NOTE

A LAW AND NORMS CRITIQUE OF THE CONSTITUTIONAL LAW OF DEFAMATION

Michael Passaportis[*]

[*] J.D., University of Virginia School of Law, May 2004. I would like to thank Professor Lillian BeVier for her invaluable guidance in writing this Note, Professor Robert Cooter for planting the seed of an idea, Lee Kovarsky for revealing to me the niceties of law and economics, and Fern Mechlowitz for her patience, support, and proofing skills.

INTRODUCTION

THE Federal Constitution prohibits laws that abridge free speech.[1] State defamation laws forbid utterances that falsely vilify community members.[2] An obvious facial tension exists between these aspects of federal and state law, both of which implicate speech. To be sure, the Supreme Court has never adopted the absolutism that the text of the First Amendment suggests.[3] In fact, defamation remained outside the Court's free-speech jurisprudence for the better part of the twentieth century.[4] When these two spheres of law eventually collided, however, the federal right to speak freely about one's peers ran roughshod over the state interest in keeping such speech truthful.[5] Precise formulations of the latter interest often appear anachronistic; talk of the "honor and reputation of the citizenry"[6] conjures up images of chivalry and good manners, far from the hardy individualism of modernity. The countervailing right to speak freely, by contrast, has inspired some of the most grandiloquent declamations in modern constitutionalism. If the ascendance of an invaluable federal right over an antiquated state interest characterizes the modern law of defamation, perhaps the Court conferred a social benefit by replacing the common law strictures with a more speech-protective regime.

This Note will argue the contrary: the federal right's colonization of this particular sphere of the common law accomplished mischief. Reputation is not the outmoded, insipid, and purely personal asset that the Court's jurisprudence suggests. Rather, it is a useful social mechanism that checks community members' selfishness and en-

[1] See U.S. Const. amend. I. The Court incorporated the First Amendment against the states in *Gitlow v. New York*, 268 U.S. 652, 666 (1925).

[2] See infra notes 63–69 and accompanying text.

[3] Although Justice Black did campaign for such an interpretation. See, e.g., Am. Communications Ass'n v. Douds, 339 U.S. 382, 448 (1950) (Black, J., dissenting) ("[T]he First Amendment forbids compromise."). See generally Charles A. Reich, Mr. Justice Black and the Living Constitution, 76 Harv. L. Rev. 673, 695–97 (1963) (discussing Justice Black's First Amendment absolutism).

[4] See infra notes 88–89 and accompanying text.

[5] See infra Section V.B.

[6] El Vocero de Puerto Rico v. Puerto Rico, 508 U.S. 147, 149 (1993).

courages cooperative behavior. Citizens who desire social commendation find public-spirited behavior more attractive. Reputation increases such behavior by amplifying the social benefits accruing to the actor. In this manner, pursuing a good name contributes to social welfare. The Court failed to recognize this beneficial aspect of reputation when it weakened the common law of defamation. As a consequence, the balance of competing interests underlying the constitutional replacement yielded too much to free speech and retained too little for reputation.

To elucidate the social utility of reputation, this Note will begin by explaining the social perplexities that reputation overcomes—collective action problems. Part I of the Note will summarize conventional rational choice theory's understanding of collective action problems. Attempts to explain community cooperation in the face of such problems, from within rational choice theory, have met with limited success. In response, one school of thought has relaxed customary rational choice premises to account for social cooperation. Part II will describe the law and norms literature, which complements the traditional premises with the sociological insight that community judgment circumscribes human behavior. This new field of legal theory explains how such judgment can encourage cooperative behavior and efficiently resolve social dilemmas. The beneficial effect of norms depends on the social circulation of true reputational information about community members. For the law and norms literature, therefore, false reputational information has the negative externality of indirectly preventing the resolution of collective action problems.

Part III will explain the mechanics of this externality, noting how community members have incentives to instigate defamatory falsehoods about their peers. The externality of false negative speech implies that the law should harshly punish communication of such speech. Efficient social norms require strict defamation regimes. Despite this requirement, the actual law of defamation has evolved in the opposite direction. Part IV will summarize the strictures of defamation at common law and the weakening of this tort at the hands of the federal Constitution in the mid-twentieth century. The constitutional law of defamation rests upon a defensible understanding of the incentives to engage in political speech and the effect of defamation law on such incentives, which Part V will reca-

pitulate. Part V will also describe the Court's evident failure to consider the full reputational costs of false negative speech. As a result, the particular balance of interests underlying the constitutional law of defamation yields too much to free speech and retains too little for reputation. Part VI will provide a graphical rendition of this analysis and suggest that the constitutionalization of defamation may have upset a remarkably efficient common law regime.

I. COLLECTIVE ACTION PROBLEMS AND RATIONAL CHOICE THEORY

Social dilemmas—situations in which "the action required for achieving the collectively best outcome or goal is different from (and in conflict with) the action required for achieving the individually best outcome"[7]—have long attracted academic attention. The interest in studying these dilemmas is due to their apparent ubiquity in social life.[8] For instance, the "tragedy of the commons,"[9] an expanded application of the basic prisoner's dilemma, predicts overuse of public goods—an outcome that regularly occurs in contemporary societies.

Models of collective action problems also provide an elegant and seemingly rigorous justification for state interventionism. Traditional public choice theorists envisioned governmental regulation as the solution to social dilemmas.[10] By altering actors' payoffs, and by making cooperation more attractive and defection more costly,

[7] Raimo Tuomela, On the Structural Aspects of Collective Action and Free-Riding, 32 Theory & Decision 165, 166 (1992).

[8] See Robert Axelrod, The Evolution of Cooperation 7 (1984).

[9] See Garrett Hardin, The Tragedy of the Commons, 162 Sci. (n.s.) 1243, 1243–44 (1968). The following payoff matrix and inequalities describe the general conditions for this type of dilemma:

		Group Action	
		n or less choose cooperation	More than *n* choose cooperation
Individual Action	Cooperate	*X*	*X* + *Y*
	Defect	0	*Y*

$Y > (X + Y) > 0 > X$.

[10] See, e.g., Douglass C. North, Structure and Change in Economic History 20–32 (1981) (defending the efficiency of governmental intervention to resolve social dilemmas).

legal sanctions ensured the optimal resolution of these problems—punishment prevented defection.

Despite the attractiveness of this concise etiology of law, solutions to collective action problems often occur in the absence of central regulation. Psychological experiments, for instance, consistently reveal cooperation in noncoercive experimental settings. These experiments generally construct pecuniary payoffs to make defection or free-riding the dominant strategy, but cooperation the efficient outcome.[11] Despite rational choice predictions, researchers have found that "when pecuniary incentives appear to compel defection, 'many subjects do *not* defect.'"[12] Real-world examples also exist: the civil rights movement, for instance, required that a small group of dedicated activists risk large personal costs to achieve benefits for whole minority communities.[13]

Attempting to explain this "residuum of cooperation,"[14] rational choice theorists distinguished one-shot games from repeated interactions. In one-shot games, where the players did not expect to meet again, defection remained the dominant strategy.[15] In repeated games, however, the possibility of long-term gains made cooperative strategies more attractive.[16] Simulations seemed to demonstrate the costliness of uniform defection, and that more co-

[11] In the "public goods" experiment, for instance, participants are to contribute a portion of a five dollar endowment to a group project. The contributions are placed in envelopes, which the researcher opens, totals, doubles, and then distributes evenly among the participants. The "tragedy of the commons" logic dictates that each participant will retain his five dollars in an attempt to receive fifty cents to the dollar of the other participants' total contribution. See John O. Ledyard, Public Goods: A Survey of Experimental Research, *in* The Handbook of Experimental Economics 111, 111–13 (John H. Kagel & Alvin E. Roth eds., 1995).

[12] Richard H. McAdams, Cooperation and Conflict: The Economics of Group Status Production and Race Discrimination, 108 Harv. L. Rev. 1003, 1011 (1995) [hereinafter McAdams, Cooperation and Conflict] (quoting John M. Orbell et al., Explaining Discussion-Induced Cooperation, 54 J. Personality & Soc. Psychol. 811, 811 (1988)).

[13] See, e.g., Genna Rae McNeil, Charles Hamilton Houston: 1895–1950, 32 How. L.J. 469, 473 (1989) ("At that time, to champion civil and human rights for persons of African descent was always dangerous and costly.").

[14] McAdams, Cooperation and Conflict, supra note 12, at 1012.

[15] See Thomas C. Schelling, Micromotives and Macrobehavior 216–19 (1978).

[16] See Axelrod, supra note 8, passim (describing various possible strategies and testing them against each other to determine relative success and robustness); see also Randal C. Picker, Simple Games in a Complex World: A Generative Approach to the Adoption of Norms, 64 U. Chi. L. Rev. 1225, 1286 (1997) ("[I]nteractions will often lead to convergence on the superior norm.").

operative behavior actually maximized long-run returns regardless of peer strategy.[17] Subsequent studies, however, suggested that these observations rested on unduly optimistic assumptions unlikely to hold in the real world.[18]

The attempt to explain cooperative behavior from within rational choice theory thus met limited success. The existence of such behavior in human affairs, in turn, suggested either that the theory's assumptions were inadequate or that state intervention was indeed necessary to enforce social cooperation.

While some rational choice theorists adopted the latter conclusion,[19] behavioral economists explored the former conclusion and sought to adjust rational choice theory assumptions to explain the divergent empirical findings of pervasive cooperation. These adjustments often drew on insights from other social sciences and generally questioned the tenets of strict rationality and the exclusive pecuniary motivation on which rational choice theory rested.[20]

II. BEHAVIORAL ECONOMICS AND NORMS

One prominent avenue of academic investigation addressed rational choice theory's conception of personal motivation. Commentators questioned whether social actors only pursued pecuniary

[17] Axelrod, supra note 8, at 45. In particular, "tit for tat," a strategy in which a player cooperates, but defects after a co-player's defection and then cooperates only after the co-player's cooperation, scored the highest average in a tournament against a variety of other strategies. Id.

[18] See Paul G. Mahoney & Chris W. Sanchirico, Competing Norms and Social Evolution: Is the Fittest Norm Efficient?, 149 U. Pa. L. Rev. 2027, 2029 (2001) (noting that cooperative norms achieve evolutionary success only when their efficiency gains exceed the mismatch risk of non-cooperative norms).

[19] See, e.g., id. at 2058 (claiming that the improbable conditions for social cooperation justify governmental intervention); see also Gregory Mitchell, Why Law and Economics' Perfect Rationality Should Not Be Traded for Behavioral Law and Economics' Equal Incompetence, 91 Geo. L.J. 67, 69–71 (2002) (arguing for the retention of rational choice theory's assumption of perfect rationality on the basis of its coherence).

[20] See, e.g., Russell B. Korobkin & Thomas S. Ulen, Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics, 88 Cal. L. Rev. 1051 (2000) (critiquing rational choice theory's assumption of perfect rationality and calling for further scholarship). See generally Donald C. Langevoort, Behavioral Theories of Judgment and Decision Making in Legal Scholarship: A Literature Review, 51 Vand. L. Rev. 1499 (1998) (summarizing the application of behavioral social science to law and economics).

payoffs.[21] Drawing on insights from sociology, psychology, and anthropology, these theorists posited a more catholic universe of human goals.[22] In legal literature, such nonconformism has created the burgeoning field of law and norms. One model of norms proceeds from the sociological insight that community judgment circumscribes human behavior.[23]

The chief exponent of this model criticizes rational choice theory for failing to recognize that people "receive benefits from cooperation, or avoid costs from defection, that are not part of the formal, pecuniary payoff structure of the game."[24] Conceding that material self-interest counsels defection as the dominant strategy *in vacuo*, Professor McAdams argues that fellow game-players' perceptions reduce the defection payoff and increase the gains from cooperation when games are set in a social milieu.[25] He equates these additional costs and benefits with the intuitive concept of esteem.[26] Esteem, or perception of "relative social rank or social status," is a "basic pleasure," and a social actor experiences disutility when her peers observe her performing a shameful act.[27] Community judg-

[21] See, e.g., Joseph Henrich et al., In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies, 91 Am. Econ. Rev. (Papers & Proc.) 73, 77 (2001) (noting that a study of small societies demonstrated that "the canonical model of the self-interested material payoff-maximizing actor is systematically violated").

[22] See, e.g., Linda Hamilton Krieger, The Content of Our Categories: A Cognitive Bias Approach to Discrimination and Equal Employment Opportunity, 47 Stan. L. Rev. 1161, 1213 (1995) (arguing that the traditional legal account of race discrimination fails to consider psychological circumscription of human cognition); Richard H. McAdams, An Attitudinal Theory of Expressive Law, 79 Or. L. Rev. 339, 342 (2000) ("[I]ndividuals have a preference for esteem; they care what others, even strangers, think of them as an end in itself.").

[23] Professor McAdams's model has been described as "the most comprehensive theory to explain the origin and regulation of norms." Robert E. Scott, The Limits of Behavioral Theories of Law and Social Norms, 86 Va. L. Rev. 1603, 1604 n.2 (2000). The following discussion examines only this theory of norms in depth. A complete taxonomy includes at least two other models. See Robert Cooter, Do Good Laws Make Good Citizens? An Economic Analysis of Internalized Norms, 86 Va. L. Rev. 1577, 1580–81 (2000) (defending an "internalization" model of norms); Eric A. Posner, Law and Social Norms 18 (2000) (defending a conception of norms as a method of signaling discount rates). I consider the implications of my argument for these models below. See infra notes 48–54 and accompanying text.

[24] McAdams, Cooperation and Conflict, supra note 12, at 1019–20.

[25] Id.

[26] Id.

[27] Id. at 1020.

ment imposes a cost in such circumstances and makes conformity with favored customs more attractive.

Esteem can overcome or mitigate social dilemmas. When the commendable behavior corresponds to "cooperation" in the prisoner's dilemma, for instance, esteem resolves the dilemma if it renders cooperation more individually advantageous for each actor than defection.[28] In esteem theory, norms are merely those behavioral regularities which, when followed, generate esteem or disesteem from the relevant peer group.[29] Many norms are efficient in that they resolve collective action problems by making defection comparatively costlier and cooperation relatively cheaper.[30]

In fact, the relationship between esteem and collective action problems in Professor McAdams's model of norms is even more virtuous than initially appears. Esteem-seeking behavior results in particularly high levels of cooperative behavior because competition for esteem is a zero-sum endeavor.[31] Community members have "negative relative preferences" for esteem from their peers.[32] A relative preference is a "function in which one derives pleasure or displeasure from the fact of another's consumption level *in relation to* one's own."[33] Social actors have *relative* preferences for esteem because they value their status level in relation to their peers'

[28] More formally, and with reference to the general conditions for an *n*-person prisoner's dilemma or "tragedy of the commons" discussed above, esteem adds benefits *B* to the "Cooperate" row and adds costs *C* to the "Defect" row. See supra note 9. Esteem can resolve the dilemma if the added benefits and costs resolve the inequality $0 > X$ or $Y > X + Y$. This occurs if $B + C > -X$. See id.; see also McAdams, Cooperation and Conflict, supra note 12, at 1021 n.66, 1029.

[29] See Richard H. McAdams, The Origin, Development, and Regulation of Norms, 96 Mich. L. Rev. 338, 350–53 (1997) [hereinafter McAdams, Regulation of Norms].

[30] As a human motivator, esteem is distinct from reciprocity. Reciprocity—seeking others' approval to serve one's own long-term welfare—can represent a long-run self-interested strategy. See Axelrod, supra note 8, at 124–41; see also James D. Morrow, Game Theory for Political Scientists 264–68 (1994) (discussing iteration's effect on dominant strategies). Social actors seek esteem, by contrast, as an end in itself. See McAdams, Regulation of Norms, supra note 29, at 355 n.77; cf. Robert C. Ellickson, Order Without Law: How Neighbors Settle Disputes 56–57, 130–31 (1991) (describing how norms of reciprocity, enforced by gossip, scorn, ostracism, and even material retaliation, develop among a rural population).

[31] McAdams, Cooperation and Conflict, supra note 12, at 1030.

[32] Id. at 1020.

[33] Richard H. McAdams, Relative Preferences, 102 Yale L.J. 1, 8 (1992) [hereinafter McAdams, Relative Preferences].

positions. These preferences are *negative* because community members want to occupy a higher status level than their peers.[34]

Expenditures to satisfy relative preferences normally represent social waste.[35] If the acquired good suffers from under-production, however, relative preference satisfaction can resolve a market failure by inducing closer to optimal production of that good. Therefore, if individuals have a relative preference for esteem, and if group members provide esteem for cooperative behavior, relative preferences will induce closer-to-optimal production of cooperative behavior. In this manner, esteem's effect on the provision of public goods is socially beneficial, despite the inherent social waste of competing with the peer group to improve one's social status.[36]

For instance, consider a village inhabited by one hundred people and infested with mosquitoes. Each villager would pay fifteen dollars to avoid mosquito bites and their associated diseases. The insects breed in a swamp next to the village and an entrepreneur decides to drain the swamp. He can finance this venture at a cost of $1,000. To do this, he must collect ten dollars from each resident. Each resident reasons that, in the absence of his payment, a sufficient portion of the village will still contribute money, $1,000 will be raised, and the entrepreneur will drain the swamp. As a result, none of the residents contribute, the required $1,000 does not materialize, and the swamp continues to stagnate. Suppose, however, that village members esteem those who contribute to public projects and disesteem those who fail to contribute. If all the villagers have an identical relative preference for esteem, all will contribute

[34] McAdams, Cooperation and Conflict, supra note 12, at 1020. McAdams, Relative Preferences, supra note 33, at 9. Class rank in academic institutions provides a good example. Class rank is a purely relative measurement, yet its existence elicits hard work from students who are motivated, at least partly, by a desire to perform better than their classmates. Students have a negative relative preference for class rank.

[35] See McAdams, Relative Preferences, supra note 33, at 27. Relative preferences are inherently incapable of satisfaction—an equal investment in pursuit of relative status by all members of a group necessarily fails to satisfy any group member's relative preference. Those investments, however, may have satisfied an absolute preference and the foregone opportunity for such satisfaction represents social waste. Id. at 55–56. The Coase Theorem suggests that group members should bargain to an efficient resolution. Transaction costs, however, prevent this resolution—the temptation to cheat and buy the status good while others refrain from doing so is too great. See id. at 64.

[36] See McAdams, Relative Preferences, supra note 33, at 60.

ten dollars, seeking to elevate their relative social status.[37] The entrepreneur will raise enough money to drain the swamp, thus benefiting each resident by fifteen dollars. Each villager receives a net gain of five dollars that he or she would not have gained absent his or her esteem-seeking behavior.

These beneficial effects of norms do not necessarily require that community members have *relative* preferences for esteem. Even if esteem preferences are absolute, if quantities of available esteem are finite, esteem-seeking behavior will still increase production of public goods.[38] Neither do the positive consequences of norms require that community members have *identical* relative preferences for esteem.[39] The esteem theory of norms requires only that a community laud cooperative behavior and censure defection, and that its members have negative preferences for esteem. Under these conditions, social perceptions will increase production of cooperative behavior towards the social optimum. In this manner, esteem ameliorates collective action problems.

III. ESTEEM, GOSSIP, AND FALSE GOSSIP

One important and necessary element of Professor McAdams's model of norms is gossip. By increasing the number of community members aware of a norm violation, gossip amplifies the shameful

[37] The bargaining costs that prevent the Coasean solution, which are pernicious when they cause social waste, are beneficial when they aid resolution of a social dilemma. See id. at 27, 56, 64.

[38] See McAdams, Regulation of Norms, supra note 29, at 357. Returning to the mosquito example, *X* may have an absolute preference for esteem and desire only that she have a "good reputation" for donating to public projects. Nevertheless, if the remainder of her peer group has a finite amount of esteem to allocate, *X* will have to compete with other villagers to secure it. Further, since esteem is, by definition, *X*'s perception of how others regard her, her subjective conception of "good reputation" cannot deviate too much from others' actual perception of her social status. To have a "good reputation" for donation, *X* must have roughly the esteem of an average community member. It would be quixotic and highly idiosyncratic for *X* to believe that she had a "good reputation" for donation when the community actually thought of her as its worst donor.

[39] Relative preferences may conflict and generate optimal production of public goods even when they are not identical. In a community of *A*, *B*, and *C*, for instance, if *A* desires top ranking, and *B* and *C* desire average ranking, their preferences remain intrinsically incompatible. If *A* ranks highest, then either *B* and *C* are tied and neither occupies the average rank, or *B* is higher than *C* (or vice versa) in which case *C* (or *B*) ranks last. See McAdams, Relative Preferences, supra note 33, at 52 n.215.

effect of that violation and thus increases the likelihood of cooperative behavior. False gossip, by contrast, has exactly the opposite effect—it undermines esteem sanctions, weakens efficient social norms, and imposes costs on the community. Section III.A explains the mechanics of gossip and false gossip within the esteem theory of norms.

A. The Negative Externality of False Gossip

An individual's esteem varies directly with the number of community members aware of his norm contravention. For instance, an individual suffers most disesteem when the entire community knows of his norm violation, and an individual suffers no disesteem when his norm violation remains undetected.[40] The esteem theory of norms, therefore, implies that veracious speech, or true "gossip," is socially beneficial.[41] By disseminating news about norm violators, gossiping community members increase the pool of disesteeming peers, which in turn increases the esteem costs of norm violation and reinforces the norm. Similarly, gossip may spread news of norm adherence, increasing such action's esteem benefits and maintaining the norm's vitality. Indirectly, therefore, gossip ameliorates social dilemmas.

Yet gossip need not always convey true information—it may also spread false news of norm violation. Undeniably, such false negative gossip occurs in communities, and it must necessarily originate from some member of the peer group. In fact, the esteem model of norms reveals a community member's potent incentive to instigate false negative gossip about his peers. If esteem preferences are negative, instigating false gossip about one's peers provides a comparatively cheap method of increasing one's own relative esteem. If gossip subsequently spreads news of the false norm violation, the apparent violator receives disesteem. The instigator, who the community perceives not to have violated the norm, now appears relatively more estimable. Generating false gossip may therefore

[40] McAdams, Regulation of Norms, supra note 29, at 365.

[41] See, e.g., id. at 362 ("The conversation we call 'gossip' is often experienced as a benefit, not a cost . . . .").

increase the instigator's esteem relative to the esteem of the gossip's subject.[42]

Admittedly, false negative gossip is futile when community members have a reliable way of identifying it. For instance, in the mosquito example above, the villagers could all decide to make their donations simultaneously in the village square. The estimable behavior of donation would then become publicly observable. In this situation, false gossip loses its value, since group members have a cheap method of detecting its falsity. Estimable behavior, however, is not always public and making it public often requires costly collective action. In these instances, gossip, as a method of spreading norm information, remains a necessary condition for norm effectiveness.

False negative gossip entails certain risks. A strong norm exists against spreading such gossip, and the public revelation of the instigator's identity will result in his social censure. The expected esteem benefits of lying may, nevertheless, often outweigh its expected costs. Indeed, instigators can take inexpensive actions to minimize these costs. For instance, they may conceal their identity as the false rumor's source and avoid social censure even if the community recognizes its falsity.[43]

In such a case, the instigator of a lie can impose a net esteem cost on the lie's victim, which increases the instigator's relative esteem. When a potential esteem gain exists, lying presents a cheaper means of raising one's esteem than the conventional method of

[42] Professor McAdams appears to recognize this pathology when he notes that "[i]f individuals care how they rate in comparison to others, which is my claim, approving others might mean less room for approving oneself. Perhaps one maximizes utility by disapproving everyone but oneself." McAdams, Regulation of Norms, supra note 29, at 357 n.85. This statement implies recognition that negative esteem preferences might lead individuals to withhold gossip they should otherwise supply. It is a short step from that conclusion to the realization that individuals may have an incentive for false negative gossip. Professor McAdams does not, however, explore this implication. Id. But see Richard H. McAdams, Group Norms, Gossip, and Blackmail, 144 U. Pa. L. Rev. 2237, 2245–46 (1996) (arguing that weaker blackmail laws would discourage social propagation of news of norm violations and thereby weaken norm vitality).

[43] This occurs, for example, when the instigator of the lie passes it on to another community member with the disclaimer "*X* told me that . . ." or when the instigator anonymously provides the information to a potential gossiper.

norm adherence.[44] In these cases, preferences for esteem create incentives to instigate false negative gossip.

This gossip has several related pernicious effects on the normative behavior of community members. First, the existence of false negative gossip will distort the provision of disesteem. Community members have no accurate method of distinguishing false negative gossip from true negative gossip. To guard against the possibility of inadvertently disesteeming a peer who has not actually violated a norm, community members will withhold some disesteem that they would otherwise have granted. Withheld disesteem increases as the amount of false negative gossip increases.[45] As a consequence, such gossip decreases the cost of actual norm violation and induces more community members to violate the norm. As more members defect, the esteem benefits of following the norm decrease.[46] What formerly constituted shameful minority behavior now becomes more common and in turn more socially acceptable. False negative information, therefore, has a pernicious feedback effect on norm adherence and exacerbates the collective action problem that the norm originally addressed.

One might predict a spontaneous solution to the deleterious consequences of false negative speech. In an effort to avoid this insidious feedback, community members might agree to maintain high levels of disesteem against the individuals accused of norm violation. But this attempt to salvage the norm has its own associated problems. Understanding the chance of false negative gossip ex ante, a community member will discount the actual esteem gains from following the norm. After all, even if she conforms, the same individual may later face false negative gossip. As the esteem gains are discounted, defection becomes comparatively more attractive and more community members begin to defect. Behavior that violates the norm becomes more commonplace and the norm itself falters.

[44] Of course, if false negative gossip becomes rampant, then nobody donates and *Y* himself suffers from continued infestation—but that just demonstrates the socially pernicious effects of false negative gossip.

[45] In the limiting case, all negative gossip is false and none of it elicits any disesteem.

[46] Cf. McAdams, Regulation of Norms, supra note 29, at 367–69 (discussing the converse phenomenon of "tipping," in which increased levels of norm adherence make violation of the norm more costly).

False negative gossip's effects are essentially irremediable. A community might attempt to discount disesteem to control for falsity, but this tactic entails inadequate shaming of actual norm violation. Yet maintaining high levels of disesteem against those accused of norm violation, in an effort to disesteem real violators, wrongly shames the victims of defamation. Community members are effectively caught in a catch-22.

Of course, false negative speech injures the individual to whom it refers. She suffers frustration and a sense of injustice at the false accusation of misbehavior. The esteem theory of norms, however, further demonstrates that false gossip constitutes a negative externality and imposes costs on the community independent of the individual psychic suffering that the victim experiences. Phony allegations of norm violation threaten the norm whose violation is alleged and thereby preserve the collective action problem that the norm originally addressed.[47]

Individual community members have an incentive to lie about others' norm violations to undermine their esteem competitors' social standings. The whole community, including the instigator, would benefit if all its members could credibly promise not to instigate false negative gossip, but each member is better off instigating lies while the remainder of the peer group refrains. This is a prisoner's dilemma writ large—in other words, a tragedy of the commons. Interestingly, therefore, the esteem theory of norms, which purports to show why social dilemmas do not necessarily require government regulation, ends up entailing a social dilemma that invites government regulation.

Negative preferences for esteem vividly illustrate how incentives for false negative gossip arise, and the esteem-based theory of norms explains the costs of such gossip. The problem of false negative information, however, is not peculiar to Professor McAdams's model. False negative information will constitute a negative externality whenever norm maintenance depends on the transmission of true negative gossip through a community. When a community member values her peers' praises and attaches some costs to their

[47] The aphorism "Oh! what a tangled web we weave/ When first we practise to deceive!" is even more apt than it initially appears. Sir Walter Scott, Marmion: A Tale of Flodden Field, canto VI, stanza 17 (MacMillan 1922) (1808).

condemnations, false negative gossip, whatever its source, distorts those valuations.

The internalization model of norms, for instance, proceeds from the assumption that people have a taste for conforming to norms, which is a willingness to "pay something to obey the norm for its own sake, independent of any resulting advantage or disadvantage."[48] These tastes vary continuously across a community so that most people will pay nothing, a few will pay a little, and even fewer will pay a lot, to obey the norm. Obeying the norm also entails costs, which equal the direct costs, such as "money, inconvenience, effort, risk, or lost opportunity," less "reputational benefit" and "avoided sanction."[49] These costs are also distributed uniformly so that, as net costs increase, fewer people are willing to obey the norm.[50] An equilibrium occurs when the price people are willing to pay equals the cost of adherence. A certain level of norm adherence corresponds to each equilibrium.[51]

False negative gossip creates the same externality for the internalization model of norms as it does for the esteem model. Such gossip causes community members to withhold disesteem judgments that they would otherwise bestow. False negative gossip therefore decreases both the "reputational benefit" and the "avoided sanction" of following the norm. Assuming that "direct costs" remain constant, a reduction in "reputational benefit" and "avoided sanction" causes an increase in the net cost of obeying the norm.[52] This increase shifts the norm's cost curve outwards.[53] The shift makes it more likely that fewer people will be willing to obey the norm at the current cost, which puts upward pressure on that cost. As the cost increases, however, fewer people are willing to pay to obey the norm. False negative information, therefore, decreases the number of norm-abiding community members and represents an externality even for the internalization theory of norms.[54]

[48] See Cooter, supra note 23, at 1583.
[49] Id. at 1584.
[50] Id. at 1585.
[51] See id. at 1586–89.
[52] See id. at 1584.
[53] See id. at 1585.
[54] The other influential model of norms views normative behavior as a method of signaling a low discount rate. See Posner, supra note 23, at 18. Individuals with low

Generally, false negative gossip has an insidious effect for any theory of norms that incorporates social dissemination of reputational information as an enforcement mechanism. Absent regulation, false negative gossip will erode efficient norms and exacerbate collective action problems that these norms would otherwise ameliorate.

Of course, the strong social norm against instigation of false gossip represents some regulation and deters some would-be instigators. Despite this social deterrence mechanism, however, one would still expect significant levels of false negative gossip. Empirical evidence bears out this theoretical prediction: false negative gossip does occur in all sorts of communities, historical and contemporary.[55]

False negative gossip creates a negative externality that community members cannot resolve with a mere norm against engaging in this activity. As with other social dilemmas, legal intervention represents a possible resolution. Legal punishment can alter the payoffs associated with the dilemma, making cooperation the dominant strategy. Section III.B explores the various possible legal regimes of punishment to combat the negative externality of false negative gossip and discusses their attendant problems.

---

discount rates—or "good types"—are attractive partners while individuals with high discount rates—or "bad types"—are dangerous cooperative partners who may defect at any time. Id. Norms, which are merely arbitrarily chosen costly behavioral regularities, enable good types to separate themselves out from bad types and attract other good types with whom to cooperate. Id. at 7–8, 19. Given the emphasis that Posner places on the visibility of normative behavior, it is harder to integrate the preceding discussion of false negative gossip into his theory of norms. After all, if all normative behavior is publicly visible, then false negative gossip is a non-starter. Nevertheless, presumably not all normative behavior need be publicly visible. Further, Posner must admit that false negative gossip does in fact occur. Such gossip, therefore, represents a negative externality for his theory of norms since it makes investments in normative behavior less beneficial to the investor.

[55] See Max Gluckman, Papers in Honor of Melville J. Herskovits: Gossip and Scandal, 4 Current Anthropology 307 (1963) (summarizing anthropological research on gossip and scandal and describing examples of defamation, and its punishment, within various communities); see also Shirley G. Ardener, Sexual Insult and Female Militancy, 8 Man 422, 422–23 (1973) (recounting a type of defamation and the associated communal response among a Cameroonian ethnic group); George M. Foster, The Anatomy of Envy: A Study in Symbolic Behavior, 13 Current Anthropology 165, 172 (1972) (noting that "[i]n peasant societies envy is expressed to third persons by gossip . . . and defamation").

B. Punishment of False Negative Gossip

A vigorous prohibition on the instigation of false negative gossip might represent the best solution to its pernicious effect on norms. Punishing the instigators of false gossip would decrease the amount of false negative gossip and maintain socially beneficial norms. Of course, if the community could always inexpensively identify the instigators of false negative gossip there would be no initial market failure. The community could ignore the gossip, initiate social sanctions against the instigator, and continue to disesteem only real instances of norm violation. But it is difficult to identify the instigators of false negative gossip. The problem, therefore, requires a different punishment regime to avoid the injurious normative consequences.

Penalizing the communicators of false information represents a possible alternative regime.[56] Even if these communicators are not themselves malicious, by supplying a market for false negative gossip they indirectly encourage its instigation. By making transmission of false negative gossip costlier, punishment of its transmission decreases the levels of such gossip and increases the probable truth of all negative statements in social circulation. Actual norm violations would still attract community opprobrium and norm adherents could be more confident that their reputations would not be falsely attacked in the future. Penalizing communicators, therefore, represents a seemingly attractive solution to the problem of false negative gossip.[57]

[56] A similar rationale of punishing comparatively innocent parties to deter related wrongdoers supported the statute at issue in the controversial case of *Bartnicki v. Vopper*, 532 U.S. 514, 520–21 (2001). Commenting on a federal law that criminalized the intentional disclosure of information that the discloser knew or should have known was obtained by illegal eavesdropping, a majority of the Court rejected the rationale. The majority found that "it would be quite remarkable to hold that speech by a law-abiding possessor of information can be suppressed in order to deter conduct by a non-law-abiding third party." 532 U.S. at 529–30. The Chief Justice retorted that "[t]he law against interceptions, which the Court agrees is valid, would be utterly ineffectual without these antidisclosure provisions." Id. at 551 (Rehnquist, C.J., dissenting). The majority opinion prompted some academic criticism, which echoed the Chief Justice. See, e.g., Paul Gewirtz, Privacy and Speech, 2001 Sup. Ct. Rev. 139, 148 (2001) ("But there is nothing remarkable or novel about restricting speech that is the fruit of unlawful action as a means of deterring the unlawful action.").

[57] The Court rejected the "drying up the market" rationale in *Bartnicki*. See supra note 56. But similar rationales underpin the numerous state laws against receiving sto-

Penalizing the communication of false negative gossip, however, raises its own dilemma. The punishment of false speech deters true speech when communicators cannot readily distinguish between the two.[58] Imposing harsh punishment on false speech, therefore, has the indirect and unintended consequence of deterring true speech.[59] Further, if increasing punishment deters both true and false negative gossip in roughly equal measure, then any punishment regime is as good as another. For every unit of injurious false negative speech deterred under a harsher regime, one unit of beneficial true negative speech would also vanish. Even though a harsh rule of punishment for false negative gossip also deters true negative gossip, it is still preferable to a lenient rule. Several reasons support this conclusion.

First, increased punishment may deter more false speech than true speech. While true negative gossip and false negative gossip are not readily distinguishable, they can nevertheless be distinguished. Community members may be able to identify false negative gossip after some investigation. Indeed, this investigation need not be all that costly, especially in smaller communities. If such verification is possible, harsher punishment will reduce false negative gossip more than true negative gossip. Harsher punishment for communication of false negative gossip may then have the beneficial effect of increasing the average quality of negative speech in social circulation with only a minimal decrease in its quantity.

len goods. See, e.g., Ariz. Rev. Stat. Ann. § 13-1802(A)(5) (West 1989); Ohio Rev. Code Ann. § 2913.51 (West 1992).

[58] See, e.g., Richard A. Epstein, Was *New York Times v. Sullivan* Wrong? 53 U. Chi. L. Rev. 782, 798 (1986) ("There is a danger that if the false statements are punished, then the true statements will not be made at all."); Frederick Schauer, Fear, Risk and the First Amendment: Unraveling the "Chilling Effect," 58 B.U. L. Rev. 685, 693 (1978) [hereinafter Schauer, "Chilling Effect"] ("Deterred by the fear of punishment, some individuals refrain from saying or publishing that which they lawfully could, and indeed, should."); J. Hoult Verkerke, Legal Regulation of Employment Reference Practices, 65 U. Chi. L. Rev. 115, 118 (1998) (finding in the market for employment references that "the search for second-best solutions must confront an inevitable trade-off between the quantity and the quality of available information concerning employee productivity").

[59] See Verkerke, supra note 58, at 122; see also Alain Sheer & Asghar Zardkoohi, An Analysis of the Economic Efficiency of the Law of Defamation, 80 Nw. U. L. Rev. 364, 379 (1985) ("The more forbearing liability rule discourages investment in investigation . . . *and* makes publication more likely.").

Second, harsher punishment may have beneficial long-term effects on norm maintenance. Accepting that community members have a proclivity to gossip, a harsh rule of punishment encourages the development of useful proxies that aid the identification of false negative speech. Applying these proxies enables an increase in the amount of true negative gossip without a concomitant increase in the amount of false negative gossip. A harsh rule of punishment may have the long-term consequence, therefore, of undoing the "chilling effect" that it initially imposes on true speech.[60]

Third, the aggregate individual harm to victims of false negative speech decreases as the harshness of punishment increases. This harm represents a cost quite apart from the incentive effects defamation has on community members' behavior. Thus, even if a harsh rule of punishment chills true negative speech, by also chilling false negative speech it minimizes the total associated psychic harms of defamation.

Lastly, and most importantly, under any theory of norms that recognizes social incentives to instigate false negative gossip, drying up the market for such gossip can have beneficial deterrent effects on *would-be instigators*. The esteem theory of norms predicts incentives to instigate false negative gossip: the liar hopes to depress the reputations of his esteem competitors. A regime of stricter punishment narrows the channels of transmission by decreasing the total amount of negative speech in social circulation. Assuming a fixed preference for esteem, the instigator must now adjust his behavior in order to satisfy his preference. Since lying is no longer an effective method of satisfying this preference, the instigator must now actually obey the norm to gain esteem. Drying up the market for false negative gossip, therefore, has the beneficial effect of transforming instigators of false gossip into norm followers.[61] Harsh punishment of false nega-

[60] Tort law scholars make this claim with respect to strict liability generally. See, e.g., Kenneth S. Abraham, The Forms and Functions of Tort Law 166 (1997) ("[S]trict liability creates additional research incentives, because under strict liability there is more to be gained by avoiding liability.").

[61] Even if strict punishment reduces true negative gossip to a trickle and the costs of norm violation fall dramatically, this virtuous substitution from falsehood instigator to norm follower persists. No incentives exist to propagate false news of norm adherence, or false positive gossip, which remains a dependable amplifier of esteem-following normative behavior. The esteem-seeking society member cannot, therefore, pursue his relative preference for esteem by instigating self-referential false positive

tive gossip effectively enforces the efficient Coasean bargain by surmounting the transaction costs obstructing an agreement to refrain from such gossip.

For any regime of norms in which gossip plays an enforcement role, a strict rule of punishment for the transmission of false speech appears most preferable. As punishment for false negative speech transmission becomes more lenient, therefore, the net cost in terms of norm erosion increases *even if* the more lenient regime "unchills" some true negative speech that would have remained unsaid in a harsher regime. Increased levels of false negative speech impose increased social costs regardless of whether true negative speech also increases.

If only normative concerns underlay the punishment for dissemination of false negative gossip, a harsh regime would appear socially optimal. Yet goals besides norm protection inform the actual legal regime addressing the problem of false defamatory speech. Indeed, the most important developments in defamation law in the twentieth century occurred after the Court acknowledged these competing goals under the aegis of the Constitution.

Perhaps reflecting the normative costs of negative false gossip, defamation was a strict liability tort at common law. The Court radically refashioned the common law rules, however, when it perceived these rules to restrict excessively certain kinds of political speech protected by the First Amendment. Part IV discusses the old common law regime, its constitutional replacement, and the assumptions that informed the transition.

IV. THE LAW OF DEFAMATION AND ITS CONSTITUTIONALIZATION

Prior to 1964, the Court "had gone for 170 years without finding in the first amendment any limits on libel judgments."[62] During this time, the common law of defamation had evolved into a complex body of doctrine in U.S. jurisdictions. Section IV.A traces the outline of this doctrine.

gossip. To receive the benefits of a good public name, he must actually engage in normative behavior. In short, a community can afford to largely abandon negative gossip as a means of disseminating normative information.

[62] Anthony Lewis, *New York Times v. Sullivan* Reconsidered: Time to Return to "The Central Meaning of the First Amendment," 83 Colum. L. Rev. 603, 604 (1983).

A. Defamation at Common Law

At common law, the tort of defamation required the dissemination by the defendant to another person of a false defamatory communication referring to the plaintiff.[63] A negligence standard governed the element of dissemination to a third party.[64] The remaining three components—the statement's falsity, its defamatory character, and its reference to the plaintiff—were all strict liability elements.[65] A defamatory statement was one that tended "to harm the reputation of another as to lower him in the estimation of the community or to deter third persons from associating or dealing with him."[66] The common law presumed the falsity of a defamatory statement and the defendant bore the burden of pleading and proving truthfulness.[67] Under the regime of presumed or general damages, the plaintiff could recover damages for "harm to reputation that would normally be assumed to flow from a defamatory publication of the nature involved."[68] A plaintiff eligible for presumed damages required no proof of actual injury to recover.[69] Despite the strictures of defamation liability at common law, a limited set of privileges immunized certain defendants from damages for defamation.[70] Two types of privilege existed—absolute and qualified.

[63] See Restatement (Second) of Torts § 580B cmt. b (1977); see also M. Linda Dragas, Curing a Bad Reputation: Reforming Defamation Law, 17 U. Haw. L. Rev. 113, 125–27 (1995) (discussing the tort of defamation at common law). At common law, a distinction existed between oral defamation, or slander, and defamation in a written or otherwise more permanent form, or libel. The common law of defamation made libel substantively easier to prove than slander. See Abraham, supra note 60, at 252.

[64] Restatement (Second) of Torts § 580B cmt. b (1977).

[65] Id.

[66] Id. § 559; see also Zinda v. Louisiana Pac. Corp., 149 Wis. 2d 913, 921 (Wis. 1989) (providing a similar definition of "defamatory").

[67] W. Page Keeton et al., Prosser and Keeton on the Law of Torts 839 (5th ed. 1984) [hereinafter Prosser & Keeton].

[68] Restatement (Second) of Torts § 621 cmt. a (1977). Some jurisdictions at common law permitted recovery of presumed damages in all libel actions and a certain subset of slander actions. See Abraham, supra note 60, at 252. For the remaining slander actions, damages required particular proof. Id.

[69] Restatement (Second) of Torts § 621 cmt. a (1977).

[70] See Rodney A. Smolla, Law of Defamation § 8 (1986).

Absolute privileges completely absolved the defendant of liability, without regard to his motive in uttering the falsehood.[71] These privileges attached only to statements made during judicial and legislative proceedings, executive communications, interspousal dialogue, political broadcasts, and statements made with the consent of the plaintiff.[72]

Qualified or conditional privileges, by contrast, did not defeat liability in all instances and could be overcome by a showing of abuse.[73] Abuse occurred when the defendant had no reasonable grounds for believing the truth of the statements asserted.[74] A defendant who had made statements to protect his own legitimate interest could avail himself of a conditional privilege.[75] Such a privilege also attached to statements made to protect the legitimate interests of a third party, particularly when the speaker had a moral or legal duty to protect that party's interest.[76] Employment references also enjoyed a conditional privilege,[77] as did statements made by speakers who shared common interests with their audience.[78]

Lastly, the common law recognized a "fair comment" qualified privilege. Born of the realization that "'[f]air discussion is essentially necessary to the truth of history, and the advancement of science,'"[79] this privilege immunized the publication of good faith, but

[71] Id. § 8.01[2]; see also Abraham, supra note 60, at 253 (describing absolute privileges at common law).

[72] See Prosser & Keeton, supra note 67, at 816–24; see also Smolla, supra note 70, §§ 8.02–.06 (describing situations where absolute privilege attached at common law).

[73] See Smolla, supra note 70, § 8.07[2].

[74] See, e.g., Zuschek v. Whitmoyer Labs., 430 F. Supp. 1163, 1166 (E.D. Pa. 1977) (holding that the conditional privilege lost when defendant had no reasonable grounds for believing truth of matter asserted), aff'd mem., 571 F.2d 573 (3d Cir. 1978). Other showings also defeated the conditional privilege—proof of "excessive publication," proof of publication in spite of an absence of belief in the necessity for publication in light of the reason for the privilege, proof of common law "ill-will malice," or proof of "actual malice." Smolla, supra note 70, § 8.07[2].

[75] Smolla, supra note 70, § 8.08[1]. Instances of this conditional privilege include: statements made to defend one's reputation in response to attack by another, statements made while retrieving stolen property, and statements made in the course of collecting a bona fide debt. Id.

[76] Id. § 8.08[2][a]. Examples of this conditional privilege include: a doctor's statement to protect a patient and an attorney's statement on behalf of a client. Id.

[77] Id. § 8.08[2][d]; see also Verkerke, supra note 58, at 160–61.

[78] Smolla, supra note 70, § 8.08[3][a].

[79] Id. § 6.02[1] (quoting Tabart v. Tipper, 170 Eng. Rep. 981, 982 (1808)).

defamatory, statements about matters of public importance.[80] Although the original fair comment privilege protected only statements of pure opinion from liability,[81] a large minority of states eventually extended the privilege to false statements of fact.[82] For states adhering to the more restrictive regime, some courts originally required that the underlying facts reasonably imply the stated opinion.[83] Nevertheless, even in these jurisdictions the privilege gradually came to protect any opinion based on true facts, regardless of its reasonableness.[84] The requirement that opinions within the privilege comment on matters of public importance assumed an expansive interpretation at common law. Courts deemed such matters to include "a great variety of subjects . . . matters of public concern, public men, and candidates for office."[85]

At common law, therefore, defamation was essentially a strict liability tort, with small pockets of negligence, and even smaller pockets of absolute immunity.[86] Observing this legal landscape Jus-

[80] Robert C. Post, The Constitutional Concept of Public Discourse: Outrageous Opinion, Democratic Deliberation, and *Hustler Magazine v. Falwell*, 103 Harv. L. Rev. 601, 627 (1990); see also Smolla, supra note 70, § 6.02 (exploring the occasions for fair comment at common law); David A. Logan, Tort Law and the Central Meaning of the First Amendment, 51 U. Pitt. L. Rev. 493, 502–03 (1990) (same).

[81] See Rodney A. Smolla, Let the Author Beware: The Rejuvenation of the American Law of Libel, 132 U. Pa. L. Rev. 1, 67 (1983); see also Susanna Frederick Fischer, Rethinking *Sullivan*: New Approaches in Australia, New Zealand, and England, 34 Geo. Wash. Int'l L. Rev. 101, 116 n.85 (2002). Of course the line between opinion and fact is blurry and common law courts had to make fine distinctions to police this line. See Logan, supra note 80, at 503. The common law rule that fair comments could not imply the existence of undisclosed defamatory facts compounded this blurriness. See David A. Anderson, Is Libel Law Worth Reforming? 140 U. Pa. L. Rev. 487, 505 (1991).

[82] Sheer & Zardkoohi, supra note 59, at 371 n.19.

[83] Logan, supra note 80, at 503–04.

[84] Smolla, supra note 70, § 6.02[1].

[85] See Coleman v. MacLennan, 98 P. 281, 285 (Kan. 1908); see also Smolla, supra note 70, § 6.02[2] (quoting Mashburn v. Collin, 355 So. 2d 879, 882 (La. 1977)) (confirming the broadness of fair comment at common law).

[86] Commentators thus label the common law defamation regime "strict liability." See, e.g., Smolla, supra note 70, § 1.02[2] ("[T]he law of defamation . . . was essentially a strict liability tort with most rules stacked in the plaintiff's favor."); Fischer, supra note 81, at 114 ("[L]iability is strict. . . . [T]his common law privilege was relatively narrow."); Robert E. Keeton, Federal Influences on the Treatment of Law and Fact in Tort Litigation, 55 Md. L. Rev. 1344, 1361 (1996) ("'At common law the defamation actions, libel and slander, were strict liability torts. . . . Absent truth or privilege, the actor was subject to liability without regard to his fault or lack of fault.'" (quoting

tice Holmes quipped: "'Whatever a man publishes he publishes at his peril.'"[87] As a tort, defamation remained squarely within state jurisdiction. Although the First Amendment facially implicated any state laws touching on speech, prior to the 1960s jurists and legal commentators generally placed regulation of defamatory falsehoods outside the Constitution's ambit. "[L]ibelous words," Justice White later noted, were "wholly unprotected by the First Amendment."[88] All this changed, however, when the Court decided *New York Times Co. v. Sullivan*.[89]

B. The Constitutional Law of Defamation

The plaintiff in *Sullivan*, an Alabama police chief, claimed that an advertisement in the *New York Times* defamed him.[90] The advertisement provided information and opinion about the civil rights movement, described student protest at Alabama State College, alleged official abuses, and requested financial support.[91] The text contained some inaccuracies—Dr. Martin Luther King had been arrested four times by Alabama officials, not seven times as the advertisement claimed; the protesting students had sung the National Anthem, not "My Country 'Tis of Thee"; and the student dining hall had not been padlocked "in order to starve [the protestors] into submission."[92]

The trial court charged the jury that, under Alabama law, the statements were libelous per se,[93] that no fair-comment privilege existed because the errors were factual, and that the jury should find the defendant liable if the statements were "of and concerning" the plaintiff.[94] The jury found that the statements were indeed

Fowler V. Harper et al., The Law of Torts § 5.0, at 3–5 (2d ed. 1986 & Supp. I 1995) (emphasis removed))).

[87] Peck v. Tribune Co., 214 U.S. 185, 189 (1909) (quoting The King v. Woodfall, 98 Lofft, 776, 781 (K.B. 1774) (Lord Mansfield)).

[88] Gertz v. Robert Welch, Inc., 418 U.S. 323, 370 (1974) (White, J., dissenting).

[89] 376 U.S. 254 (1964).

[90] Id. at 256.

[91] Id. at 256–59.

[92] Id. at 257–59.

[93] At common law a statement was libelous per se if the defamatory meaning was clear from the face of the words. Smolla, supra note 70, § 1.04[5]. A libel per se implied injury from the bare fact of publication. *Sullivan*, 376 U.S. at 262.

[94] *Sullivan*, 376 U.S. at 262.

"of and concerning" Sullivan, even though they did not mention him by name, because they referred to the Alabama police, which he commanded.[95] The jury awarded $500,000 under Alabama's presumed and punitive damages regime.[96] Affirming the judgment, the Alabama Supreme Court rejected a First Amendment defense.[97]

The Supreme Court of the United States reversed, holding that, because the advertisement addressed a major public issue, it qualified for constitutional protection under the First and Fourteenth Amendments.[98] According to the Court, "the background of a profound national commitment to the principle that debate on public issues should be uninhibited, robust, and wide-open"[99] required a constitutional rule limiting tort recovery in defamation suits by public officials. A public official could recover for a defamatory falsehood regarding his official conduct only if he proved, by clear and convincing evidence,[100] that the "statement was made with 'actual malice'—that is, with knowledge that it was false or with reckless disregard of whether it was false or not."[101] Since the record supported at most a finding of negligence against the *Times*, the newspaper possessed a constitutional right to avoid censure for publishing the advertisement.[102]

The petitioner's counsel urged, and the Court accepted, an analogy of defamation of public officials to seditious libel.[103] The Sedition Act of 1798 criminalized defamation of the President, Congress, or the federal government as a whole.[104] Just as "the court of history" rejected the Act because of "the restraint it imposed upon criticism of government and public officials," so too must strict liability for defamation of public officials' governmental actions fail, explained the Court.[105] "The right of free public discussion of the

[95] Id. at 288.
[96] Id. at 256.
[97] See N.Y. Times Co. v. Sullivan, 144 So. 2d 25, 40 (Ala.1962).
[98] *Sullivan*, 376 U.S. at 271.
[99] Id. at 270.
[100] Id. at 285–86.
[101] Id. at 279–80.
[102] Id. at 288.
[103] Lewis, supra note 62, at 606.
[104] Id.
[105] *Sullivan*, 376 U.S. at 276.

stewardship of public officials," the Court stated, "was thus . . . a fundamental principle of the American form of government."[106]

Since *Sullivan* involved a "classic example of an activity . . . of 'governing importance,'"[107] the Court did not have to delimit precisely the constitutional privilege it created. The Court did not decide "how far down into the lower ranks of government employees the 'public official' designation would extend" nor "the boundaries of the 'official conduct' concept."[108] Later cases inevitably posed these questions.

The Court first extended the constitutional privilege to defamatory statements about public figures in general.[109] This extension rested upon the conclusion that, since the guarantees of free speech "'are not the preserve of political expression or comment upon public affairs' . . . freedom of discussion 'must embrace all issues about which information is needed . . . to enable the members of society to cope with the exigencies of their period.'"[110] Finding no "rational distinction" between criticism of public officials and criticism of private citizens who "seek to lead in the determination of . . . policy,"[111] the Court granted constitutional protection to speech that is defamatory of such "public figures."

The Court defined "public figure" to include any private citizen who had voluntarily involved himself in an important public controversy,[112] or who retained power over such controversies by virtue of his social position.[113] Access to means of public rebuttal, which

[106] Id. at 275.

[107] William J. Brennan, Jr., The Supreme Court and the Meiklejohn Interpretation of the First Amendment, 79 Harv. L. Rev. 1, 14 (1965).

[108] *Sullivan*, 376 U.S. at 283 n.23.

[109] See Curtis Publ'g Co. v. Butts, 388 U.S. 130 (1967) (the companion case to *Walker v. Associated Press*).

[110] Id. at 147 (quoting Time, Inc. v. Hill, 385 U.S. 374, 388 (1967); Thornhill v. Alabama, 310 U.S. 88, 102 (1940)).

[111] Id. at 147–48 (quoting Pauling v. Globe-Democrat Publ'g Co., 362 F.2d 188, 196 (8th Cir. 1966)). Chief Justice Warren justified the expansion by noting that public officials no longer exclusively determined policy in modern society. Id. at 163–64 (Warren, C.J., concurring).

[112] See id. at 155. Walker, for instance, was a retired general who allegedly led a violent anti-integration rally in Mississippi. Id. at 140, 159 n.22.

[113] See id. at 155. Butts, for instance, was a nationally known football coach at the University of Georgia. Id. at 135–36.

the Court ascribed to public figures, partially justified the extension of the actual malice rule to their defamation.[114]

The Court continued its development of the constitutional privilege in *Rosenbloom v. Metromedia, Inc.*[115] A plurality of justices abandoned the public figure test and extended the constitutional privilege even further to "all discussion . . . involving matters of public or general concern."[116] Justice Brennan also repudiated the concepts of voluntary exposure and self-defense that had informed the notion of public figure in *Curtis Publishing Co. v. Butts*. He asserted that "[e]xposure of the self to others in varying degrees is a concomitant of life in a civilized community"[117] and stated that access to means of rebuttal depended upon "the unpredictable event of the media's continuing interest in the story."[118]

The Court reined in the expansion of the *Sullivan* rule in *Gertz v. Robert Welch, Inc.* and once more reserved to states some discretion in fashioning their defamation laws.[119] The plurality found that the *Rosenbloom* rule gave insufficient weight to the legitimate state interest in protecting private citizens' reputations.[120] For private persons, states could "define for themselves the appropriate standard of liability for a publisher or broadcaster of defamatory falsehood" when "the substance of the defamatory statement 'makes substantial danger to reputation apparent.'"[121] Without proof of ac-

[114] See id. at 155; id. at 164 (Warren, C.J., concurring).

[115] 403 U.S. 29 (1971) (plurality opinion).

[116] Id. at 43–44 (plurality opinion). Justice Brennan's plurality opinion failed to define "matters of public or general concern" but held that arrest of a private citizen for alleged distribution of obscene materials clearly qualified. Id. at 44–45. Commentators bemoaned that, after *Rosenbloom*, the mere fact of media publication became proof of public interest. See, e.g., Joel D. Eaton, The American Law of Defamation Through *Gertz v. Robert Welch, Inc.* and Beyond: An Analytical Primer, 61 Va. L. Rev. 1349, 1398 (1975) ("[N]early all cases found news media reports . . . to be of such concern.").

[117] *Rosenbloom*, 403 U.S. at 47 (quoting Time, Inc. v. Hill, 385 U.S. 374, 388 (1967)).

[118] Id. at 46.

[119] 418 U.S. 323 (1974). The six separate opinions in this case reveal just how contentious an area of law defamation had become.

[120] *Gertz*, 418 U.S. at 346 (plurality opinion). The Court found private plaintiffs more vulnerable to defamation than public figures because the former lacked access to means of rebuttal and deserved greater protection because they had not voluntarily assumed the risk of public scrutiny of their affairs. Id. at 344.

[121] Id. at 347–48 (quoting Curtis Publ'g Co. v. Butts, 388 U.S. 130, 155 (1967)). Strict liability remained unconstitutional, however. Id.

tual malice, however, recovery could not exceed the plaintiff's actual damages.[122] With regard to public figures, the *Gertz* plurality reaffirmed the extension of the actual malice rule in *Butts* and *Walker* and clarified that presumed and punitive damages remained open to successful plaintiffs.[123]

Since *Gertz* made the substantive elements of the tort turn upon the plaintiff's renown, the precise meaning of the public figure test announced in that case assumed great importance. The Court recognized three groups of public persons: all-purpose, limited-purpose, and involuntary.[124] Both all-purpose and limited-purpose public figures "have assumed roles of especial prominence in the affairs of society."[125] Limited-purpose public figures, by distinction, "have thrust themselves to the forefront of particular public controversies in order to influence the resolution of the issues involved."[126] Such figures "engage the public's attention in an attempt to influence [a public issue]."[127] The limited-purpose public figure test, therefore, appeared to echo the voluntary-involvement and self-protection ideas that had informed *Butts* and *Walker*.[128]

Gertz's status-based taxonomy of defamation law represented an equilibrium of sorts. Subsequent cases expounded the category of limited-purpose public figure[129] and clarified constitutional constraints on procedure in defamation cases.[130] There have also been alterations to the substantive constitutional law of defamation,

[122] Id. at 349.

[123] Id. at 342, 349–50.

[124] Id. at 344–46. The Court cryptically and unhelpfully asserted that "it may be possible for someone to become a public figure through no purposeful action of his own, but the instances of truly involuntary public figures must be exceedingly rare." Id. at 345.

[125] Id. at 345.

[126] Id.

[127] Id. at 352.

[128] See supra text accompanying notes 109–11.

[129] Various cases clarified the definition of public figure. See Wolston v. Reader's Digest Ass'n, 443 U.S. 157 (1979); Hutchinson v. Proxmire, 443 U.S. 111 (1979); Time, Inc. v. Firestone, 424 U.S. 448 (1976).

[130] See Bose Corp. v. Consumers Union of United States, 466 U.S. 485, 511 (1984) (holding that appellate court must independently review record to determine whether public figure plaintiff in fact demonstrated actual malice by clear and convincing evidence); Herbert v. Lando, 441 U.S. 153, 175–76 (1979) (holding that public figures are entitled to pre-trial discovery of editorial process on the issue of actual malice).

some of which were significant.[131] For the purposes of this Note, however, the broad contours of the constitutional law of defamation, outlined in the preceding discussion of *Sullivan*, *Butts*, and *Gertz*, and the policy rationales propounded in those cases, suffice to frame critical comment.

In order to facilitate such comment, Part V gleans from the aforementioned opinions the various justifications for encouraging speech defamatory of public figures or defamatory speech implicating matters of public concern. Part V then considers how these justifications square with theories of norms that rest upon social dissemination of true negative gossip.

V. THE CONSTITUTIONAL LAW OF DEFAMATION AND NORMS

The revolution in constitutional defamation jurisprudence plainly rested upon certain descriptive premises regarding the value and prevalence of particular categories of speech. Section V.A identifies and explains these premises and their connection with the principles expounded in *Sullivan* and its progeny.

A. The Problem of Under-Produced Political Speech

The *Sullivan* Court recognized that some true defamatory speech creates significant positive externalities. Justice Brennan stated that "'[t]he maintenance of the opportunity for free political discussion to the end that government may be responsive to the will of the people and that changes may be obtained by lawful means . . . is a fundamental principle of our constitutional system.'"[132] Citing a state supreme court precedent, the Court emphasized:

[131] The most important revision occurred in *Dun & Bradstreet, Inc. v. Greenmoss Builders*, 472 U.S. 749 (1985). In this case, the Court reintroduced a subject matter criterion for ascertaining the level of constitutional protection attaching to defamatory speech. Id. at 759–61. A majority held that, in a case involving speech not of public concern, the First Amendment did not require a showing of actual malice for an award of general or punitive damages. Id. at 761. Combining *Sullivan*, *Gertz*, and *Dun & Bradstreet* is no easy task. Justice O'Connor provided a useful reconciliation in her majority opinion in *Philadelphia Newspapers, Inc. v. Hepps*, 475 U.S. 767, 771–75 (1986).

[132] *Sullivan*, 376 U.S. at 269 (quoting Stromberg v. California, 283 U.S. 359, 369 (1931)).

> "It is of the utmost consequence that the people should discuss the character and qualifications of candidates for their suffrages. The importance to the state and to society of such discussions is so vast, and the advantages derived are so great, that they more than counterbalance the inconvenience of private persons whose conduct may be involved . . . ."[133]

Later cases echoed this theme. Justice Harlan observed that "'criticism of private citizens who seek to lead in the determination of . . . policy'" is no less important to the public interest than criticism of public officials.[134] Similarly, the *Rosenbloom* Court noted that "'[f]reedom of discussion, if it would fulfill its historic function in this nation, must embrace all issues about which information is needed or appropriate to enable the members of society to cope with the exigencies of their period.'"[135] Collectively, these comments demonstrate the Court's realization that certain instances of public speech benefit not merely the speaker and his immediate audience, but the community at large. The analogy to seditious libel in *Sullivan* confirms this characterization—the earlier Court had found seditious libel similarly conducive to public welfare.

Ignoring for the moment the controversy over just which types of speech implicate this external benefit,[136] the Court's recognition of this characteristic of the speech at issue admits of more formal expression. Standard microeconomic theory predicts that levels of production will reflect the point at which marginal private benefit equals marginal private cost. For a typical publisher or speaker, this elementary equation determines levels of speech—a community member, for instance, will derive certain private benefits from speaking and will do so until he can more gainfully use his time otherwise. Some speech, however, benefits other community members by providing them with information that informs their

[133] Id. at 281 (quoting Coleman v. MacLennan, 78 Kan. 711, 724 (1908)).

[134] Curtis Publ'g Co. v. Butts, 388 U.S. 130, 147–48 (1967) (quoting Pauling v. Globe-Democrat Publ'g, 362 F.2d 188, 196 (8th Cir. 1966)).

[135] Rosenbloom v. Metromedia, Inc., 403 U.S. 29, 41 (1971) (quoting Thornhill v. Alabama, 310 U.S. 88, 102 (1940)).

[136] The Court's extension of the actual malice rule in *Sullivan*, *Butts*, and *Rosenbloom*, its retreat in *Gertz*, and its puzzling complication in *Dun & Bradstreet* exemplify this controversy. See supra Section IV.B.

governmental decisions and that is otherwise difficult to obtain.[137] Well-informed governance decisions create associated democratic gains that benefit the whole society, not merely the original supplier of the information and his auditors. Such decisions ensure that those controlling public resources are adequately disciplined by their constituents and guarantee the long-run maximization of public welfare.

The individual in possession of information informing these decisions cannot convert the external public benefit associated with this information into personal revenue. His expected private gain, which constrains his decision whether to speak, is less than the expected benefit of this speech to society. Since he cannot capture the positive externality of the self-government speech, the citizen will refrain from speaking even when the community as a whole would prefer that he speak.

In such circumstances, a legal regime that subsidizes speech benefits society at large. By prompting an increase in the level of self-government speech, a legal subsidy causes a shift towards the socially optimal amount. Provision of just this subsidy motivated the Court to replace the strict liability common law regime with an actual malice rule in *Sullivan*.[138] The net decrease in expected liability occasioned by this shift provides incentives for private speakers to utter more speech critical of public figures' governance actions. The actual malice rule apparently elicits closer to the optimal amount of speech defaming public figures.

Despite the facial plausibility of the public subsidy argument, the Court's justification for the actual malice rule begs an important objection. Although true defamatory speech about public officials or matters of public concern provides a self-government benefit, false defamatory speech on such topics imposes a corresponding social harm. If the true information in such speech contributes to democratic governance, false information equally causes democ-

[137] Decisions of this sort include whom to vote for, what social and political causes to support, what stance to take on international affairs, and what reforms of the governance system to support. Speech informing such decisions is hereinafter termed "self-government speech."

[138] See infra note 140.

ratic distortion.[139] A harsh regime of punishment for false defamatory statements, such as strict liability at common law, does deter some true defamatory statements, as the *Sullivan* Court recognized:

> A rule compelling the critic of official conduct to guarantee the truth of all his factual assertions—and to do so on pain of libel judgments virtually unlimited in amount—leads to . . . "self-censorship." . . . Under such a rule, would-be critics of official conduct may be deterred from voicing their criticism, even though it is believed to be true and even though it is in fact true, because of doubt whether it can be proved in court or fear of the expense of having to do so.[140]

Yet a lenient regime, such as the actual malice standard, does just the opposite: it induces "would-be critics of official conduct" to utter their criticism, even though it may well be false ex ante and even though it is in fact false ex post. Lenient regimes of punishment "unchill" both true and false defamatory statements.

It is clear, therefore, that for all the Court's high-minded pronouncements about "'unfettered interchange of ideas'"[141] and "'breathing space'" for freedom of expression,[142] *Sullivan* rests upon the implicit claim that the benefits of increased defamatory truths outweigh the harm of increased defamatory falsehoods un-

[139] This point has not been lost on the Court. See, e.g., Dun & Bradstreet, Inc. v. Greenmoss Builders, 472 U.S. 749, 769 (1985) (White, J., concurring) (noting that the actual malice rule entails that "the stream of information about public officials and public affairs is polluted and often remains polluted by false information"). Commentators have also noted this phenomenon. See, e.g., Epstein, supra note 58, at 814 ("If it is important for the public to know that Jones has been a faithless public official, it is equally important for the public to know that Jones has been a diligent public official falsely accused . . . ."); Sheer & Zardkoohi, supra note 59, at 379 n.47 ("It is also possible that a published false statement may mislead society, and thereby create external societal costs.").

[140] *Sullivan*, 376 U.S. at 279. This dynamic describes the "chilling effect" which punishment of false defamatory speech has on true defamatory speech. The chilling effect reflects the notion that "robust free speech systems protect speech not because it is harmless, but despite the harm it may cause." Frederick Schauer, Uncoupling Free Speech, 92 Colum. L. Rev. 1321, 1321 (1992) [hereinafter Schauer, Uncoupling]. For an exhaustive exploration of the chilling effect in First Amendment law see Schauer, "Chilling Effect," supra note 58.

[141] *Sullivan*, 376 U.S. at 269 (quoting Roth v. United States, 354 U.S. 476, 484 (1957).

[142] Id. at 271–72 (quoting NAACP v. Button, 371 U.S. 415, 433 (1963)).

der an actual malice rule.[143] Only on this assumption can the actual malice rule generate a net social gain.

Of course, the Court does not adopt this proposition by fiat. Several mechanisms that the Court identifies support the claim that true defamation's advantages exceed false defamation's costs. Chief among these is the idea that public figures have access to wide-ranging means of rebuttal when defamers falsely accuse them of shameful behavior.[144] The possibility of public rebuttal allows the community to capture the benefits of increased true defamation while avoiding the attendant costs of increased false defamation in the following manner. The actual malice rule increases both types of defamation. Falsely defamed public figures who have access to a public forum for their self-defense will successfully exculpate themselves. Actual offenders, by contrast, will tacitly admit their guilt by failing to mount an exculpatory campaign or conducting an unsuccessful one. The community will benefit from the knowledge of exculpated individuals' virtues and the inculpated individuals' vices. In this manner, the self-help remedy of public rebuttal supports the claim that increased true defamation's benefits outweigh increased false defamation's costs.[145]

To be sure, there are objections to the reasoning behind the self-help remedy.[146] As such, the notion that the actual malice rule could secure the benefits of increased true defamation while avoiding the costs of increased false defamation remains dubious. Yet the Court might have defended this notion on other grounds. The absence of

[143] *Sullivan* "recognizes only the error costs that run in one direction: those which lead to the reduction in the quantity of speech." Epstein, supra note 58, at 798.

[144] The Court describes and acclaims this self-help remedy on several instances. See *Dun & Bradstreet*, 472 U.S. at 756; Gertz v. Robert Welch, Inc., 418 U.S. 323, 344 (1974); Curtis Publ'g Co. v. Butts, 388 U.S. 130, 155, 164 (1967). Some commentators have endorsed the Court's reasoning, see Sheer & Zardkoohi, supra note 59, at 379 n.47, while others have questioned it, see Anderson, supra note 81, at 526. The Court itself has repudiated the reasoning on at least one occasion. Rosenbloom v. Metromedia, Inc., 403 U.S. 29, 46–47 (1971).

[145] See *Gertz*, 418 U.S. at 344.

[146] One obvious line of criticism proceeds from the recognition that even guilty public figures will invariably respond to negative publicity with public protestations of innocence. If the public cannot easily distinguish false protestations from true protestations, the claim that true defamation's positive externality outweighs false defamation's negative externality weakens. Cf. id. at 344 n.9 (admitting that "the truth rarely catches up with a lie").

a substantiality requirement from the element of falsity, for instance, meant that common law defendants could face large judgments for statements that undeniably conveyed the correct overall impression of the defendant's malfeasance.

In any event, the Court did not press the argument that actual malice could secure self-government gains at no cost to its logical conclusion. After all, if the benefits of increased true defamation outweighed the costs of the concomitant increase in false defamation, surely the Court should have awarded an absolute privilege for all defamation of public figures.[147] Why did the *Sullivan* Court halt the erosion of strict common law defamation at actual malice? Given the comparatively small costs of false negative statements corrected by public rebuttal, and given the comparatively large benefits of the "unchilled" true defamations in a more lenient regime, an absolute immunity for defamation of public figures would surely have maximized the public good.

The Court declined such an extension of constitutional protection precisely because it recognized the competing interests at stake in defamation law. *Sullivan* "stops short of that extreme position" because "[a] world without any protection against defamation is a world with . . . too much misinformation."[148] The underproduction of political information is not the only dynamic underlying laws against false negative speech. Indeed, the Court explicitly credited a state interest in preserving its citizens' reputations on multiple occasions.[149] The recognition of this interest at least reveals that the Court's zeal for self-government speech did not blind it to possible social costs that a regime subsidizing such speech might cause.

Nonetheless, the understanding of this countervailing interest, both in the legal opinions and the academic commentary, reveals a conception of reputation unreflective of its theoretical richness as informed by theories of norms. The Court's understanding completely fails to consider reputation as a social device for regulating

[147] Justice Black, joined by Justice Douglas, actually urged such a privilege in *Sullivan*, 376 U.S. at 293 (Black, J., concurring), and in *Butts*, 388 U.S. at 170–72 (Black, J., concurring).

[148] Epstein, supra note 58, at 798–99.

[149] See, e.g., *Butts*, 388 U.S. at 147 (recognizing the "competing consideration" of reputational protection); Rosenblatt v. Baer, 383 U.S. 75, 86 (1966) (same).

normative behavior. As a consequence, the Court may well have underestimated the benefits of a regime that accords this behavior maximal protection by minimizing the possibility of false negative gossip. In other words, the Court may have misjudged the benefits of a strong law of defamation. Conversely, the Court may have undervalued the social costs of reducing the punishment for false defamation, notwithstanding any attendant self-government benefits.

Section V.B briefly recapitulates theories of norms that recognize social diffusion of true reputational information as an incentive for normative behavior. This Section then describes, by way of contrast, the Court's exclusive focus on reputation as a purely individual asset. This contrast illuminates the Court's underestimation of the costs of increased defamatory speech under the actual malice rule. As a result, the rule gave too much to self-government speech and retained too little for reputation and norms.

B. The Actual Malice Rule and Normative Behavior

Part IV noted that false negative gossip, or defamation, has a negative feedback effect on norm adherence. Community members cannot punish instigation of such gossip directly. The penalization of transmission of false negative information represents the second-best solution. Admittedly, such penalization chills some true negative speech. Even if harsh punishment for diffusion of false defamation deters some true defamation, however, such punishment secures the socially optimal level of defamatory speech. Instigators, who are thereby denied a market for false negative speech, must now adhere to the norm to satisfy their preferences for esteem. A strong law of defamation, therefore, has a favorable effect on overall norm adherence within a community.

An esteem theory of norms recognizes two types of costs for false defamatory speech: individual victim costs and social costs of decreased norm obedience. The Court's defamation jurisprudence certainly contemplates the former type of defamation costs. Justice Powell recognized, for instance, that "[t]he legitimate state interest underlying the law of libel is the compensation of individuals for the harm inflicted on them by defamatory falsehood."[150] Similarly,

[150] *Gertz*, 418 U.S. at 341.

Justice Stewart emphasized that "[t]he right of a man to the protection of his own reputation from unjustified invasion and wrongful hurt reflects no more than our basic concept of the essential dignity and worth of every human being."[151] Yet these quotations are illustrative of the Court's one-dimensional conception of reputational harm. The Court never considers the harmful consequences of unchecked false defamation for normative behavior. Reputation commands the Court's attention solely as an individual asset, never as a social mechanism for encouraging cooperative behavior and alleviating collective action problems.

This narrow focus also pervades academic commentary on *Sullivan* and its progeny. "A plaintiff suing in tort for psychic injury arising out of a defendant's speech," notes one observer, "may vindicate one or more of the following state interests: the preservation of mental tranquility, the right to be left alone, and the right to a good name."[152] Even those commentators who would otherwise fault the Court for undervaluing reputational costs in the *Sullivan* calculus implicitly concede the individual nature of these costs. Professor Schauer, for instance, admits that the increased levels of self-government speech under *Sullivan* create a "general societal benefit" but questions why the costs associated with this benefit "must be borne exclusively or disproportionately by a small subset of the beneficiaries."[153] But the notion that the costs of false defamation fall solely upon the defamed victims misconceives the nature of these costs. False defamation causes decreased norm adherence and resurgent social dilemmas, which affect society at large, not merely the parties to a defamation dispute.

Some commentators allude to the general social costs of increased false speech in their discussion of modern defamation law. A critic of the actual malice regime, for instance, notes that "[r]eputation is not some lifeless abstraction, but the summation of all the possibilities for gainful interactions—economic, social, and political—with others that are stripped away by false statements."[154] Another observer describes reputation as a means by which "an individual personally identifies with the normative characteristics

[151] *Rosenblatt*, 383 U.S. at 92 (Stewart, J., concurring).
[152] Logan, supra note 80, at 567 (footnote omitted).
[153] Schauer, Uncoupling, supra note 140, at 1322.
[154] Epstein, supra note 58, at 798.

of a particular social role and in return personally receives from others the regard and estimation that society accords to that role."[155] Nonetheless, these colorful but vague allusions show that the specific relationship between defamation, reputation, and normative behavior remains unexplored in the legal literature. As a consequence, the costs of increased false negative speech under *Sullivan* remain overlooked.

Even if self-government speech is underproduced, the shift from strict liability all the way to actual malice oversteps the mark. An esteem theory of norms predicts the overproduction of false negative gossip and supports a harsh regime of punishment for the transmission of such gossip.[156] A corollary of this proposition is that social costs will increase as the punishment for defamation decreases. The Court in *Sullivan*, by effectively subsidizing the propagation of false defamatory speech, caused just such a decrease. The actual malice regime, therefore, imposes social costs in the form of norm erosion. Even if the constitutional law of defamation represents the ideal reconciliation between self-government speech benefits and reputational costs as the Court understood them, it does not represent the ideal reconciliation of these interests per se. To the extent that the Court overlooked reputation as a social mechanism for ensuring normative behavior, the *Sullivan* regime yielded too much to increased political speech and retained too little for reputation.

The Court was not wrong to adjust the common law strict liability regime in *Sullivan*. For whatever reason, that regime did appear to chill excessively self-government speech.[157] That the adjustment

[155] Robert C. Post, The Social Foundations of Defamation Law: Reputation and the Constitution, 74 Cal. L. Rev. 691, 699–700 (1986) (footnote omitted).

[156] See supra notes 58–61 and accompanying text.

[157] Procedural aspects of common law defamation may have imposed excessive burdens on defendants and disabled them from effectively defending defamation suits. For instance, at common law, defendants had to plead and prove truth. See supra note 67 and accompanying text. *Sullivan* assigned the converse burden to the plaintiff. 376 U.S. at 279–80. Similarly, some substantive aspects of defamation at common law were obviously excessive. The falsity element of the tort, for example, had no substantiality requirement. Defamation defendants could face large money judgments as a result of minor factual errors. See, e.g., supra notes 92 and 96 and accompanying text (describing the minor errors at issue in *Sullivan* itself and the large judgment that resulted). Indeed, one commentator suggests that the Court might have done better to

transformed a strict liability rule into a recklessness rule would only prompt general criticism if the associated costs to reputation were purely individual. People can reasonably disagree about how much a citizen values her unsullied good name. Yet reputation has a peculiar social flavor too—it is a mechanism that modulates normative behavior. The Court's failure to appreciate this function of reputation suggests at least that *Sullivan* went too far and may even show that *Sullivan* was incorrect to go anywhere at all.

This is the basic thrust of the norms-based critique of the actual malice rule. To more vividly explicate the possible injuriousness of the actual malice rule, Part VI offers a graphical representation of the preceding arguments. Part VI then uses this graphical tool to express the rationale behind *Sullivan* and to explain the critique of that rationale from the standpoint of norms.

VI. THE COMMON LAW VERSUS *SULLIVAN* FROM A LAW AND ECONOMICS PERSPECTIVE

A. The Economics of Strict Liability

Tort theorists have long explored the economic justifications for strict liability.[158] Defamation, although a peculiar tort, is susceptible to a similar analysis. Under the Hand formula, courts will impose negligence liability whenever the injurer's marginal cost of precaution is less than the associated marginal reduction in probability of injury, multiplied by the gravity of that injury.[159] Assuming that jurors apply this standard faultlessly, negligence liability will deter just those accidents whose cost exceeds the cost of avoidance.[160] Injurers will adjust their levels of care to avoid only those accidents

constitutionalize a substantial falsity requirement for defamation of public figures. See Epstein, supra note 58, at 794.

[158] See William M. Landes & Richard A. Posner, The Economic Structure of Tort Law 54–84 (1987); Richard A. Posner, Economic Analysis of Law 177–82 (6th ed. 2003) [hereinafter Posner, Economic Analysis of Law]; John Prather Brown, Toward an Economic Theory of Liability, 2 J. Legal Stud. 323 (1973); Guido Calabresi, Optimal Deterrence and Accidents, 84 Yale L.J. 656 (1975); Richard A. Posner, A Theory of Negligence, 1 J. Legal Stud. 29 (1972); Steven Shavell, Strict Liability Versus Negligence, 9 J. Legal Stud. 1 (1980).

[159] See United States v. Carroll Towing Co., 159 F.2d 169, 173 (2d Cir. 1947); Posner, Economic Analysis of Law, supra note 158, at 168 n.2.

[160] See Abraham, supra note 60, at 160.

that they can efficiently avoid. For the remaining accidents, injurers will not take additional precautions, because they are better off causing these accidents and paying the victims damages.[161]

In such circumstances, strict liability, which holds injurers responsible even when avoidance costs outweigh expected accident costs, will cause no change in levels of care.[162] Strict liability will secure no extra safety and may even have the pernicious effect of undermining accident *victims*' incentives to adopt optimal levels of care.[163]

Although strict liability cannot generate added incentives to perform any particular activity more safely, it can create incentives for injurers to engage in safer activities.[164] Even without any effect on care-level decisions, strict liability can induce efficient activity-level decisions.[165]

Negligence liability turns only upon the level of care adopted. As such, it does not create optimal incentives for injurers to consider whether to engage in the risky activity at all, or the amount of risky activity to pursue.[166] Strict liability induces injurers to substitute safer activities for riskier ones. This substitution can cause a "more efficient allocation of resources among possible activity levels and activities."[167] Strict liability is not, however, universally preferable to negligence. As noted above, for many types of accidents, victim behavior can alter the probability and severity of injuries. Imposition of strict liability for such accidents causes inefficiency by re-

[161] See id.

[162] See id.

[163] See William M. Landes & Richard A. Posner, The Positive Economic Theory of Tort Law, 15 Ga. L. Rev. 851, 873 (1981). This perverse effect of strict liability does not undermine its application to defamation, which is best characterized as an "alternative care" tort. See Sheer & Zardkoohi, supra note 59, at 391–92 (describing cases in which the victim's ability to avoid injury is "relatively ineffective or insubstantial" as "alternative care" torts); cf. Shavell, supra note 158, at 1 (describing situations in which "the actions of injurers but not of victims are assumed to affect the probability or severity of losses" as "unilateral" accidents). The marginal cost of accident prevention for the potential defamation victim, who may often not know of the defamation until after publication, far exceeds the marginal cost of accident prevention for the defaming party. Therefore, defamation is both an alternative care tort and a unilateral accident.

[164] See Abraham, supra note 60, at 164.

[165] See id.

[166] See Shavell, supra note 158, at 2.

[167] Abraham, supra note 60, at 165.

moving the incentives for victims to avoid accidents, even when victims are better situated than injurers to do so. A shift from negligence to strict liability may have other drawbacks. When the activity in question has significant positive externalities, the decrease in activity associated with the imposition of strict liability may generate a net social cost. For activities that generate external benefits for the community at large "it is no longer clear that full internalization of victim losses is desirable."[168] The following series of graphs explores the conditions under which imposition of strict liability causes social cost.[169] Figure 1 depicts the conventional law and economics defense of strict liability.

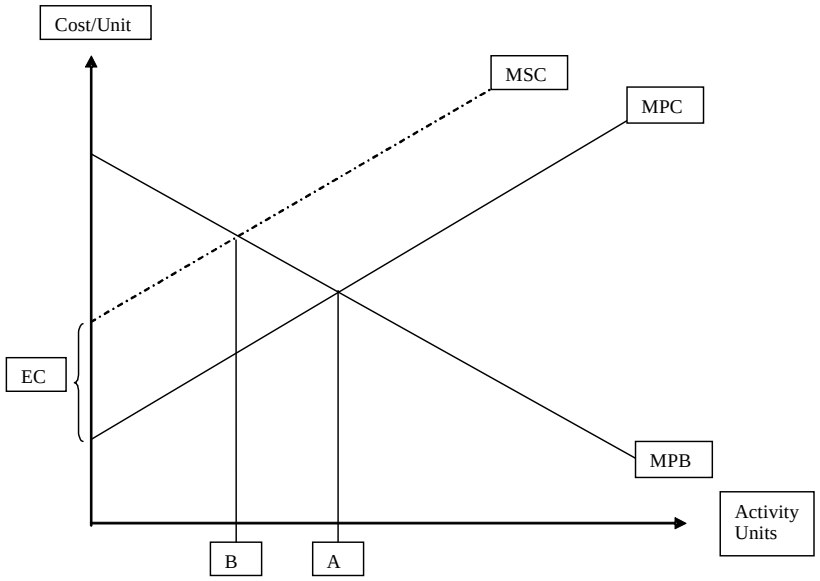

Figure 1

In Figure 1, the horizontal axis measures the injurer's activity level. The downward-sloping solid line measures the marginal benefits accruing to the injurer from additional activity, or the amount he would pay to increase his activity by a small amount

[168] Keith N. Hylton, A Missing Markets Theory of Tort Law, 90 Nw. U. L. Rev. 977, 984 (1996).

[169] Figure 1 and its variants infra are adapted from Hylton, supra note 168, at 986.

(labeled "marginal private benefits" or "*MPB*"). The upward-sloping solid line tracks the injurer's marginal costs from additional activity (labeled "marginal private costs" or "*MPC*"). The injurer's own optimal level of activity occurs at *A*, where *MPB* and *MPC* intersect. For any activity level below *A*, the injurer can increase his activity for a net private benefit; for any level above *A*, the injurer can decrease his activity for a net private benefit.

The Hand formula will elicit the amount of activity corresponding to *A* from the injurer—negligence liability influences only care-level decisions and does not otherwise alter the injurer's privately optimal activity-level choice. In the absence of externalities, *A* is also the socially optimal activity level. When the activity in question causes a negative externality, however, the marginal total costs at each level of activity are greater than the marginal private costs. In Figure 1, the broken line "marginal social cost" or "*MSC*" reflects such a negative externality. The magnitude of this cost, at every level of activity, corresponds to the "external cost" or "*EC*" value marked on the vertical axis. In these circumstances, the socially optimal level of activity matches the intersection of *MPB* and *MSC*, or point *B* on the horizontal axis. Strict liability, by causing an activity-level reduction from *A* towards *B*, causes the injurer to adjust his activity level towards the social optimum.

When the activity in question creates significant external benefits, however, negligence liability may remain appropriate. Figure 2 depicts the situation in which negligence liability is socially optimal.

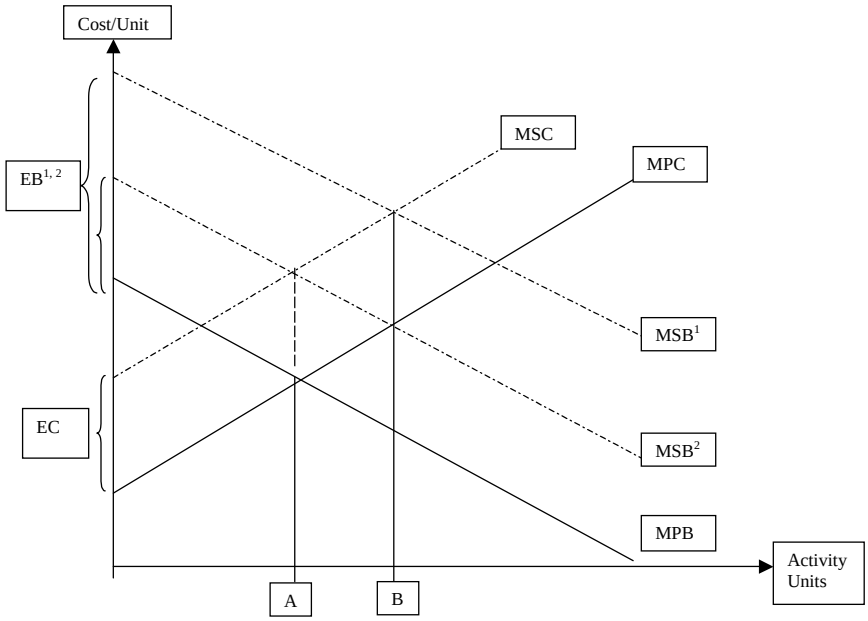

Figure 2

In Figure 2, the activity in question causes both a positive externality and a negative externality. As in Figure 1, the broken line *MSC* represents the negative externality. The positive externality ("external benefit" or "*EB*" on the vertical axis) is the amount of external benefit bestowed upon the public at each level of activity. The line labeled "marginal social benefit" or "*MSB*" represents the amount by which the social benefit exceeds the injurer's private benefits. The intersection of *MSC* and *MSB* corresponds to the socially optimal level of injurer activity. When the external benefits and external costs largely offset each other, the social optimum approximately equals the private optimum. In Figure 2, for instance, the intersection of MSB^2 and *MSC* corresponds almost exactly to *A* on the horizontal axis. This correspondence follows from the almost exact equivalence of EB^2 to *EC*. In such a situation, negligence liability remains appropriate—it does not cause the injurer to deviate from his privately optimal level of activity, which is also the social optimum, *ex hypothesi.*

When the positive externality of a certain activity substantially exceeds its negative externality, however, even negligence liability may be too harsh. In Figure 2, the larger positive externality EB' outweighs EC by a significant amount. As a consequence, MSB' intersects with MSC at point B. B, which represents the social optimum, exceeds the privately optimal activity-level A by a significant measure. In these circumstances, even negligence liability is too harsh because it elicits a suboptimal private activity level. A relaxation of negligence liability, by increasing the activity level towards B, approaches the social optimum.

Courts cannot accurately measure external benefits and external costs for each type of activity. Therefore, the practical legal rule suggested by the foregoing discussion is that courts should impose strict liability when the ratio of external costs to external benefits significantly exceeds one. Similarly, courts should impose negligence liability when these amounts offset each other and a weaker liability regime (recklessness or intentionality) when the ratio of external costs to external benefits is significantly less than one.[170]

B. Strict Liability and Defamation

The Court's defamation jurisprudence rests upon the notion that the self-government benefits of increased true defamation exceed the self-government costs of increased false defamation.[171] In other words, the activity of engaging in defamatory speech has an associated net external benefit. The Court also recognized, however, that false defamatory speech causes individual injury to its victims. As the total amount of defamatory speech increases, false defamatory speech increases as well. The activity of engaging in defamatory speech, therefore, also has an associated external cost. Figure 3 depicts these claims.

[170] Hylton, supra note 168, at 986.
[171] See supra text accompanying notes 141–44.

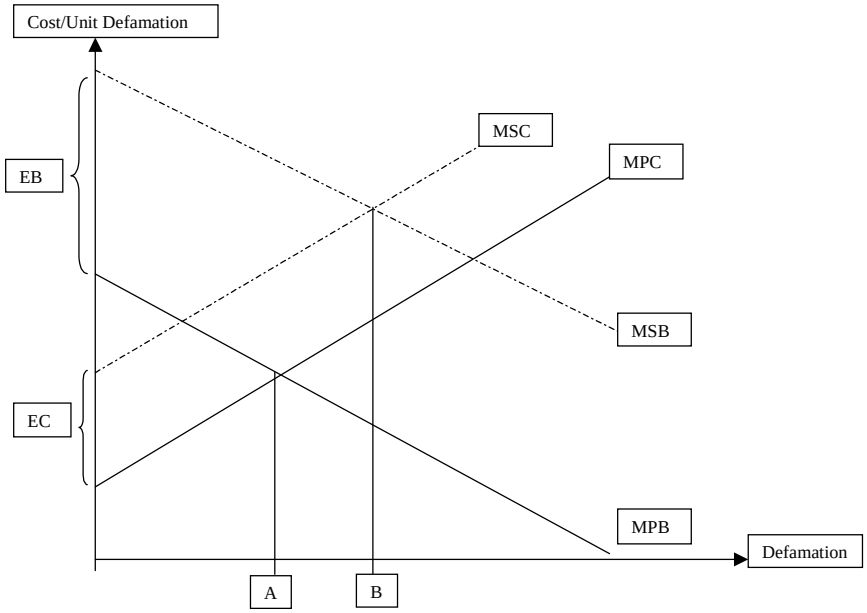

Figure 3

The "external cost" or "*EC*" in Figure 3 represents the aggregate individual psychic costs of defamatory statements. The "external benefit" or "*EB*" corresponds to the self-government benefits of increased defamatory speech. The actual malice rule announced in *Sullivan* follows from the Court's perception that *EB* substantially exceeds *EC*, as in Figure 3, at least for a certain kinds of defamatory speech.[172] By replacing the strict liability common law regime with a recklessness rule, the Court aimed to cause an activity-level increase and a shift towards *B* units of defamation, the social optimum.

[172] The Court has not been able to decide definitively what kinds of defamatory speech create this external benefit. See supra note 136 and accompanying text. The *Sullivan* Court equivocated, finding no need to discuss the contours of the speech that the actual malice rule protected. The *Butts* Court suggested that any speech about a public figure qualified. The *Rosenbloom* Court extended this characterization to any speech on matters of public concern. In *Gertz*, the Court again restricted the characterization to speech about public figures. The confusion after *Dun & Bradstreet* exemplifies this uncertainty. See supra note 136 and accompanying text.

The Court, however, failed to consider the true normative costs of defamatory speech.[173] At a minimum, this failure entails that negligence was a more appropriate replacement for the common law strictures than the actual malice rule. At worst, this failure demonstrates strict liability's optimality and entails that the Court in *Sullivan* upset a sensible reconciliation of the competing costs and benefits. Figure 4 summarizes the preceding claims.

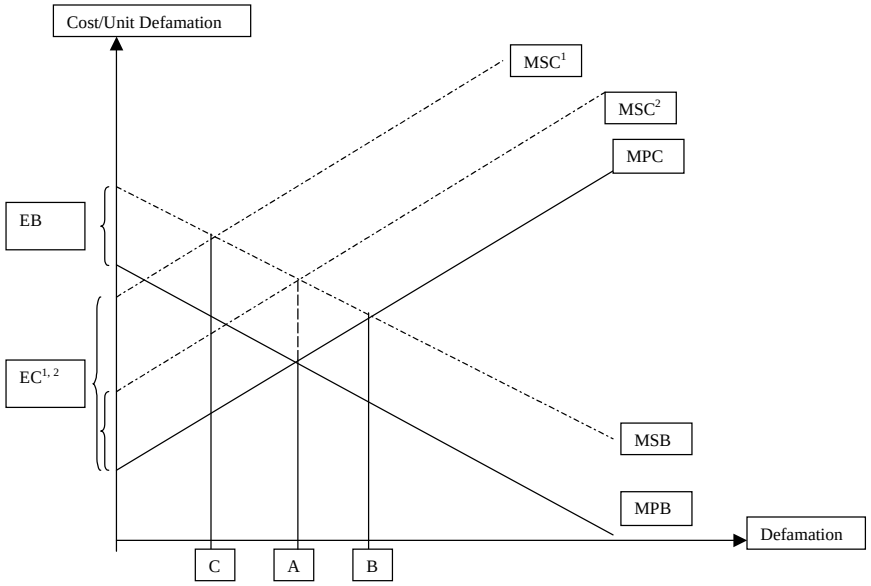

Figure 4

Unlike in Figure 3, the external benefit *EB* in Figure 4 no longer exceeds the external cost *EC*. Figure 4 represents two alternatives to the situation depicted in Figure 3. The broken line MSC^2 reflects the possible equality of defamation's external benefit and its external cost (that is, when $EC^2 = EB$, in Figure 4). The broken line MSC^1 represents the circumstances in which the external cost of such speech exceeds its external benefit (that is, when $EC^1 > EB$, in Figure 4).

The normative criticism of *Sullivan* is that the Court misestimated the size of the external cost. Believing that the true social

[173] See supra notes 150–53 and accompanying text.

costs and benefits corresponded to Figure 3, the Court decided to effectively subsidize defamatory speech by replacing strict liability with actual malice. By doing so, the Court hoped to cause an increase in the level of defamation towards the social optimum. If Figure 4 is a more faithful depiction of the social costs and benefits, however, the actual malice rule is not optimal. In the case in which the self-government benefits of defamation are approximately equal to the normative costs (that is, when $EC = EB$ in Figure 4), the actual malice rule does too much: it replaces a strict liability regime with a recklessness regime when a negligence regime would have served as a better replacement. In the case in which the normative costs exceed the self-government benefits (that is, when $EC > EB$ in Figure 4), the common law rule of strict liability represented the social optimum. Given this optimality, the Court's shift away from this rule towards actual malice represents social waste.

The central claim of this Note is that once the calculus includes the normative effects of false negative speech, as opposed to merely the individual psychic costs of false defamation, any regime of punishment more lenient than negligence necessarily causes social harm. The normative costs of false speech are at least as large as the self-government benefits of the concomitant increase in true defamation.[174] When these two values are equivalent, actual malice is as bad as strict liability, if not worse.[175] At best, therefore, incorporation of normative costs into the social calculus implies that the Court's *Sullivan* revolution was indefensible.

[174] See supra Section V.B.

[175] Even under an equality of these costs and benefits, a strict liability regime may remain superior to a recklessness regime. A recklessness or negligence test has high associated administrative costs. See Abraham, supra note 60, at 163. The forensic inquiry into whether the defendant ignored a substantial risk of falsity does not admit of easy resolution. See generally Brian C. Murchison et al., *Sullivan's* Paradox: The Emergence of Judicial Standards of Journalism, 73 N.C. L. Rev. 7, 11–12 (1994) ("[B]y permitting the use of circumstantial evidence of journalistic behavior to prove the journalist's state of mind, the *Sullivan* rule has spawned a *de facto* set of judge-made standards that covers all aspects of journalistic behavior."). In standard tort theory, the increased number of claims under a strict liability regime could offset the administrative costs of a negligence or recklessness regime. See Abraham, supra note 60, at 163–64. Commentators have noted, however, that an increase in defamation suits occurred after *Sullivan*. See Smolla, supra note 81, at 4. This increase suggests that the actual malice regime may have secured the worst of both worlds.

The one-dimensionality of the Court's conception of reputation reinforces the argument against actual malice.[176] The Court conceived reputation as a purely individual asset, not a social mechanism for allocating esteem and generating normative behavior. As a result, the Court underestimated the normative costs of the increase in false negative speech under the actual malice rule. If the Court found actual malice appropriate, even in its ignorance of the normative costs of false negative speech, it should follow that consideration of such costs entails the optimality of a negligence regime.

Nevertheless, as commentators have admitted, estimating values for the externalities at issue[177] ultimately reduces to a matter of conjecture.[178] An apologist for the Court might claim that, despite the facial absence of normative costs from the Court's calculus in *Sullivan*, the magnitude of the external benefit from self-government speech still justifies the actual malice rule, even after consideration of such costs. Of course, the correct arbiter of such a claim is an empirical study of the effect of the constitutionalization of defamation law on self-government benefits and normative costs. Without such a study, the debate over the correctness of the *Sullivan* calculus is in danger of descending into a shouting contest between unverifiable assumptions about the magnitude of externalities.

At least one additional piece of evidence, however, counsels in favor of the common law—its inner logic. In an effort to provide further evidence against the actual malice regime, the remainder of this Note discusses how the common law tort of defamation, as outlined in Section IV.A above, exhibits an internal coherence that the contemporary constitutional law of defamation lacks. Given this coherence, the common law was quite plausibly efficient.[179] If so, only a modest revision of doctrine should have occurred in *Sullivan*, as opposed to the actual revolution that the case initiated.

---

[176] See supra notes 150–51 and accompanying text.

[177] *EB* and *EC* in figs. 1–4, supra.

[178] See Hylton, supra note 168, at 986.

[179] The claim that common law doctrines evolve towards efficiency is a mainstay of law and economics. See, e.g., Posner, Economic Analysis of Law, supra note 158, at 252 (asking "[h]ow is it possible . . . for the common law—an ancient body of legal doctrine, which has changed only incrementally in the last century—to make as much economic sense as it seems to do?").

C. Was the Common Law of Defamation Efficient?

An apparent morass of doctrinal intricacy characterized defamation at common law.[180] One treatise frankly admits that "a great deal of the law of defamation . . . makes no sense."[181] Consideration of the normative justification for a strong prohibition on negative false speech, however, clarifies some of this complexity.

At common law, many defamation plaintiffs could recover presumed damages upon proving the elements of the tort. The standard explanation for such damages points to the difficulty that plaintiffs faced in proving actual damages.[182] A normative standpoint, however, suggests a more elegant justification for the common law regime of presumed damages. Under a theory of norms, the statement's falsity and its abstract defamatory character, as opposed to any direct effects on the plaintiff's well-being, represent social cost. By providing a market for false negative statements, the transmission of such statements encourages their instigation. This instigation imposes a social cost because it diverts resources away from norm adherence. The social damage caused by false negative gossip thus turns more on its falsity and its defamatory nature and less on its reference to the plaintiff. In this vein, the defendant's affront to the plaintiff is less important than the defendant's amplification of a false negative statement.

Several commentators have noted the oddity of permitting the recovery of large damage awards without proof of actual harm.[183] Indeed, the large award of presumed damages in *Sullivan* itself

[180] See supra Section IV.A.

[181] Prosser & Keeton, supra note 67, at 771.

[182] See Stanley S. Arkin & Luther A. Granquist, The Presumption of General Damages in the Law of Constitutional Libel, 68 Colum. L. Rev. 1482, 1483 (1968) (footnote omitted); see also Dunn & Bradstreet, Inc. v. Greenmoss Builders, 472 U.S. 749, 760 (1985) ("[T]he rationale of the common-law rules has been the experience and judgment of history that 'proof of actual damage will be impossible in a great many cases where, from the character of the defamatory words and the circumstances of publication, it is all but certain that serious harm has resulted in fact.'") (quoting William L. Prosser, Law of Torts, § 112, p. 765 (4th ed. 1971)).

[183] See John O. McGinnis, The Once and Future Property-Based Vision of the First Amendment, 63 U. Chi. L. Rev. 49, 94 n.185 (1996) ("[T]he Court has rightly been suspicious of strict liability and presumed damages in the libel context."); see also Lewis, supra note 62, at 604 ("Seen from twenty years later . . . [a] state libel rule that allows an official to recover large damages for trivial errors of fact . . . seems grotesque.").

contributed to the Court's denigration of the common law practice.[184] The normative social consequences of defamatory falsehoods may help to explain what otherwise appears abhorrent.

The absolute privileges available under the common law also exhibit a remarkable coherence, viewed from a normative standpoint. In all the instances in which this privilege attached, the chance of false negative speech causing actual harm to the reputation of the defamed individual remains small. As a result an instigator of false speech would not consider, ex ante, transmissions covered by the privilege to constitute a good market for her falsehoods. Harsh punishment of defamation in these circumstances is unnecessary since there is little false negative speech to deter. In other words, the negative externality *EC* for all speech covered by absolute privileges at common law remains small. Further, the positive externality *EB* for true speech that enjoyed an absolute privilege is relatively large. The costs and benefits, therefore, most closely approximate Figure 3, making immunity from defamation appropriate.

An absolute privilege adhered to statements made during the course of judicial and legislative proceedings. Under these circumstances, the forensic capabilities of the two pertinent bodies—courts of law and legislatures—coupled with the heightened motives of various interested parties to ferret out falsehoods,[185] suggest the eventual public exposure of any false defamatory utterances. As such, the potential instigator would discount the value of legislatures and courts as markets for false negative speech. The negative externality of such speech in these situations, therefore, remains small. Counterintuitively, in fact, false negative speech in courts and legislatures may create a public benefit of sorts. If the threat of harsher legal consequences dissuaded duplicitous parties, such as dishonest witnesses, from uttering falsehoods, their subsequent public reprimand would become less likely in turn. To the extent such reprimand represents a social good, any rule stricter than the privilege would undercut this public benefit. The benefits

[184] See *Sullivan*, 376 U.S. at 294 (Black, J., concurring).

[185] An advocate, for instance, has an obvious incentive to identify and publicize false statements uttered by the opposing party or that party's witnesses. Similarly, legislators from rival parties have an incentive to reveal instances of lying by their political adversaries.

of true negative speech, furthermore, exceed the general positive externality associated with regular defamatory speech. An impeaching witness, for instance, who successfully undermines the character of an opposing witness, does the court and the community at large a particularly important service. The smaller external cost and larger external benefit make immunity from liability for defamations uttered in courts and legislatures socially optimal.

An absolute privilege also attached to statements made with the consent of the plaintiff and interspousal dialogue. In both of these cases, the external normative costs of false negative speech are small, perhaps even nonexistent, justifying immunity from defamation actions.

Prohibiting the plaintiff who consented to publication from subsequently suing for defamation merely permits similarly situated plaintiffs to credibly contract around the default rule. Without this absolute privilege, idiosyncratic potential plaintiffs, who honestly desire publication of a statement defaming them, would find it very difficult to secure such publication. Ex post, the law must enforce the plaintiffs' evident desires for publication to make these desires credible ex ante for similarly situated plaintiffs. Since these plaintiffs desire publication of statements defaming them, the negative externality associated with false negative speech cannot be great. Indeed, from a normative perspective, the negative externality cannot exist at all since the instigator and the defamed individual are the same person. In such a situation, it would be nonsensical to speak of the instigator taking advantage of the mechanism of social gossip to disesteem a social rival. Figure 3, therefore, most closely approximates the social costs and benefits, making immunity from defamation socially optimal.

For communications between spouses defamatory of third parties, the policy rationale of encouraging marital candor may trump the reputational interests of defamed parties and the associated norm effects. Even punishing knowingly false statements uttered between spouses creates no beneficial normative effects. The defaming spouse's evident desire that her partner not associate with the defamed person will ensure marital disesteem of that person, even if the law affords him a remedy. Alternatively, the existence of a close relationship such as marriage furnishes a good proxy for the veracity of defamatory speech within such a relationship. In

such circumstances, the instigator of a defamatory falsehood will discount the value of the interspousal market for defamatory falsehoods about his social rivals. Either this market will benefit him by happenstance (when the defaming spouse knowingly communicates a falsehood in order to ensure nonassociation for reasons independent of the defamatory information) or this market will weed out false speech rendering its instigation less useful.[186] Interspousal defamation, therefore, creates a small negative externality and has a large positive externality, making immunity from liability socially optimal.

In all the instances in which the common law afforded an absolute privilege, the attendant social costs and benefits most closely correspond to Figure 3. This correspondence makes no liability preferable even in comparison to negligence.

A similar unifying normative explanation also underlies the common law conditional privileges. Recall that a showing of lack of reasonable grounds for belief in the defamatory facts defeated all conditional privileges. Such privileges, therefore, effectively created a negligence regime for certain categories of negative speech. Larger positive externalities and smaller associated normative costs characterize all these species of communication. Unlike run-of-the-mill defamation, the speech to which a conditional privilege attached has externalities that offset each other. Figure 2 represents the appropriate descriptor and common law conditional privilege secured the social optimum.

A commentator has already demonstrated the suboptimality of a strict liability regime for negative employment references[187] and defended the negligence standard for such speech.[188] That such references enjoyed a conditional privilege at common law should come as no surprise.

A conditional privilege also attached to statements made to protect the legitimate interests of a third party to whom the defamer

[186] Knowing defamation of third parties within close relationships such as marriage may even represent a social good. Such defamation often partakes of a commiseration with the partner rather than vindictiveness towards the third party. To fully sympathize with a disappointed spouse, for instance, a partner may utter white lies to vindicate the spouse's emotions. Permitting the expression of such sympathy and the intimacy that it creates supports the extension of an absolute privilege at common law.

[187] See Verkerke, supra note 58, at 159–60.

[188] See id. at 123 n.23.

owed a moral or legal duty of protection, and to statements made by speakers who shared common interests with auditors. In both these circumstances, the magnitudes of the positive and negative externalities of defamatory speech make negligence the optimal liability rule. For instance, the conditional privilege for speech made to protect a third party's interests covered statements made by a lawyer on behalf of his client. Relationships such as this, involving delegated protection of interests, take advantage of gains from specialization. The lawyer can advocate the client's interests more ably than the client himself. In Figure 2, these gains represent an increased *EB* and effectively shift the *MSB* curve outwards, making equilibrium near *A* more likely. In addition, the presumption of a fiduciary relationship suggests that negative speech about third parties in these circumstances has a small associated *EC*. Lawyers have professional and pecuniary interests in ensuring the veracity of speech they utter on their clients' behalves. Potential instigators would therefore discount the value of this particular market as a method of disseminating false defamation about their social rivals. A decrease in *EC* and an outward shift in the *MSC* curve in Figure 2 reflect this dynamic, again making equilibrium near *A* more likely. Negligence represents the social optimum for speech made to protect a third party's interests.

Of particular importance for this Note, a conditional privilege applied to fair comment on matters of public concern. At common law, matters of public concern assumed an expansive definition.[189] Even the common law, therefore, recognized the positive externality of self-government speech, which will remain underproduced without a governmental subsidy. Admittedly, a majority of U.S. jurisdictions forbade such privilege to false statements of fact.[190] Further, the notoriously muddy distinction between opinion and fact may have had a similar kind of chilling effect on self-government speech as underlies the modern vagueness doctrine.[191] Unsure of whether courts would find speech opinion or fact, would-be speakers, erring to avoid liability, would have refrained from uttering even some clear opinions. Nonetheless, the proper corrective for

[189] See supra note 85 and accompanying text.

[190] Sheer & Zardkoohi, supra note 59, at 371 n.19.

[191] For judicial exposition of the chilling effect of vague statutes regulating speech, see Grayned v. City of Rockford, 408 U.S. 104, 108–09 (1972).

such a chilling effect is an extension of the fair comment privilege to all factual assertions. A rule of actual malice does too much and increases defamatory speech beyond the social optimum. Ignoring the jurisprudential quirks, therefore, speech covered by the fair comment at common law has offsetting positive and negative externalities, making negligence liability socially optimal.

Given the common law's formal elegance, it is surely correct to claim that "[f]ederal judges should be at least aware of the possibility that these common law rules contain a greater inner coherence than first meets the eye."[192] If the law that *Sullivan* replaced did in fact possess such coherence, and if the intricacies of the common law doctrine point to social optimality, any shift away from that law carries risks of social cost. These risks are at their highest when the change is radical rather than incremental. *Sullivan* began just such a radical transformation.

That other common law nations have retained harsher defamation regimes further buttresses the foregoing defense of the pre-*Sullivan* common law of defamation. Indeed, this retention has occurred despite scrutiny of the governance impact of strict liability for defamation by the highest courts of the pertinent jurisdictions. The Canadian Supreme Court, for instance, resisted adoption of the actual malice rule for defamation of public officials. The Court found that:

> False and injurious statement [sic] cannot enhance self-development. Nor can it ever be said that they lead to healthy participation in the affairs of the community. Indeed, they are detrimental to the advancement of these values and harmful to the interests of a free and democratic society. . . . A good reputation is closely related to the innate worthiness and dignity of the individual. It is an attribute that must, just as much as freedom of expression, be protected by society's laws.[193]

[192] Epstein, supra note 58, at 791.

[193] Hill v. Church of Scientology of Toronto [1995] 126 D.L.R. (4th) 129, 159–60 (Can.); see also Theopanous v. Herald & Weekly Times (1994) 124 A.L.R. 1, 22–23 (Aust.) (refusing to adopt actual malice for defamation of political figure, finding that the test "gives inadequate protection to reputation"); Derbyshire County Council v. Times Newspapers, 1 All. E.R. 1011 (H.L. 1993) (declining to adopt actual malice test). See generally Fischer, supra note 81 (examining defamation law from England, Australia, and New Zealand).

Given such pronouncements by other learned legal bodies, defenders of the actual malice rule must awkwardly argue that democratic governance suffers abroad at the hands of defamation or that peculiar aspects of U.S. life make political speech particularly fragile. A consideration of the political traditions of these nations, however, suggests the more plausible conclusion that strong defamation laws can coexist with sensible free speech in a liberal democracy. Indeed, protection from negative false speech may actually enhance public life by sustaining normative solutions to collective action problems.

CONCLUSION

This Note has applied some implications of the law and norms literature to the apparently unrelated field of defamation law. In particular, this Note has expounded a norms-based critique of the relaxation of defamation law begun in *Sullivan*. The Court considered reputation solely as an individual asset and conceptualized damage to reputation only as an individual cost in its various constitutional defamation opinions. Under an esteem theory of norms, however, reputation denotes also the disesteem which individuals experience upon public exposure of their norm violations. As such, reputation represents an artful social mechanism for the amelioration of collective action problems—a benefit that accrues to all community members. Underestimating the costs of increased false negative speech, the Court fixated on the chilling of true negative speech under strict liability and instituted the actual malice rule as a corrective. Even this rule, however, credited some countervailing state interest in reputation. If this interest is larger than the Court realized, as this Note claims, then it at least follows that negligence, not recklessness, should have replaced strict liability. The common law may even have been largely optimal, and the Court may have upset a sensible reconciliation of free speech benefits and normative costs when it decided *Sullivan*.

In concluding, this Note examines briefly three objections to the thesis of this Note. The first objection takes issue with a premise of the argument—that norms make human behavior more efficient. The second objection is more general and asks whether the ability to freely criticize elected government officials is essential to a healthy democracy. The third objection is similarly overarching

and questions whether a decision like *Sullivan* can ever be divorced from the historical forces informing it.

First, false negative gossip weakens the norm whose violation is alleged. When a norm resolves a social dilemma, false negative gossip has a socially pernicious effect. Not all norms, however, resolve collective action problems. Some norms cause rather than ameliorate inefficiencies. The norm of nonassociation with blacks, for instance, which motivated whites in the Jim Crow South, prevented some socially advantageous relationships. For norms such as these, the effect of false negative gossip may represent a social good—false negative gossip's erosion of an inefficient norm benefits the whole community. An objector to this Note's thesis might claim that the law should not punish gossip that falsely claims violation of an inefficient norm because any resulting norm failure actually benefits society.

The premise of this objection—that some norms are inefficient—is surely correct. Further, the conclusion that the law should sometimes sponsor false defamation helps explain some otherwise puzzling cases.[194] Nonetheless, separating efficient norms from their inefficient cousins presents no great jurisprudential difficulty. To the extent that a court can accomplish such a partition, the preceding objection justifies only the modification, not the abandonment, of a strict liability regime.

Second, despite the harsh picture this Note paints of speech criticizing government officials, such speech surely constitutes a necessary feature of democratic society. Citizens must be able to air their bona fide grievances in public and complain when they perceive those governing them to have offended the popular will. The ability to engage in such criticism not only benefits the speaker, her listeners, and the community at large; it also allows

[194] One such case is *Connelly v. McKay*, 28 N.Y.S.2d 327 (Sup. Ct. 1941). The plaintiff maintained a service station and rooming house primarily patronized by interstate truck drivers. Id. at 328. He alleged defamation by a statement that he was informing the Interstate Commerce Commission ("ICC") of the names of truck drivers who violated ICC rules. Id. The court refused to give him relief, finding that an allegation of "giving information of violations of the law to the proper authorities" could not "constitute a foundation upon which to build an action for slander." Id. at 329. Quite plausibly, the court declined to afford the plaintiff relief because the norm of failing to inform on criminal truck drivers, which the plaintiff allegedly violated, is socially inefficient.

citizens to hone their analytical and oratory skills, to develop in themselves and their audience a public spiritedness that strengthens republicanism, and sometimes even to persuade others to progressively alter their worldview.

The ability to criticize one's rulers, however, may also harm the community when the object of derision is innocent of the bad acts alleged. Further, public speech creates opportunities for less virtuous human tendencies to flourish—influencing one's audience by deceit and scare-mongering, eliciting from them vindictiveness in place of public spiritedness, and even debasing public discourse generally with sensationalism. The law should be alert to these dangers and balance them against the benefits of speaking out against the government. A regime of negligence for defamation touching on public matters accomplishes just such a balance.

Third, *Sullivan* was decided during the heyday of the civil rights movement. Potent historical and social factors influenced this piece of jurisprudence just as much as they influenced other landmark decisions of the era. To ignore these factors and construct a criticism of the actual malice regime with anemically ahistorical formalism risks missing the forest for the trees.

Undoubtedly, Southern state governments used substantive law to oppress blacks. Strong defamation laws deterred commentators from disseminating news of the racial situation in the South to the rest of the country. State defamation law thereby forestalled the national political pressure for reform that eventually came to bear on the South. Southern defamation law, in deterring individual instances of false negative speech, prevented dissemination of the true negative picture to the remainder of the country. Viewed in this light, the Court's holding in *Sullivan* may have been an attempt to ensure free diffusion of the Southern racial situation to the national community.

While this historical commentary may justify *Sullivan*'s narrow holding, it does not justify the precedential shockwaves it created. In preventing the South from abusing its own defamation law, the Court thwarted all states from benign use of such law to secure socially beneficial goals. A state need not co-opt defamation law for such nefarious purposes, and more astute jurisprudential solutions to this peculiar Southern problem surely existed. A scalpel excises a localized injury without damaging healthy tissue better than a

saw, even when the injury is as nasty as that which afflicted the South.

Of course, a First Amendment absolutist might shrug her shoulders at the normative costs of the actual malice regime and remind us that the Constitution often requires expensive disregard for competing social values. The Court, however, has not required free speech to trump all other ends, and acknowledging the social milieu in which defamation occurs gravitates in favor of a harsher regime than exists today.