completely" burdened, that is, extinguished, if that interest is insubstantial in comparison with the government interest (and the extent to which it is achieved). *See* Part VI.B.3 *infra*.

361. The first amendment overbreadth test, interestingly enough, is quite similar to the equal protection and due process balancing tests. A law may be invalidated for first amendment overbreadth because the means-end "fit" is inadequate, because there are less restrictive alternatives despite close "fit," or because (whatever the alternative means) the state interest is insufficient to justify the restriction upon speech. Israel, *Elfbrandt v. Russell: The Demise of the Oath?*, 1966 SUP. CT. REV. 193, 217-18; *see* W. LOCKHART, Y. KAMISAR & J. CHOPER, CONSTITUTIONAL LAW 819-22 (4th ed. 1975).

362. *See* Part IV.A *supra*.

minimum rationality constraint. Three inquiries are relevant in this process. *How* different must the burdened and the unburdened be in the extent to which they pose the harm? When must the state seek less restrictive "reasonable" alternatives? When is perfect fit itself not enough? After a preliminary comment, these three questions will be examined in turn.

Critics of equal protection balancing might argue that the minimum rationality constraint, supplemented by substantive due process balancing, guarantees all the "fairness" the equal protection clause demands.[363] This argument has some force, for a stronger interpretation of substantive due process *does* mitigate the need for an equal protection scrutiny beyond "minimum rationality," and the Supreme Court's reluctance to invoke substantive due process after the *Lochner* debacle has certainly spurred the development of equal protection "strict scrutiny."[364]

This argument is ultimately unconvincing, however. A classification may satisfy both the minimum rationality constraint and substantive due process and yet on balance be unfair, as where a large difference in treatment turns upon a trivial difference in social harm. For example, had the statistics in *Craig* indicated that 1.9% of girls drove drunk as compared to 2.0% of boys, the Court could have invalidated the law despite its "rationality" and despite its satisfaction of substantive due process (the interest in purchasing beer is hardly fundamental).[365] What is fair on balance depends, as we will see, upon the extent of the differential in harm and the availability of alternatives in light of the competing state and personal interests.

1. *The Extent of the Differential in Harm*

A court's first task in equal protection balancing is to assess the *extent* of the differential in harm according to the significance of the personal and governmental interests, the severity of the burden, and the extent to which the government interest is served. The absolute amounts of overinclusion and underinclusion should not affect this balance (except insofar as they reflect the degree of differential harm). Of course, a case of perfect fit describes the greatest possible differential in harm, and in this sense equal protection balancing favors an absolute reduction in misfit. But the balance may validate a lesser differential in harm, and then it is unsatisfactory simply to

363. *See* Craig v. Boren, 429 U.S. 190, 226-28 (1976) (Rehnquist, J., dissenting). Although Justice Rehnquist does not explicitly endorse a substantive due process approach, he concedes that the irrebuttable presumption/due process question "call[s] for a balance of the State's interest against the harm resulting from any overinclusiveness or underinclusiveness." 429 U.S. at 226. Thus, Justice Rehnquist's approach is similar to the argument presented in the text.

364. *See Developments, supra* note 1, at 1131-32.

365. *See* Craig v. Boren, 429 U.S. 190, 227 (1976) (Rehnquist, J., dissenting).

point to the absolute amounts of misfit to demonstrate the "irrationality" (or insufficient rationality) of a law.

Turner v. Fouche[366] illustrates the distinction between the absolute amount of misfit and the differential in harm. The Supreme Court in *Turner* held that a statute limiting county school board membership to property owners failed even the weakest equal protection test.

> It cannot be seriously urged that a citizen in all other respects qualified to sit on a school board must also own real property if he is to participate responsibly in educational decisions, without regard to whether he is a parent with children in the local schools, a lessee who effectively pays the property taxes of his lessor as part of his rent, or a state and federal taxpayer contributing to the . . . annual school budget.[367]

This argument only points to the underinclusiveness of the classification. It does not demonstrate that there is *no* greater likelihood that property owners, as a class, will participate responsibly in educational decisions than nonproperty owners. In the portion of the opinion concerning the state's interest in ensuring that board members have an "attachment to the community and its educational values," the Court more usefully describes its reasoning:

> However reasonable the assumption that those who own realty do possess such an attachment, Georgia may not rationally presume that that quality is necessarily wanting in all citizens of the county whose estates are less than freehold. Whatever objectives Georgia seeks to obtain by its "freeholder" requirement must be secured, in this instance at least, by means more finely tailored to achieve the desired goal.[368]

Here the Court concedes a possible differential in "harm" and thus must be implicitly balancing what it perceives to be a small differential against the interest of nonfreeholders in school board membership.

Determining how great the differential must be between the burdened and the unburdened requires that the personal interest and the extent to which it is burdened be balanced against the government interest and the extent to which it is served. But this balance should occur only at the margin:[369] the question is not whether the personal interest at issue is in the abstract more important than the state interest, but whether a particular law's infringement of personal liberty outweighs its utility in serving its social end. Measuring the actual infringement is conceptually simple—a court merely asks how many persons have been burdened and what interests are to what

366. 396 U.S. 346 (1970).

367. 396 U.S. at 363-64.

368. 396 U.S. at 364 (footnotes omitted).

369. *See* Note, *supra* note 218, at 467-68.

extent impaired. Evaluating how well the classification serves the social end is more complicated. Of course, the social end is not to eradicate the social harm, since a law usually diminishes only part of a harm. In *Craig*, for example, the government's side of the balance cannot contain the undifferentiated interest in eliminating drunk driving by eighteen- to twenty-year-olds. The interest relates only to eighteen- to twenty-year-old *males* and reducing their drunk driving, since the regulation will not be fully effective. The Court should weigh only that portion of the harm which the law actually alleviates.[370] However wise it may be to presume that a law usually achieves what it purports to,[371] evidence that the burden is inefficacious should be weighed in the equal protection balance.[372] Thus, the Court in *Craig* was properly influenced by the fact that boys could easily obtain beer despite the regulation.

Unfortunately, scholarly discussion of the extent to which means serve ends has often missed the distinction between absolute amount of misfit and differential harm. Professor Gunther, for example, has spoken generally about the need to scrutinize the means-end fit in equal protection cases, and he seems to assume that simply adding the absolute amounts of overinclusion and underinclusion tells us whether the fit is justifiable.[373] But the issue should not be whether,

---

370. One benefit of this analysis is that it makes narrowly contrived purposes self-defeating. A narrow purpose reduces misfit only at the expense of narrowing the state interest in imposing the burden. Conversely, a more general purpose has a greater social value but results in a greater degree of misfit. A natural equilibrium is achieved in which the state is allowed significant leeway in defining purposes but is forced to accept the natural consequences of that definition.

For instance, if the purpose of the prohibition against carriers advertising the wares of others in Railway Express Agency v. New York, 336 U.S. 106 (1949), was to minimize traffic hazards, as the Court hypothesized, this significant public interest would be included in a calculus of misfit which would probably show that the classification was both grossly overinclusive and underinclusive. If, instead, the purpose was merely to protect existing carriers against an influx of carriers supported by large advertisers, a more precisely tailored classification would be accomplished at the expense of a lesser (or less obviously permissible) state interest. Of course, if economic discriminations are to be afforded minimal review, the classification might withstand analysis under either interpretation. The example nevertheless illustrates that this approach lessens the problem of the state reducing the effective degree of scrutiny by narrowly articulating its purpose.

371. *See* text accompanying notes 256-61 *supra*. Although the effectualness argument is discussed there in terms of due process, the discussion applies to equal protection as well.

372. If a burden is ineffective, both the state interest in imposing the burden and the personal interest in avoiding it are less than they would be if the burden were completely effective. Thus, the Court in *Craig* might have observed that the ease with which the prohibition could be circumvented made plaintiffs' complaint less appealing. Whether these effects cancel each other perfectly in an equal protection balance is an intriguing question. (Arguably equal protection complainants suffer some injury simply in being targeted as members of a class appropriately burdened, even if the arrows go astray. But a parallel argument can be offered for the state: many laws have a symbolic or "propaganda" value even if they are not enforced. *See* text at notes 149-50 *supra*.)

373. Gunther, *supra* note 2, at 47-48. In addition, see Simson, *supra* note 5, at 679-80, who apparently makes the same assumption.

in the abstract, some persons who are included should be left out or whether some persons who are left out should be included.[374] Rather, it is whether those who are included are "sufficiently" different from those who are left out, where sufficiency is judged according to the weight of the personal interests actually infringed and the weight of the social harm that the law actually alleviates.

2. *Prima Facie Fairness and Reasonable Alternatives*

This section asks when a court should concern itself with possible alternative means of accomplishing a law's purpose, and how the reasonableness of alternatives should be determined. In balancing with respect to differential harm, a court should make a prima facie determination of fairness before considering alternatives. The court may find that a certain differential is, "on balance," either fair or unfair regardless of the existence of reasonable alternatives. The driving test illustration employed in this Note is probably an example of a law that is fair irrespective of alternatives, if we can assume that the classification presents some differential in harm. Even if every other state in the nation had a less costly, more accurate test for screening dangerous drivers, no court would invalidate this state's test, given the relative unimportance of the individual interest in driving and the significance of the social interest in safe highways. In contrast, while denying food stamps to members of unrelated households might be the only feasible safeguard against households formed only to obtain them,[375] a court could well find the law "on balance" unfair.[376]

A court which concludes that a law's fairness can be judged without evaluating alternatives has in effect concluded that the law presents an extreme case under the balancing test. That is, it has found the law unmistakably fair or unfair. Greater precision as to the standards for "prima facie fairness" is probably unattainable, for

---

374. The Supreme Court often analyzes misfit in these terms and infers irrationality from the mere fact that the classification creates *some* overinclusion or underinclusion, without pausing to examine the *degree* of differential between T and ~T(equal protection misfit) or of due process overbreadth. *See, e.g.*, Jimenez v. Weinberger, 417 U.S. 628, 637 (1974), and Rinaldi v. Yeager, 384 U.S. 305, 310 (1966) (equal protection); Moore v. City of East Cleveland, 431 U.S. 494, 499-500 (1977) (plurality opinion) (due process).

375. *See* United States Dept. of Agriculture v. Moreno, 413 U.S. 528 (1973).

376. *See also* Craig v. Boren, 429 U.S. 190 (1976), in which the Court struck down a legislative attempt to reduce drunk driving by means of an alcohol purchase limitation—even though there may be "no apparent way to single out persons likely to drink and drive," 429 U.S. at 227 (Rehnquist, J., dissenting).

In one sense, of course, United States Dept. of Agriculture v. Moreno, 413 U.S. 528 (1973), and *Craig* are not laws which are unfair "regardless" of alternatives, but cases in which the Court will only accept the "alternative" of invalidation. But this is not a sensible characterization; the Court is holding that it is unsatisfied with the fairness of the law, not that it is satisfied with the reasonableness of the alternative.

those standards will reflect the premises with which a court approaches balancing.

A court which cannot make a prima facie judgment about a law's fairness must examine the reasonableness of its possible alternatives. The Supreme Court purports to look at alternatives only under the compelling state interest test,[377] but its analyses under the irrebuttable presumption doctrine and rational basis test reveal its interest in alternatives. The Court nominally considers only those alternatives that are at least *equally effective* in achieving the state's goal[378] (and also, of course, less restrictive). Nevertheless, under the compelling state interest test the Court does *not* in fact demand "equal effectiveness,"[379] and under the irrebuttable presumption doctrine it is openly undisturbed by the additional administrative costs of the alternative it imposes.[380] This divergence between theory and practice can be cured by formulating an explicit standard for evaluating alternatives under equal protection balancing. Simply stated, the reasonableness of an alternative should depend on the prima facie fairness of the law, on the social costs of the alternative, and on how much less restrictive the alternative is. This standard requires some elaboration.

The strict scrutiny "necessity" requirement has classically meant that there must be (a) no less restrictive means which (b) serves the government interest equally well.[381] The second element of this definition must be treated skeptically, for few alternative means serve an end *just* as well as the existing means. As Professor Brest has remarked,[382] the presence of such an alternative would demonstrate that the law was not Pareto optimal[383] and thus that its restrictiveness could be alleviated at *no* cost[384] even though the law was

377. *See Developments, supra* note 1, at 1102-03, 1122.

378. *See* Storer v. Brown, 415 U.S. 724, 736 (1974) ("the Constitution does not require the State to choose ineffectual means to achieve its aims"); American Party of Texas v. White, 415 U.S. 767, 781 (1974).

379. For example, in Carrington v. Rash, 380 U.S. 89 (1965), the Court stressed that more precise tests were available to determine the bona fide residency of servicemen and servicewomen who moved to the state, and it invalidated the broad denial of the right to vote to all such residents. It is obvious, however, that *some* additional administrative costs would be incurred by replacing the presumption with the more precise tests.

380. *See* cases cited in note 203 *supra*.

381. *See* Shapiro v. Thompson, 394 U.S. 618, 634-35, 637 (1969). The second element of the requirement, equal effectiveness, is often stated less explicitly than the first. *See* note 378 *supra*.

382. P. BREST, *supra* note 116, at 991-92.

383. The Pareto criterion states: "Any change which harms no one and which makes some people better off (in their own estimation) must be considered to be an improvement." W. BAUMOL, ECONOMIC THEORY AND OPERATIONS ANALYSIS 527 (4th ed. 1977). Here, if the alternative means is just as effective as the existing means and burdens personal interests less, then the existing means is not Pareto optimal, since a change in the alternative would not undermine the state interest (it would "harm no one") and would make some people better off.

384. *See also* Note, *supra* note 218, at 469.

"rational" in the sense of satisfying the minimum rationality constraint. Only a government with a bizarre or Spartan desire to minimize the happiness of its citizens, or, more plausibly, a government hostile to those citizens burdened by the law under review,[385] would fail to seek Pareto optimality. No court should give any weight to these government interests. Even if thoughtlessness or ignorance rather than malice motivated the irrationality, courts should require the "efficient" solution.

The Court applies, as it must, the "less restrictive means" test even where the state interest is less well served by the alternative means. Of course, the effectiveness of the available alternatives should be considered, but they need not be as effective as the challenged means. Before turning to the question of how "reasonable" the alternative means must be, we should briefly consider the other element of the classic test—the "less restrictive" requirement.

A "less restrictive" means is an alternative that either creates less misfit or imposes a lighter burden, or both. The "less misfit" interpretation is relatively uncomplicated.[386] It is the approach in irrebuttable presumption cases, in which the Court allows the legislature to confer the same benefits as the law under attack, but requires that they be conferred on a larger class. It is also the approach most often used in fundamental interest cases, in which the

---

385. *See id.* at 469 n.27; Nimmer, *The Meaning of Symbolic Speech Under the First Amendment*, 31 U.C.L.A. L. REV. 29, 41-42 (1973).

386. The kind of misfit which a less drastic alternative reduces is usually overinclusion (or, in the case of a benefit, underinclusion), for that kind of alternative directly aids the plaintiff; he will no longer be burdened despite his innocence or excluded from the benefit despite his need. But a means which creates less *under*inclusion with respect to a burden (or *over*inclusion with respect to a benefit) might also be considered less restrictive. For example, if in *Craig* the state could have easily reduced drunk driving by the alternative means of education and stronger penalties, then the law's underinclusiveness *and* overinclusiveness might have been lessened: girls who could buy beer previously and would have been drunk drivers would be discouraged from driving while drunk, and boys who were safe drivers before but could not buy beer could do so.

But reducing the burden-underinclusiveness (or benefit-overinclusiveness) of a law is less restrictive only in the sense that the alternative is *fairer* than the original, not in the sense that the alternative burdens personal interests less severely. (In the above example, boys would not be more lightly burdened simply because the girls who posed a harm under the original law would not under the alternative.) Courts usually employ "less restrictive alternative" in its second sense, possibly because the phrase "less restrictive" suggests the meaning "burdening plaintiffs less" or because a trivial alternative can always be found which is less restrictive in the first sense. (In *Craig*, girls would be forbidden from purchasing beer.) Nevertheless, to the extent that an alternative reduces underinclusion courts should consider it less restrictive, since if it were adopted, the plaintiff would be less "burdened" in the sense of having a weaker equal protection claim: he has less reason to complain that the similarly situated are not treated equally.

It might be objected that in an equal protection case, courts should never characterize a less overinclusive alternative as less restrictive, since overinclusion (with respect to a burden) is essentially a *due process* problem. But the objection fails to note that a less overinclusive classification will necessarily produce a greater differential in harm and thus a lesser equal protection misfit. It also fails to note that a less overinclusive alternative burdens personal interests less, thus increasing the state's comparative position in the equal protection balance.

Court often notes how easily the legislature could achieve its goals by better fitting means.[387]

The lesser burden[388] interpretation of "less drastic means" appears less often. Its rationale is clear enough: even if an alternative means burdens the same class of persons, the means may be less restrictive if the burden it imposes on those persons is lighter, for the burden may implicate different and less important personal interests or may infringe the same interests less seriously. For example, one justification offered for the welfare residency requirement in *Shapiro v. Thompson*[389] was that it would prevent the fraudulent receipt of benefits from more than one state. But the Court noted: "Since double payments can be prevented by a letter or a telephone call, it is unreasonable to accomplish this objective by the blunderbuss method of denying assistance to all indigent newcomers for an entire year."[390] That is, the alternative would have imposed only a mild "burden," namely, a check on whether they received welfare benefits from another state. Although the same class of persons—all welfare recipients who arrived in the state within a year—would be burdened by this alternative, the enormous reduction in burden would make it much less restrictive.

With these explanations of "less restrictive" and "equally effective" in mind, we may turn to the merits of the problem: how do we determine whether an alternative is "reasonable"? "Reasonableness," of course, is not a static notion. If a law is almost prima facie unfair regardless of alternatives, then even the presence of a somewhat costlier alternative might invalidate the law. If the law was almost invalid even had there been no other means to achieve the end, then it would be anomalous to uphold the law simply because the alternative that does exist imposes more than nominal costs. In the converse situation, in which the law is almost prima facie fair, an alternative that imposes almost *any* significant costs on the state may be "unreasonable," for similar reasons.

The "reasonableness" of an alternative should depend not only on the prima facie fairness of the law, but also on the social costs of the alternative and how much less restrictive the alternative is. That is, the more "unfair" the law, the less restrictive the alternative, and the less the costs imposed by that alternative, the more willing courts

387. *See* note 379 *supra*.

388. "Burden" also refers to denial of a benefit. The "lesser burden" interpretation, as applied to benefits, means that the alternative does not completely *deny* benefits to underincluded would-be recipients; it gives a lesser benefit, or the same benefit subject to conditions not applied to other recipients. The remedy ordered in an irrebuttable presumption case is generally less restrictive in this sense, since it does not *grant* the benefit to plaintiff, but only allows him to rebut the presumption that he does not deserve it.

389. 394 U.S. 618 (1969).

390. 394 U.S. at 637.

should be to invalidate the law for failing to employ a less drastic means. In essence, we are simply comparing the general fairness of the law with the general fairness of an alternative[391] and asking whether the additional fairness achieved by the latter justifies some increase in social costs.

The social costs incurred by an alternative will mainly be of two kinds. One cost may be the decreased effectiveness of the alternative means. *Vlandis v. Kline*,[392] though nominally an irrebuttable presumption case, illustrates such a cost.[393] The Court found constitutionally infirm the statutory irrebuttable presumption that certain applicants for in-state tuition at a university were nonresidents. Specifically, married students who had legally resided outside the state when they had applied and unmarried students who had lived outside the state during the preceding year were deemed to be "nonresidents" for their stay at the university. The Court believed that a more refined test of residence was practicable. The Court carefully noted that it was *not* forbidding reasonable durational residence requirements that could be met while one was a student.[394] The alternative of, say, a one-year residence test,[395] even if it used the same presumptions as the statute in *Vlandis*, might increase social costs. Although the original four-year presumption was too overinclusive, it did screen out some nonresidents who would not have been screened out by a one-year presumption (assuming that after a year students can more easily qualify as residents). The alternative of a one-year presumption therefore alleviates somewhat less effectively the perceived "harm" of granting in-state tuition to nonresidents.

A second social cost an alternative can incur is its additional administrative cost. In *Vlandis* the alternative of requiring hearings at which individuals could rebut the presumption of nonresidence would not be significantly less effective in limiting in-state rates to bona fide residents. The state in *Vlandis* objected to this alternative mainly because of the second kind of social cost: the hearings mandated by the Court were more expensive than the presumption.[396]

391. The term "general" is significant, for the less restrictive alternative approach should *not* require the state to choose alternative B over existing means A if A and B are "equally" restrictive, even if the actual plaintiff is burdened by means A but not by means B. Giving this plaintiff a remedy would only create a new and equivalently "burdened" class of persons discriminated against.

392. 412 U.S. 441 (1973).

393. Similar principles apply to "necessity" problems in equal protection and in due process/irrebuttable presumption cases; the only difference is in the kind of misfit to be remedied (*i.e.*, differential harm vs. overbreadth).

394. 412 U.S. at 452.

395. *See* Starns v. Malkerson, 326 F. Supp. 234 (D. Minn. 1970), *aff'd. mem.*, 401 U.S. 985 (1971).

396. The two kinds of social cost created by an alternative are often interrelated, for an alternative that less effectively promotes the state's goal can usually be made more precise if

(This, of course, is the usual objection to the alternative of allowing a presumption to be rebutted.) As was suggested earlier,[397] the cost of hearings can be significant and should not be dismissed as blithely as it has been in many sex discrimination and irrebuttable presumption cases.[398]

Alternatives, it is clear, are not easily evaluated. Simple standards for evaluating alternatives that lower courts could readily and correctly apply are elusive (though this is a problem endemic to any refined, case-sensitive balancing test).[399] And such standards are necessarily difficult to apply because the actual effectiveness of alternatives can usually not be predicted. When a court evaluates the prima facie fairness of a law, it may have a legislative record, or at least detailed justifications by government counsel, describing the law's probable effect. In some cases the court may even have evidence of the law's actual effectiveness. But evaluations of alternatives that the court may have conceived sua sponte and that may never have been evaluated in the legislature must often rest on unwarranted factual assumptions. In these cases the dangers of judicial legislation loom large.

Nevertheless, a court which doubts its competence to judge alternatives and which fears it is arrogating the legislative function may rest its decisions simply on the prima facie fairness or unfairness of the law without engaging in a more detailed balancing. Moreover, some difficulties may be eased by shifting burdens of proof. One commentator has suggested that critics of a government regulation have the burden of showing by clear and convincing evidence that a less restrictive alternative exists.[400] In sharp contrast, the funda-

---

the state is willing to shoulder the extra administrative cost. Of course, this relationship does not invariably hold; the state may sometimes be forced to accept a less effective classification even when it cannot be made more precise. In other words, sometimes the state just may not have the second alternative of retaining the classification's level of effectiveness by increasing administrative cost.

397. *See* Part IV.B.2 *supra*.

398. Administrative cost includes not only the additional out-of-pocket expenses—hiring hearing examiners, evaluating more detailed and "individualized" categories of proof—but also less tangible costs—loss of predictability, possible delay in imposing the benefit or burden, and greater opportunity for abuse of discretion—which jeopardize both classificatory accuracy and individual freedom. Where a hearing is already provided to some and the alternative only requires that it be more complete or detailed or that it be extended to others, the burden to the government is not as great as when new administrative machinery must be established. These and other factors affect the significance which should be attached to the administrative cost of conducting hearings.

399. Consider, for example, the unhappy experience under the fifth amendment "voluntariness" test for confessions, a test ultimately rejected in favor of the per se rule of Miranda v. Arizona, 384 U.S. 436 (1966). *See* W. LOCKHART, Y. KAMISAR & J. CHOPER, CONSTITUTIONAL LAW 655-57 (4th ed. 1975).

400. Struve, *The Less-Restrictive-Alternative Principle and Economic Due Process*, 80 HARV. L. REV. 1463, 1471 (1967). For a general discussion of burdens of proof under the minimum rationality standard, see P. BREST, *supra* note 116, at 1005-10.

mental interest test apparently requires the *government* to prove that there are no less restrictive means.[401] This approach cannot be taken too literally since a party cannot reasonably be asked to anticipate and refute every possible challenge to its position. But if the government is asked only to disprove plausible alternatives proffered by the plaintiff or the court, burden of proof will have been reasonably allocated in cases where an important personal interest is significantly infringed or where the classification is almost prima facie unfair. Where lesser personal interests are infringed, it is sensible to require plaintiffs to prove that a less restrictive alternative exists.

3. *When Perfect Fit Is Not Enough*

Under an equal protection balancing test with "bite," perfect fit might not be enough to validate a law if the state interest "perfectly" served is insufficiently important compared to the burden on personal interests. The Supreme Court's existing tests do appear to transcend perfect fit, for while most classifications need only relate to a "legitimate" state interest, a gender-based classification must serve an "important" goal, and classifications touching fundamental interests (or employing suspect traits) must be necessitated by a "compelling" interest, irrespective of fit with a lesser goal. A good example of this "transcendental" technique was the Court's decision that once a state grants criminal defendants an appeal as of right, it must also appoint counsel for the indigent to ensure that the poor as well as the rich may have a meaningful appeal.[402] The state could have argued[403] that its failure to furnish free counsel was "perfectly" related to the interest of having defendants pay their own way in the criminal justice system. But such an interest is too slight to justify differential treatment with respect to the important personal right at issue, namely, the meaningful chance to avoid the severe consequences of conviction.[404]

The view that perfect fit with respect to a non-compelling interest may not be enough to satisfy equal protection has been criticized as a

401. *See Developments, supra* note 1, at 1122.

402. Douglas v. California, 372 U.S. 353 (1963).

403. There is no indication in the opinion that this argument was in fact made. But a similar argument was propounded by Justice Harlan in his dissent in Rinaldi v. Yeager, 384 U.S. 305, 311 (1966). The majority had held that a state could not recoup the cost of trial transcripts furnished to indigent incarcerated prisoners yet decline to seek such reimbursement from indigents who are not imprisoned. Harlan reasoned in dissent:

> Surely the State might reasonably choose to reimburse itself for such transcript costs out of prison allowances, but deem it not worth the added time and trouble, *or even advisable*, to attempt to extract such charges from a convict not in prison who must support himself on his own resources.

384 U.S. at 311 (emphasis added).

404. *See* Simson, *supra* note 5, at 691-94.

barely disguised return to substantive due process.[405] This criticism is not strictly accurate, since equal protection *may* properly employ this transcendental method, but it is true that that method is usually more appropriate in due process cases. A sensitive understanding of substantive due process, of equal protection balancing, and of the definition of "legitimate" discriminations, makes less necessary any resort to the radical notion that "perfect fit is not enough."

An insufficiently developed equal protection balancing test partly accounts for the hierarchy of purposes—from legitimate, to important, to compelling—in the current equal protection tests. Almost all classifications create some misfit under the model presented in this Note. That is, most laws are not, on their face, self-justifying. Further analysis is required to find a classificatory *purpose* that is "legitimate," and there will then ordinarily be some misfit to evaluate. The question whether that misfit is justifiable is, as we have seen, answered through a balancing test that weighs the differential in harm between the included and the excluded according to the strength of the government and personal interests.[406] Thus, when the government's interest is insignificant, the differential in harm may be insufficient to validate the law, and there may be no need to move up the hierarchy and examine misfit only with respect to a more important purpose.

One important reason why the Court has used this "more than perfect fit" approach is that the Court misconceives saving administrative costs as the same kind of government interest as the substantive purpose of the statute. If saving cost is described as *the* purpose of a classification, then it is easy to describe "*the* purpose" as insignificant irrespective of the closeness of the fit. Yet laws are usually not passed in order to save money but in order to *do* something, subject to a cost constraint. Cost constraints are not unimportant, but every law costs something, and costs will always affect the evaluation of the reasonableness of alternatives. Perhaps the Court requires an "important" interest in gender-based classifications to underscore its view that administrative cost is an insufficient rationale. But all of the gender cases reaching the Supreme Court have involved concededly inaccurate generalizations with respect to the substantive goal of the statute; since the means invariably did not perfectly fit the goal, the "more than perfect fit" approach in these cases has been unnecessary.

As this discussion suggests, courts may also have resorted to "more than perfect fit" because of their tendency to ascribe a narrow and not "legitimate" purpose to a statute. In *Douglas*, for example, if

405. *See Developments, supra* note 1, at 1132.

406. Also relevant, of course, are the extent to which the government interest is served and the extent to which the personal interest is burdened.

the purpose is described as the desire to save the money that providing free counsel would cost, the purpose is not legitimate.[407] The Court must reach a higher or more generalized level of purpose before evaluating misfit—not because "perfect fit is not enough" at this level, but simply because there is no reason to credit a purpose that is not obviously legitimate. (It would not be proper to define the purpose of our driving test as screening from the road people who fail driving tests and then to analyze that purpose as insignificant.) In short, the requirement that misfit be evaluated only with respect to legitimate goals largely serves the function courts believe the "perfect fit" approach performs.

Despite these observations, the "more than perfect fit" approach is significant when even a legitimate purpose is too insubstantial. This approach is also important when the goal is "inevitably arbitrary" in the special sense discussed above,[408] for such goals by definition cannot sensibly be evaluated under a balancing test, one of whose elements is misfit—their misfit cannot even be measured. Thus, such goals should be treated as though the classificatory fit were perfect and should be invalidated only when perfect fit is not enough.[409] If, for example, a school principal decided that physical segregation of student protesters wearing black armbands[410] was aesthetically pleasing, this inevitably arbitrary discrimination could only be invalidated on the ground that the speech and privacy interests of the students required that a more important interest be invoked. It would be futile to dispute the aesthetic pleasure the principal derived from the scheme; the closeness of fit is unmeasurable.

When equal protection demands more than perfect fit, the balance resembles straightforward substantive due process balancing. But, in theory and occasionally in practice, the tests are distinct. They diverge when an *equal* burdening of a fundamental or important interest does not quite violate substantive due process, but a differential burden does violate equal protection. For example, the Supreme Court in *Douglas* only held that an indigent should have the same "meaningful" access to the courts that the state afforded a wealthy defendant. It did not hold that "meaningful" access is a

407. The purpose is not legitimate because it does not explain why the costs of an indigent's appeal should be imposed on the indigent rather than on society generally.

408. *See* Part V *supra*.

409. In the case of laws with neutral goals (such as social experiments), this statement should be qualified: the law should be invalidated only when the *combination* of perfect fit toward the substantive goal *and* the independent state interest in classifying neutrally (*e.g.*, the interest in obtaining knowledge from an experiment) are insufficient to outweigh the individual's side of the equal protection balance. *See* text at notes 339-42 *supra*.

410. *Cf.* Tinker v. Des Moines Independent Community School Dist., 393 U.S. 503 (1969) (recognizing free speech right of high school students to wear black armbands in nondisruptive antiwar protest).

substantive requirement, and it thus apparently did not forbid the state to deny the use of transcripts on appeal to rich and poor alike.[411] Similarly, in *Shapiro v. Thompson*[412] the Court purported to strike down a residency requirement for welfare benefits only because the fundamental right to travel was *differentially* infringed. The Court presumably would have *upheld* the requirement had the state also imposed qualitatively similar burdens on new residents other than potential welfare applicants.[413]

As these cases illustrate, the theoretical distinction between equal protection and substantive due process in this area often lacks practical substance. A more equal infringement of the right to travel than occurred in *Shapiro* would probably not be constitutional. And the Court's frequent assertion that there is no constitutional right to appeal a criminal conviction[414] can probably not be taken at face value. Even if that assertion is correct, a state which has granted an appeal as of right may probably not encumber it to the extent of denying defendants the use of trial transcripts. But in some areas, such as voting rights, the due process/equal protection distinction[415] does seem to have practical as well as theoretical meaning. Many state offices can be appointive, but once the state makes them elective it may not restrict any group's interest in voting in elections for those offices, absent a compelling state interest.[416]

Despite these objections, the notion that equal protection may require more than perfect fit is sometimes of practical consequence. How should this notion be expressed in an actual balancing test? One possibility would be to establish a finely graded continuum in

411. Support for this observation is perhaps to be found in Mayer v. City of Chicago, 404 U.S. 189 (1971), which required that an indigent defendant be furnished with a sufficiently complete trial record to permit proper consideration of his claims even though the conviction was punishable by fine only—and even though the right to counsel (a substantive right) has not yet been extended to such convictions. *See* Argersinger v. Hamlin, 407 U.S. 25, 37 (1972) (noting that the question has yet to be decided).

412. 394 U.S. 618 (1969).

413. Although these cases did not involve perfect fit, the Court's analysis in effect assumed that the fit was perfect.

414. *See, e.g.*, Ross v. Moffitt, 417 U.S. 600, 606 (1974), citing with approval McKane v. Durston, 153 U.S. 684 (1894).

415. The distinction might be further refined to the following: substantive rights/due process/equal protection. Even though a citizen may have no substantive right (*e.g.*, under the first or sixth amendment) to obtain a certain benefit, it might be a denial of (substantive) due process or equal protection to *condition* the grant of the benefit in certain ways, for instance by invading substantive due process interests or by distinguishing on the basis of race or wealth. Notice that substantive due process can sometimes be the source of a right to the particular benefit as well as the source of a right not to have the benefit conditioned upon the nonassertion of protected interests. For example, as just noted, due process might plausibly support both a constitutional right to a criminal appeal and a constitutional right not to be penalized for the assertion of a statutory appeal right, *see* North Carolina v. Pearce, 395 U.S. 711, 725 (1969) ("vindictiveness against a defendant for having successfully attacked his first conviction must play no part in the sentence he receives after a new trial").

416. Kramer v. Union Free School Dist. No. 15, 395 U.S. 621, 628-30 (1969).

which any increase in the significance of the personal interest (or the extent to which it is burdened) would require a more significant "threshold" state interest below which even perfect fit would not validate the law. Of course such a continuum would not be judicially manageable, and the Court should require only a small number of intervals. (The present test has three intervals,[417] "legitimate," "important," and "compelling," though the middle interval is applied not to intermediate-level interests but to an intermediately "suspect" classificatory trait, sex.) A two-, three-, or four-tiered test is perhaps appropriate here—but unlike the present tiers, these tiers would describe only the importance of the requisite government interest and not the required closeness of fit or the availability of alternatives. Within any single tier, a careful means-end analysis would still be necessary.

In establishing these thresholds, care should be taken not to require too compelling a purpose at the upper tier lest equal protection mandate pervasive equality. Under the present strict scrutiny test, for example, the compelling state interest requirement is nearly impossible to satisfy.[418] The state must treat similarly all who are exercising their fundamental interests in voting, travelling, seeking access to the criminal justice process, entering marriage, and perhaps procreating, and it may not draw merely "rational" distinctions based on wealth or other criteria. Such enforced equality differs from the most invigorated rational basis test; it requires that, in addition to all its other goals, the government attempt to ensure that everyone realizes fundamental interests equally. As this approach more absolutely protects certain rights, it not only increasingly resembles a due process approach, but eventually approximates the degree of scrutiny accorded explicit substantive constitutional provisions such as the first amendment.

In sum, courts should employ the equal protection "perfect fit is not enough" approach cautiously. There are probably few interests unimportant enough that they may constitutionally be infringed equally yet important enough that unequal infringement violates equal protection regardless of the rationality of the classification in other respects. Courts should be more willing to use either substantive due process or a rational basis test which evaluates misfit to invalidate classifications they now strike down on "fundamental interest" grounds. This franker technique would probably not pro-

417. Simson, *supra* note 5, at 679, also suggests a three-tiered test, one which distinguishes between compelling, significant, and insignificant state interests. His fourth category of state interest, "unlawful," is not really a fourth "tier," for it describes a tier at which the individual always wins. *Id.* at 680 n.95.

418. The compelling interest requirement has been satisfied only in two cases, each of which involved wartime necessity: Korematsu v. United States, 323 U.S. 214 (1944), and Hirabayashi v. United States, 320 U.S. 81 (1943).

duce greatly different results, since a court more closely scrutinizing purposes by the method suggested above will almost invariably find some misfit to weigh against the purported governmental purpose. This technique would also encourage meaningful case-by-case review. Under current equal protection analysis, the Court is sometimes paralyzed when it feels it cannot accept the consequences of another "fundamental interest" baptism,[419] and it has been embarrassed by its resort to tortured reasoning to sustain laws that affect fundamental interests indirectly.[420] Under the proposed balancing test, perfectly tailored classifications would seldom be invalidated solely because they infringe important personal interests. The test provides a sensitive and logically consistent solution to the problem of protecting lesser personal liberties against discriminatory encroachment.

C. *Substantive Due Process Balancing*

A reasonable formula for substantive due process balancing may be inferred from the preceding section. Due process balances the personal interests implicated and the extent to which they are burdened, against the state interest actually served by the means employed. The balance is that expressed by the equation earlier noted.[421] It is a direct balance in which differential harm, or indeed any reference to an unclassified group, is irrelevant.[422]

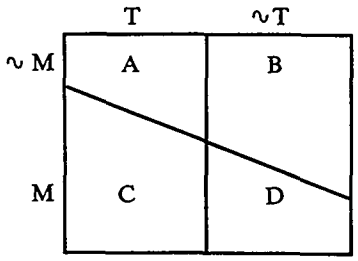
Many of the same concepts used to analyze equal protection differential treatment apply directly to due process overbreadth. Similarly, under substantive due process as well as under equal protection, "perfect fit" may not be enough to validate a law. Indeed,

419. *See* Dandridge v. Williams, 397 U.S. 471 (1970).

420. *See, e.g.*, Sosna v. Iowa, 419 U.S. 393 (1975); Ross v. Moffitt, 417 U.S. 600 (1974); Rosario v. Rockefeller, 410 U.S. 752 (1973).

421. *See* text at note 353 *supra*.

422. Such a reference *is* relevant to a limited extent, namely, to determine how *much* of the state interest is actually served by the means employed (C/(C + D), which is equivalent to C/M):



See text at notes 369-71 *supra*.

requiring more than perfect fit is more often a sensible mode of analysis in the former case than in the latter.[423]

Due process overbreadth analytically resembles equal protection differential treatment. First, a court evaluates the extent of the overbreadth and estimates its legitimacy in light of the personal interests burdened and the state interests achieved. At the "extremes" of legitimacy, the law is either upheld or invalidated, whatever the possible alternatives. Between those extremes, the court investigates reasonable alternatives, testing reasonableness according to the prima facie fairness[424] of the law, the social costs of its alternatives, and the restrictiveness of the alternatives.

Despite these formal similarities between equal protection and due process balancing, weights assigned the balanced interests may differ in the two cases, since the consequence of a due process violation is more severe than that of an equal protection violation. Ceteris paribus[425] the personal interest should have to be more important (or more heavily burdened) in a due process than in an equal protection case to constitute a violation. By the same token, the state interest should have to be more compelling (or better furthered) in a due process than in an equal protection case to justify what would otherwise be a violation. Were this not so, the equal protection clause would be gratuitous, since a liberal due process clause would invalidate even a uniform infringement of a right, and there would be no need to examine differential infringement. But precision as to *how* different the weighting of the interests should be in the two kinds of balance is elusive; the difference ultimately depends upon the relative strength with which the Court wishes to read the due process and equal protection clauses. Nevertheless, the Court needs to be sensitive to these relationships between due process and equal protection and the different functions which they perform.

Finally, as the last section suggests, in due process even more than in equal protection perfect fit might not be enough—in due process personal interests of a given level of significance must be justified in terms of state interests of a corresponding threshold significance, even if a lesser state interest is adequately, even perfectly,

423. Strictly speaking, "perfect fit" has a different meaning in the two cases. In due process, it only means the absence of any overbreadth, regardless of the amount of underinclusion. In equal protection, it means the greatest possible differential in harm between the burdened and those not burdened. This amounts to "perfect fit" as normally understood, for all and only the harmful are burdened, and there is no overinclusion or underinclusion.

424. In due process cases, a law is prima facie fair if the degree of overbreadth is not excessive; in equal protection cases, if the degree of differential harm is sufficient.

425. Of course, "comparing" the degree of overbreadth with the degree of differential harm is difficult. But the intuitive idea is that if these are roughly equivalent, then the relationships stated in the text *infra* should hold.

served. This straightforward balancing makes more sense under due process than under equal protection, for due process tends to protect personal interests more absolutely. Indeed, the uneven continuum of equal protection, substantive due process, and explicit constitutional guarantees such as the first amendment,[426] suggests that a direct balance of the relevant interests is more justifiable as the personal interest comes to deserve more absolute protection. An infringement of *anyone's* free speech rights demands an extraordinary justification. An intrusion into the autonomy of the family, by contrast, might be valid if the state could adduce a compelling interest, but it is otherwise invalid, even if there is no overbreadth.[427] An invasion of a mere equal protection interest, such as the right to procreate, deserves little absolute protection; the plaintiff probably has to show that the invasion was accomplished discriminatorily.[428]

Because due process demands more than perfect fit much more often than equal protection, misfit analysis is often less important in a substantive due process case. The Court will demand a "compelling" justification and will ignore less substantial state interests or will be more critical of the misfit created by lesser interests than under equal protection. Thus, to the extent that the Note's emphasis on misfit analysis suggests that due process invariably requires a judgment about overbreadth, this Note is misleading.[429] Nevertheless, a more active use of the due process clause to protect lesser personal interests would require a much closer attention to overbreadth analysis.[430] Both because of this possibility and because due process misfit analysis offers an instructive contrast to equal protection, overbreadth analysis is valuable.

---

426. See the taxonomy in P. BREST, *supra* note 116, at 805-06. This Note would interpose a substantive due process tier between his category 1 ("substantive constitutional rights") and his category 2 ("rights to fair and rational treatment"), which recognizes certain substantive rights of lesser constitutional significance than many of his category 1 rights, but of greater significance than his category 2 rights.

427. *See* Moore v. City of East Cleveland, 431 U.S. 494 (1977) (invalidating such a law by a simple analysis of overbreadth). Significantly, the Court did not find any compelling state interest to justify the discrimination.

428. *See* Dandridge v. Williams, 397 U.S. 471 (1970) (upholding a maximum grant limitation on welfare benefits irrespective of family size, even though effect of statute may have been to deter procreation); Skinner v. Oklahoma, 316 U.S. 535 (1942) (invalidating a compulsory sterilization law which arbitrarily distinguished between, *inter alia,* larceners and embezzlers; but upholding the general constitutionality of compulsory sterilization by its approving reference to Buck v. Bell, 274 U.S. 200 (1927)). *See also* Roe v. Wade, 410 U.S. 113, 154 (1973) (citing *Buck* in support of the proposition that one does not have "an unlimited right to do with one's body as one pleases").

429. Roe v. Wade, 410 U.S. 113 (1973), is a good example of a substantive due process case in which overbreadth analysis was irrelevant. The crucial analysis involved a direct balancing of interests to determine when each interest became "compelling" enough to outweigh the others. 410 U.S. at 162-64.

430. The irrebuttable presumption doctrine is, of course, an example of such activism.

VII. Conclusion: The Judicial Role

This Note does not address in any detail the judiciary's proper role in enforcing the equal protection guarantee. This Note does attempt to construct a general analytic framework for a moderately activist court, a framework amenable to a broad spectrum of judicial views. The Note's "balancing" terminology should not be misinterpreted as requiring an activist approach to equal protection; it is meant to explain what the existing tests actually do and to further conceptual coherence. Courts may differ as to which interests deserve special solicitude and as to how much solicitude they deserve.[431] Courts may differ as to the relative usefulness and propriety of equal protection and substantive due process rationality tests. They may differ as to the propriety (or necessity) of examining a legislature's motive. They may differ, most fundamentally, as to the relative importance of "hostile" and "unfair" discriminations. But the proposed model should be able to accommodate these differences.

"Closer scrutiny" is not reducible to a simple formula. A court which is seriously evaluating the fairness of a law must assess the differences between classes with respect to the harm posed, the interests aggrieved, and the costs incurred. The courts cannot hope for mathematical nicety[432] in these judgments, especially where classifications legitimately and inevitably reflect arbitrary choices. Nevertheless, the judgment *can* be a judgment rather than an idle invocation of slogans.

A revivified case-by-case approach to equal protection promises more genuine review; it also poses a danger of unprincipled decisions. This danger is often overdrawn, however. More exacting equal protection review should inspire not ad hoc decisions, but a body of standards applicable to different legislative subjects, personal interests, and classificatory traits—standards formulated in light of the considerations this Note has set forth. For example, discrimination against women is now treated under an "intermediate" standard: the classification must be "substantially" related to achieving "important" government objectives.[433] Such a standard, unless

431. Compare the majority opinion with Justice Marshall's dissent in San Antonio Independent School Dist. v. Rodriguez, 411 U.S. 1, 70 (1973). Commentators differ as to the central values to be protected under the equal protection clause. *See e.g.*, Karst, *The Supreme Court, 1976 Term—Foreword: Equal Citizenship Under the Fourteenth Amendment*, 91 Harv. L. Rev. 1 (1977); Perry, *supra* note 68 (distinguishing several forms of fairness rooted in "conventional morality"); Sandalow, *supra* note 7, at 670 (one such value is "to facilitate each individual's opportunity to pursue a personal conception of the good life"); Wilkinson, *supra* note 3, at 954-56 (distinguishing rights of political participation, equality of opportunity, and economic equality).

432. *See* note 55 *supra*.

433. Craig v. Boren, 429 U.S. 190, 197 (1976).

rigidly and unthinkingly applied, is not inconsistent with the view that there is in principle only one equal protection standard,[434] since the Court has concluded that sex discriminations are as a rule so unreliable that they must be justified by more than a merely conceivable set of facts.[435] And finally, however difficult it may be to fashion equal protection rules that are not only principled but also plain and workable, the present two-tiered test can no longer be retained—the methods it prescribes are not those the Court actually uses.[436]

The nagging concern that the equal protection clause raises questions courts are ill-equipped to answer perhaps explains why existing equal protection tests are mechanical and result-oriented. This concern has also been overdrawn. The Court's most important task is not to gauge numerical misfit; it is to estimate misfit roughly according to the Court's sense of how important or contrived the competing interests are. The Court can defer to the fact-finding competence of the legislature, and it can assign the plaintiff the burden of showing that a social ill does *not* exist, or would not be effectively remedied by the means chosen, or could be effectively remedied by a less restrictive alternative. And thus courts can answer the objection that judicial review of a law's fairness substitutes a court's understanding of the costs and benefits of social issues for the legislature's.[437]

Uneasiness about judicial legislation has also been fueled by the fear that judicial review will hamper legislative flexibility and freedom. That fear is hardly confined to the equal protection clause but may be particularly discomfiting in relation to it, since the *values* to be "protected equally" are not described in the fourteenth amendment.[438] But a principled theory for determining those values[439] is surely preferable to judicial abdication.

Moreover, the concern for legislative flexibility will be exaggerated if the limits of equal protection balancing are not appreciated. A court employing this Note's framework would usually defer to the state's finding that an evil exists and that the state's means alleviates it.[440] A state is entirely free to pursue any "legitimate" goals[441] except in those rare cases in which the law infringes an important per-

434. Craig v. Boren, 429 U.S. 190, 211-12 (1976) (Stevens, J., concurring).

435. *Cf.* McGowan v. Maryland, 366 U.S. 420, 426 (1961) ("[a] statutory discrimination will not be set aside if any state of facts reasonably may be conceived to justify it").

436. *See* Massachusetts Bd. of Retirement v. Murgia, 427 U.S. 307, 318-21 (1976) (Marshall, J., dissenting).

437. *See* Trimble v. Gordon, 430 U.S. 762, 784 (1977) (Rehnquist, J., dissenting).

438. *See* Professor Sandalow's discussion of "material principles," *supra* note 7, at 654-63.

439. *See* Sandalow, *supra* note 7, at 660-63.

440. *See* text at notes 256-61 *supra*.

441. *See* Part III.A.1 *supra*.

sonal liberty and must therefore be justified by a significant goal.[442] And if the goal is legitimate, the state may deploy any of the means of reaching that goal which are "accurate" enough to satisfy the appropriate level of scrutiny. Finally, as Justice Jackson wrote, equal protection does not prevent the government from acting but only requires that its actions be even-handed.[443] Though the statute may have been partly intended to avoid even-handedness, Justice Jackson does remind us that equal protection interferes with substantive legislative choices less than due process and many more explicit constitutional constraints do, since it examines only the justification for *unequally* pursuing a goal.

Reasonable persons differ over the question whether the equal protection clause prohibits only totally "irrational" classifications and discrimination against suspect groups. But if the Court wishes to subject laws to closer scrutiny than the traditional rational basis test requires, there are principled and practical ways of doing so. The Court's own decisions indicate its interest in a strengthened standard. There is no reason not to proceed.

---

442. *See* Parts VI.B.3 and VI.C *supra*.

443. Railway Express Agency v. New York, 336 U.S. 106, 112-13 (1949) (Jackson, J., concurring).