

Plea bargaining, discretionary leniency and the making of political authority

by Mary E Vogel

My purpose here is to explain why plea bargaining arose by taking a historical and macrosocial approach to the issues involved.

Though highly controversial, the origins of plea bargaining are surprisingly obscure. While often thought to be either an innovation or a corruption of the American courts after World War II, it has much deeper historical roots. Plea bargaining is defined here as a defendant's entry of a guilty plea in anticipation of concessions from the prosecutor or judge. It may be implicit or explicit and need not yield concessions in every case.

To explore the rise of plea bargaining, I examine its beginnings in antebellum Boston, Massachusetts – the first sustained instance of the practice known to exist. Boston was a national centre of legal innovation from which many legal ideas and practices spread to other cities through diffusion (Novak, 1996). Plea bargaining very probably was one such distributed legal innovation. An urban political elite, seeking to maintain its position of power, played a key role in its establishment. It was this elite's perception of crisis and threat, along with its effort to preserve social order, the legitimacy of self-rule and its own dominance, that produced in a single New England locale the practice of plea bargaining that would then become a national and, eventually, international phenomenon.

POST-INDEPENDENCE CONFLICT: CRISIS AND THE RE-MAKING OF POLITICAL AUTHORITY

In the years after the American Revolution, politicians worked to re-create political authority anew for a self-governing republican society. Yet their project faced the obstacle that this authority was to be anchored in popular self rule but, at the same time, to be constructed during the 1830s which was a period when concentration of wealth and economic inequality increased more rapidly than any other time in the 19th century. Recently, historians such as Gordon Wood (1992), have shown compellingly how intensely conflicted was the social and political landscape of the early American republic.

This “formative era” of American law was one of perceived crisis of unrest and political instability in the republic. Its timing was crucial because it occurred just as suffrage was “universally” extended. Together these events evoked new state responses to social conflict (on the importance of timing in the convergence of social and political forces, see further Nicos Poulantzas, *Political power and social classes*, New Left Books, London, 1975; Reinhard Bendix, *National building and citizenship*, Wiley, New York, 1964). As the voting public grew, uncertainty ran high as to whether self-governance would prove viable and what path politics might take. This public concern also spawned the movement for common schooling in Boston under the leadership of Horace Mann which marked the beginning of public education in the United States.

State response to the crisis was needed that would be defensible in a world of popular rule. Amidst a re-scripting of legal practices that took place, one innovation, plea bargaining, achieved special prominence. Plasticity of institutions and practices at this creative time, when judges were forging legal institutions into a modern form much of which they retain today, facilitated creation of new legal mechanisms and, once formed, allowed them to achieve a permanence not otherwise possible.

During the 1830s and 1840s, rioting and unrest were widespread. The earliest factories were constructed which changed both forms of production and working conditions. City life brought diverse strangers of unequal rank into contact. These, coupled with new waves of immigrants, created a vibrant and tumultuous urban scene. Officials, already focused on the danger that conflict posed to property, social order and growth, grew anxious (Horwitz, 1977; Nedelsky, 1990). Religious belief, previously a source of cohesion, and social consensus, which had pervaded small scale community life, were also eroding (Nelson, 1981; Lockridge, 1981). Constant spatial movement and turnover among residents in city neighbourhoods amplified the strains of inequality (Sellers,

1991). Irish immigrants began to coalesce as a major presence too. Amidst these pressures, conflict, unrest and violence, rather than harmony, was the order of the day.

Because self-rule was still new and local political capacity for responding to conflict was limited, there arose a sense of crisis and of threat to both the social order and to the elite power embedded in it. It was desire, in this context, to protect order and to reconsolidate the city elite's partisan control that elicited new state responses.

Spurred on by the election of Andrew Jackson to the Presidency, strikes by the Workingmen's movement swept the American northeast between 1833 and 1836. Their crusade was for a 10-hour working day – a goal that, by 1836, had essentially been achieved (Sellers, 1991, p 338). Yet, their discourse endured. Labour leaders charged that “capital divided society into two classes, the producing many and the exploiting few, by expropriating the fruits of labour” (Sellers, 1991, p. 338). Working men challenged growing inequality that let a privileged few flourish at the expense of many. Resentment simmered. By the 1830s, public concern was widespread about the future of republican self-rule. Workers began to use the language of republicanism in new ways that now viewed the holistic interests of the society through a new modestly socialist lens (Forbath, 1991). Social disorder, riots and strikes riveted elected officials. To defuse resentment and reassert control, they turned for help to the ideology of a “rule of law.”

By this point, conflict had gripped the public imagination. Ethnic diversity and contention soared as did images of the Irish as one major source of the turmoil. Adding to these ethnic tensions was a palpable public fear of crime – especially violence. Recently, Eric Monkkonen (1996) has presented data for New York City showing that the 1820s experienced not just heightened fear, but a very real increase in actual homicide – a trend that peaked nationally during the 1850s (Gurr, 1981). Addressing the Boston City Council on September 18, 1837, Mayor Eliot decried the threat posed by “the incendiary, burglar and the lawlessly violent” which was “increasing at a ratio faster than that of the population” (cited in Lane, 1971, p 34).

Data on arrests in Boston are available only for 1831 and 1850 so that, questions of the relation of arrests to offences actually committed aside, detailed analysis of arrests over this period is not possible. However, data on commitments of those convicted to the Boston Jail show a 45 per cent increase in just four years between 1830–34 – the only early 19th century years for which those records appear to exist – as compared to only a 25 per cent increase in the population of the City of Boston for the entire decade 1830–40 (Council of the Massachusetts Temperance Society, 1834, p 81 and Handlin, 1979, p239).

Probably the clearest sign of public fear is that the Mayor requested and obtained funds to establish a paid police

force for the city. Pointing to the “spirit of violence abroad”, Eliot argued that the residents must be protected. Whether disorder and crime actually were rising or were simply perceived now as more threatening, it is clear that violence was pervasive.

During the 1830s, a remarkable spate of riots and routs occurred. Two events, in particular, brought public distress to a fever pitch – the burning of an Ursuline Convent in Charlestown in a flare-up of anti-Irish sentiment in 1834 and the famed Broad Street riot of 1837. After the “Mount Benedict Outrage”, as the convent fire was known, public agitation soared. It was seen as “a riot with social, even political implications”. (Lane, 1971, p 30). By then, city officials, who were very aware of similar happenings in England and on the Continent, were acutely sensitive to the political potential of such events. (Lane, 1971, p 30). In 1836 and 1837 an economic downturn, followed by financial panic, further fanned fears about the fragility of the new order. Unease created by daily contact among persons of diverse ranks in the city amplified fears as the lives of the poor impinged ever more on the consciousness of the affluent (Lane, 1971).

As labour unrest, ethnic conflict and crime mounted, shockwaves were buffered less than they had traditionally been by the erosion of shared religious values and cultural commonalities (Wiebe, 1966). Thus, during the 1830s, when social conflict grew, amidst weakened cultural consensus, it produced an acute sense of crisis in the new order. Response to this crisis was shaped by its timing which caused state action to be devised in the context of two other key happenings (Poulantzas, 1975). Extension of the vote meant that any initiative must take a form that would sustain the popular consent crucial to self-rule. There was also emerging a new conscious campaign to promote social policy and the “people's welfare” through law (Horwitz, 1977). The fact that crisis emerged during this “formative era” of American law created a special window of opportunity for cultural change.

THE LEGAL ESTABLISHMENT AND THE LEGACY OF POST-REVOLUTIONARY FEDERALISM

Almost without exception, the Massachusetts bar, after the War of Independence, consisted of, first, Federalists and, later, Whigs (Warren, 1931, pp 174, 178). During the “formative era”, lawyers achieved new influence after they recovered from an immediate post-Revolutionary period of disrepute. This political collegiality in the sympathies of the bar was reinforced by a long tradition of fraternity, practical exchange of ideas and fellow feeling, on the one hand, and, first, Federalist, and, then, Whig control of the Commonwealth's judicial appointments, on the other, to ensure that the courts were presided over by judges in step with the policies of these successive elite-dominated parties.

Concern about lawyers' political allegiances was aggravated by the extensive part they were playing in state government. Nathaniel Ames argued that separation of powers was breached as lawyers wrought their influence simultaneously by their votes, their courtroom activities, and their candidacies for elected office (Warren, 1931, p 179). Denouncing the lawyers' influence in colourful terms, Nathaniel Ames wrote: "... he that is not now a Lawyer, or tool of a Lawyer, is considered only fit to carry guts to a bear in New England" (Nathaniel Ames, *Columbian Minerva*, September 6, 1803; cited in Warren, 1931, p. 180).

Thus, the "law craft", as a bastion of Federalism, possessed a distinctive ideological stance. As lawyers' status improved after the Revolution and they moved between careers in the bar, the judiciary and politics, they carried with them the unique political outlook of the Federalist/Whig elite and, with it, a clear commitment to their policies as ones that might best serve the "public good." As criminal courts innovated in their efforts to contain conflict, protect property and dampen the violence and rioting so destructive to prosperity, first, Federalist and, later, Whig ideas coloured the thinking of judges about the need for order and what policies might achieve it.

POPULAR CHALLENGE TO THE COMMON LAW

Critique of lawyers, largely on grounds of their Federalist views, gradually came during the early 1800s to be associated in the public mind with opposition to the common law. Many states had, after the American Revolution, initially adopted much of British Common Law and public attitudes toward it had been positive. After 1800, however, things changed (Horwitz, 1977, p 5). Previously the Common Law had been viewed as a fixed, customary standard. Judges envisioned their task as discovery and application of pre-existing rules (Horwitz, 1977, pp 8–9). This produced a strict conception of precedent and a popular view of law as, if not always fair, at least known.

In the closing years of the 18th century, however, signs of change appeared in both criminal and civil spheres (Horwitz, 1977). Its roots were twofold. The first was states rights constitutional theories which depicted law-finding based on precedent as a form of "*ex post facto*" law. The second was new conceptions of the basis of legitimation of political authority which portrayed the customary approach of common law as outdated in light of popular sovereignty. The constitutional challenge argued that if judges could impose criminal penalties without laws being enacted in statute, the application of precedent after an act had occurred constituted *ex post facto* law. It punished a person left in ignorance at time of the act of precisely what the law prohibited and, thus, breached constitutional limits on state power (Horwitz, 1977, pp. 11 and 14). As

ideas about the basis of political authority changed, it meant that the common law, with its roots in custom, was incompatible with authority based on laws reflecting popular sovereignty.

Initially these sentiments generated calls for abandonment of the Common Law and a move to enacted statute. The codification movement sought to recognize primacy of "the peoples" elected representatives and to move from case law to statutory enactments.

Yet, judges, who were overwhelmingly Federalist-appointed, and political leaders resisted the move to statute precisely because of the power it would have given to legislative bodies dominated by the middling and lower classes. Instead, they, together with what was then the Federalist elite, fought to maintain judicial discretion by preserving reliance on the common law (Horwitz, 1977, p 21). Although the codification movement ultimately failed, it signalled public interest in reducing the discretion of judges and in simplifying and clarifying law, before the fact, and communicating knowable legal rules and procedures to the citizenry. In part the movement foundered due to a compromise proposal advanced by Joseph Story that "a digest [be prepared], under legislative authority, of the settled portions of the common law" (Jones *et al*, 1993, p 35).

Complexity, a separate but related matter, also became a basis for challenge. In response to persistent criticism of lawyers, the courts and even the common law, Supreme Judicial Court Justice Theodore Sedgwick, a conservative from Stockbridge, urged reform arguing that only if the courts were "wise, simple and expedient" would people consider it "the most certain means of attaining justice" and use them rather than extra-judicial means for resolving conflicts (Jones *et al*, 1993, p 30).

EXTENSION OF THE FRANCHISE AND THE POLITICS OF CONSENT

By the end of Jackson's second Presidential term in the mid-1830s, "universal" suffrage was a fact of life and reconstituting politics. By easing restrictions, such as property ownership and the poll tax, states extended the vote to new segments of the labouring classes though a goodly share of Boston's citizens had already voted before. Artisans and workers now produced more representative assemblies though a lingering tradition of deference meant that the result was not immediate. Elected leaders, in turn, faced new constraints as their decisions increasingly required at least some broad based popular consent. This abetted a move, already under way, to challenge the political control of Boston's Federalist elite (Lane, 1971). It also aroused worries about what other forms, particularly with respect to property, contestation might take. Joseph Story noted, at one point, that the lawyer's most "glorious and not infrequently perilous" responsibility was to protect the "sacred rights of

property” from the “rapacity” of the “majority” (Story, 1829; cited in Mensch, 1982).

While proprietors complained that conflict marred quality of life, city leaders worried about even more far-reaching consequences (Lane, 1971). Familiar as they were with the rioting and revolt in Europe during the 1820s and 1830s, Boston’s politicians worked feverishly to restore order, reconsolidate their partisan base, and cement popular commitment to the institutions of the republic. Because the franchise precluded solutions to disorder and unrest that jeopardized voter support, new responses had to be devised not only to violence, property crime and riot but to growing political tensions too.

The courts, which provided Americans’ primary experience of the state before local political parties formed in the 1840s, now assumed a key role (Skowronek, 1982). Beginning in the 1820s, a first wave of court reform had established the Boston Police Court. It was a reform spearheaded by Boston’s leading citizens and it aimed to re-establish the lower courts as a respected and well-used forum for resolving conflict (Hindus, 1980; Dimond, 1975). This responded both to the demands of the propertied for security and, even more, to the “claims [for a just forum on the part] of a [lower] class [whom they felt it] unsafe to deny” (Lane, 1971, p 23). Additional institutional and cultural changes followed very soon.

By the 1830s, Boston’s local officials were “no longer so firmly united by ties of class and [state] party [affiliation] as their predecessors [had been]” (Lane, 1971, p 46). The city maintained a one-party tradition where “candidates labelled Democrat ... had [in most years virtually] no chance of ... [electoral] success” (Lane, 1971, p 47). However, the times were creating intractable dilemmas for the beleaguered [Federalist and, then, Whig] municipal authorities and “hopes for the material future were [increasingly] balanced by fear for the political” (Lane, 1971, pp 47, 60). Under pressure, elite Bostonians experimented with new alternatives. To take one step “backward” to reconsolidate elite power, this city with its tradition of single-party Federalist/Whig control was forced to take several small steps forward in the service of consensus-building and reform.

Strategies to restore order were conceived, then, at a time that precluded politics as usual. The Whigs feared threats, not only to property per se, but even more to the stability in day to day affairs that investment and growth required. (Adam Smith, in *The Wealth of Nations* (1776), cited the task of providing the security and predictability needed for commerce as one of two essential roles of the post-mercantilist state. Insurance companies were working at precisely the time of this study to rationalize and diminish risk (Steinberg, personal communication with the author, 1997)). Fearing for the future, leaders worked to nurture order and predictability in public life and to cultivate the consent of citizens to both institutions of self-

rule and the stewardship of their party. To this end, they approached social control, not through overtly coercive means, but in ways that underscored the party’s claim to serve the will of the people.

City officials and elite civic leaders accomplished this by appealing to the pre-eminent social discourse of the day – that of a “rule of law”. By common agreement, they argued, social life must proceed according to a body of rules specified in advance and oriented to fairness. Such rules, they contended, apply universally to every citizen and prescribe equal treatment for each accused person in court. Then, in a dramatic claim, it was argued that, even when such rules depart from popular opinion, they must, unceasingly, be observed. Only through adherence to legal principles and procedures, leaders argued, could the new project of self-rule be sustained. By appealing to the language of the widely revered “rule of law” as a basis for cultural imagery, they hoped to bolster both social order and the legitimacy and authority of republican institutions. With order restored, they believed they could re-secure their hold on power.

The language in which the reforms were advanced reveals how officials viewed them. When, in 1822, the Police Court had been established, Mayor Josiah Quincy unveiled his plan by denouncing the potential for social conflict inherent in the previous system of Justices of the Peace. Quincy argued that “whenever confidence ... [lapses] in the lower tribunals, there is no justice ... [for] the poor, who cannot afford to carry their causes to the higher ones” (Quincy, 1822, pp 7-8). Such injustice, he proclaimed, corrupts the morality and political commitment of citizens. Quincy referred, among other things, to the prior fee structure whereby magistrates had prospered the greater the number of cases heard. Anticipating a point later made by Max Weber, Quincy argued that where political authority anchors its legitimation in legal rules, the danger is especially great when that law is perceived as unjust. The risk is that laws, so viewed, may be treated as no law at all and that political authority itself will then be undercut. Following quickly upon court reform, other major new institutions including prisons and reformatories, a house of industry and a professional police force were also set in place.

As new institutions moved into motion, judicial decision-making and court procedure also changed – although more informally and incrementally. Judges’ decisions took on a policy focus (Horwitz, 1977). In the criminal courts, pardons, the *nolle prosequere*, the plea of *nolo contendere* and grants of immunity had already begun to be used in new, explicitly conditional ways to further specific policy goals.

Plea bargaining now made its debut in the courts. Here Max Weber’s insights into legal change are helpful. Weber argues that development of new legal norms and practices has always been the product of innovation, or the

construction of new lines of action, in settings where an existing repertoire does not suffice (Weber, 1978, pp 753–784). Typically these changes are initiated, Weber contends, by status groups acting on the basis of interest. As time goes on, the innovations first acquire the power of habit, then, of norms and, finally, they are formalized in law. Plea bargaining developed very much in such a progression.

Although the practice of plea bargaining arose during a period of reform, it was not advanced as a unitary plan or formal initiative. Instead, it emerged as an informal and pragmatic accretion of small changes in the customary practice of the courts that was, only then, culturally codified. In light of the small gradual shifts in court practice through which plea bargaining emerged, the origin of this practice resembles other key legal developments, such as the rise of the prosecutor, described by John Langbein (1973).

LAW AS AN INSTRUMENT OF SOCIAL POLICY: JUDICIAL ACTIVISM

During the “formative era”, judges began to reconceptualize American law as an instrument of social policy. This transformation in law, combined with state structure, made it likely that political response to crisis would come through the courts. As judges changed the way they envisioned their role, they increasingly crafted their decisions with an eye to policy implications beyond their case at hand (Horwitz, 1977). In private law, case decisions aimed to facilitate healthy markets and economic growth (Horwitz, 1977). In the criminal courts, judges sought to assure behaviour that would uphold social order and, especially, foster the security and predictability needed for development (Vogel, 1999).

The early 1800s had been a “disruptive and potentially radical period” (Mensch, 1982, p19). As American leaders and jurists worked to re-establish post-independence political authority, they came to rely heavily on the courts where the role of judges was changing. It was the effort to reconcile the tension between judicial discretion and popular will, mentioned above in the context of the common law, that contributed mightily to what Horwitz (1977) has called the “transformation of American law”. Judges increasingly bridged the gap between common law and “the people” by envisioning themselves as agents of “popular sovereignty”. They came to view their role as that of activist and innovator functioning on behalf of the “common good” (Horwitz, 1977, p 30). In the course of this change, judges began to view law as a policy instrument (Horwitz, 1977). Increasingly, judges articulated decisions and used law as a tool to shape the path of social change. In Mark DeWolfe Howe’s words, “it was as clear to laymen as it was to lawyers that the nature of American institutions ... was largely to be determined by the judges ... (and that) questions of ... law were ... considered as questions of social policy” (Howe, 1947–50;

cited in Horwitz, 1977). Howe’s words bespoke a conscious turn by the state to the courts, among other institutions, to promote its policies.

Judicial discretion, specifically in sentencing, was no exception. Mayor Josiah Quincy emphasized the existence of such discretion in his address to the Grand Jury of Suffolk County when he observed “There is, indeed, a discretion invested in judges” (Quincy, 1822, p 12). That he believed such discretion should be informed by social policy in shaping sentencing Quincy left no doubt. He proclaimed that “The utility of a concentrated system of penal and criminal law in which punishment shall be graduated by the nature and aggravation of crime and *adapted to the actual state of society and public sentiment* [emphasis this author’s], ... [is] appreciated” (p 14). That judges’ discretion in the criminal sphere centred on sentencing policy, Quincy also emphasized. He noted that a judge’s discretion included selecting “time and place [of imprisonment]” as well as other aspects of the severity of sanction (p 12). Public knowledge of such policy uses of prosecutorial and judicial discretion in sentencing was widespread and the practice met with legislative approval (House Report, Massachusetts Legislature, No 4, January 1845).

EMERGENCE OF PLEA BARGAINING: CHANGES IN THE PRACTICE OF LAW

Given the paucity of local political institutions and the many challenges of the day, during the early to mid-19th century, the courts emerged as central in shaping the relation of citizens to the state. At this point, they stepped forward as agents of the state to promote political stability, to enhance the legitimation of institutions of self-rule by nurturing political authority and to create conditions conducive to healthy economic development. (For more detailed analysis of the historical origins of plea bargaining, see Mary E Vogel, “The Social Origins of Plea Bargaining: Conflict and the Law in the Process of State Formation, 1830–1860”, *Law and Society Review*, vol 33, no 1, March 1999).

Reaching back into the traditions of the common law, the courts turned to mechanisms of discretionary, or episodic, leniency. Through these practices, leniency was frequently, but not always, accorded and so could not be counted on and taken for granted. What was unique about the tradition of leniency was that to qualify for it one relied on the intercession of what were essentially character witnesses to whom one was known. As Thompson (1975) and Hay *et al* (1975) have shown, in England where litigation was also widespread, practices of leniency created incentives to appreciate, nurture and reciprocate social ties and bonds of patronage with those more privileged. In this way, one might benefit from the goodwill to cause a prosecution to be foregone or to have a powerful patron to plead for mercy if one ran afoul of the law.

The result, in England, was a system of justice that reinforced the stability of the class structure, despite vast material inequality, through these social ties that it fostered at the same time that it bolstered political legitimacy by affirming a formal message of universality (ie, law applies to all) and equality (ie, formally equal treatment procedurally) before the law. In the United States, plea bargaining emerged as the most widespread form of episodic leniency – one which also promoted stability – but now in new ways in a context of popular electoral politics.

First signs of plea bargaining appear in the lower court of Boston in the 1830s (Vogel, 1988, 1999) – in Boston, the lower court was the Police Court (later renamed the Municipal Court) which is the equivalent of a country district court. Before that time, both bargained guilty pleas (either explicit or tacitly implicit) and, in fact, guilty pleas altogether were quite rare (Alschuler, 1979; Langbein, 1978). This same “backwardness of the courts in receiving a plea of guilty” is again mentioned in Blackstone’s Commentaries on the Law of England (1765, cited by Alschuler) and approved by Chitty (1816, cited by Alschuler). In fact, judges exhorted defendants to exercise the hard won rights of the republic against self-incrimination and warned of penalties if they failed to do so. Nor, according to prior research by John Langbein and others, did the practice exist in England or elsewhere before the 19th century. Although leniency in the form of pardons and grants of clemency has a long history, those did not involve direct exchange and never achieved the pervasive, routine use that plea bargaining did.

Yet, in Boston, during the 1830s and 1840s, this changed with the beginnings of plea bargaining. Guilty pleas, the first element of bargaining that appears together with concessions in disposition or sentencing, were first entered in significant numbers during the 1830s and, by 1840, were widely accepted – a pattern that continued into the 20th century (Vogel, 1999). Overall guilty pleas surged from less than 15 per cent of all cases entered in the docket in 1830 to 17 per cent in 1840 and then to a high of 88 per cent in 1880 (Vogel, 1999). This dovetails with Raymond Moley’s (1929) New York finding that by 1838, 15 per cent of all felony convictions in Manhattan and Brooklyn were the product of guilty pleas. By 1869 he found that guilty pleas accounted for 75 per cent of all convictions.

However, defendants’ tendency to plead guilty varied among different types of offences with bargained pleas, initially, most common in the Boston Police Court for property offences and least for offences against the moral order (Vogel, 1999). Fisher (2000) suggests that some charge bargaining may have occurred in the mid-tier courts in Massachusetts for several decades in regulatory and capital cases from the late 18th into the early 19th centuries and then abruptly stopped. However, the extent to which the record books from which he draws cases for

study may dwell primarily on complex cases that cannot be encompassed in the docket is not fully known.

Thus, any attempt to generalize from cases “sampled” from the record books to the population of cases appearing before those courts must be viewed with greatest caution. Nor does this study distinguish between negotiated pleas of *nolo contendere* which began in the higher courts and were consummated by lawyers, on the one hand, and the much newer bargained guilty pleas accomplished initially by a defendant before a judge and without counsel, on the other.

Turning to cultural traditions of the common law, Bostonians, during the late 1830s and 1840s, reworked elements of the tradition of discretion and episodic leniency into the practice of plea bargaining which while closing cases in a much and vociferously sought reform, retained for the courts considerable control over both sentencing and its implementation. Judges took standard vehicles of leniency, such as the pardon in which leniency was traditionally granted after conviction, and moved it up to a point before a decision was yet made – giving it a more contractual quality. In the case of pleas of *nolo contendere*, which were used with some frequency, especially in regulatory cases, conditions might be specified explicitly for the grant of leniency. Much less complicated and almost always conditionless was the guilty plea bargain which emerged in criminal cases – especially those of larceny and assault.

At a time when the Bar was under challenge to allow any man legitimately hired by a litigant to argue a case in court (a proposal ultimately defeated), the simplicity of plea bargaining and absence of arcane legal formalities had popular appeal. Pressures to popularize the practice of law were strong enough that the legislature, in 1835–48, removed control over admission to the bar from the Suffolk County Bar Association to its own agents. In response, the Bar Association disbanded in 1836 and did not reconstitute itself formally until after mid-century, although it continued to function as an extremely powerful private club with strong personal ties and networks of connection to the bench. Though informally constituted, the bar maintained its influence over the legal profession and the law in both old and new ways. In addition to personal and family ties, along with lingering fondness for mentors of apprenticeship days, lawyers met at the Social Law library in Boston and for meals, and, most importantly, “frequently corresponded with Supreme Judicial Court justices such as Lemuel Shaw, requesting copies of legal opinions and commenting on points of law” (Jones *et al*, 1993, p 38).

Since many lawyers were also involved in politics, we have every reason to believe that judges’ awareness of Federalist and, later, Whig policy objectives was high. Specific reforms, such as the 1841 request for a change in chancery rules, were sometimes proposed (Jones *et al*,

1993, p 32). Such exchange of ideas dated back to at least 1814 when the court considered complaints from lawyers that the circuit delayed and, thus, denied justice in Suffolk County (Jones *et al.*, 1993, p 33). A permanent court for the county was soon established.

Plea bargaining appears to have been espoused by old political and social elites whose electoral power was under siege because of the continued control it gave them, in a broad sense, through judicial discretion over sentencing policy. In Boston, during the early decades of the 19th century, virtually the entire Bar consisted of former Federalists – now Whigs. While committed to a “republican” vision, their outlook, like that of many of Boston’s elite Brahmin families, was a different, more forward looking variant of Jefferson’s bucolic trust in the freeman farmer that accepted manufacturing, commerce and industrialization as the inevitable path of change. Thus, a fundamental commonality of political vision, along with strong social and institutional ties, created both affinity and frequent mingling among the members of the bar, the judiciary and Boston’s elite privileged inner circle.

Defendants, largely lower class persons in the lower court, accepted the practice because it held out a sense of leniency, the appearance of control over one’s fate through negotiation, and the elimination of intrusive state oversight of the lives of defendants through what had been the increasingly frequent practice of leaving cases “open” on file. By offering leniency, closure of cases and some control over the outcome the Whigs hoped to draw conflicts into the courts before they could escalate in other realms – thus promoting the stability needed for growth. The opportunity to gain experience making decisions among structured options in a public setting was also used as a process to educate the masses in both a conception of citizenship and a sense of the responsibilities it involved.

Plea bargaining emerged as a significant phenomenon during the 1830s and by 1840 the practice of granting concessions in cases where such a plea had been entered was set in place and continued into the 20th century (Vogel, 1999). Plea bargaining did not emerge as a full blown plan or scheme. Instead it was the product of gradual incremental improvisation by a Whig political elite seeking to bolster social order so vital to the healthy functioning of markets and to economic development and, with it, their own flagging political fortunes. Though informal, plea bargaining responded to various criticisms of criminal justice afoot in that day – namely claims that 18th century justice had been too expensive, hard to understand and slow. With its simplicity, rationality and regularity, plea bargaining offered a routinization that was appealing. In its exchange, albeit symbolic, the practice drew the attention of not only defendants but also the public to the precise costs of criminal acts in a way favoured by consequentialists. Popular court vignettes appeared in city newspapers and communicated both case

outcomes and often a moral lesson to readers on a daily basis.

THE SOCIAL DYNAMICS OF CONSENT

Besides providing advantages to city officials, to Boston’s social elite, and to defendants, plea bargaining, once established, held out specific advantages for judges, prosecutors and defence attorneys. While not presented here as causes of the rise of plea bargaining, these advantages explain its “acceptance” by the court. For judges, the practice provided a rejoinder to criticisms of court discretion and reliance on precedent. While the codification movement, which had sought to restrict judicial discretion by moving to legislative statutes, had failed, the threat posed by its underlying sentiment remained. For judges, plea bargaining offered a new, more conciliatory, customary means of maintaining discretion – yet in a depersonalized, knowable and relatively predictable market-like form that was more palatable to the masses.

Justices in the lower courts, whose salaries were annually appropriated, also had reason to believe, rightly or not, that they faced subtle pressure for consonance with the policies of governor and legislature because of initiatives proposed in the legislature to examine the performance of judges individually during the appropriation process – an abortive attempt at political review of the judiciary. Politically motivated court reorganizations, that turned out all sitting judges and appointed new ones, had also historically been common. While judges were appointed by the governor for life subject to good behaviour, the early decades of the 19th century repeatedly saw court reorganizations, at both federal and state levels, motivated at least in part by politics, in which entire benches of sitting judges were turned out and new ones appointed.

In this context, plea bargaining provided a low profile and implicit form of discretion that facilitated sentencing consistent with prevailing policies and purposes of punishment.

In addition to Whig influence through judicial appointments and social policy, there existed by 1840 a tradition of judges, justices of the peace and district attorneys who had careers that mixed judicial and political life. Eventually, after 1858, district attorneys were elected and prosecutors were linked to politics directly. This heightened the value of the discretion that plea bargaining accorded judges and prosecutors in cases that could colour their political prospects. This is not to say that judges and prosecutors crafted positions with an eye to political gain. Reliance on plea bargaining, however, did accord them latitude in high profile situations of consequence. This connection between judges and prosecutors, on the one hand, and elected office, on the other, is not one that existed in England.

Plea bargaining also had other bureaucratic consequences that served prosecutors and defence

attorneys as well. Cases in the lower court were usually expeditiously handled by a judge alone with public prosecutors rarely involved before 1850. While district attorneys were salaried and so had no financial interest in case outcomes, the 1830s saw the legislature first require annual reports detailing court caseloads and dispositions. This appears to have been part of the court reform movement to establish impersonal and regularized justice. Such rationalized reporting meant that a process which inherently produced a high conviction rate grew desirable as the century wore on and public prosecutors handled more cases in the lower courts. While such required reports contributed to growing emphasis on efficiency and rational criteria of performance, their effect was limited in the lower courts where cases were typically handled without attorneys for either defence or prosecution. Perhaps most salient, bargaining provided a daily power resource for the prosecutor.

For defence attorneys, criminal cases were not particularly lucrative, and so they stood to lose little in fees as a result of expeditious bargaining. Though attorneys often defended serious criminal cases, most lower court cases, before mid-century, were resolved without defence counsel so that attorneys lost virtually nothing at all. When defence attorneys did appear, plea bargaining enhanced their discretion as it did that of the prosecutor. Bargaining, thus, closely safeguarded the prerogatives of judges, and to the extent that they gradually came to serve in the lower courts, of prosecutors and defence attorneys too. Because plea bargaining served each actor well, it was variously embraced or tolerated, rather than opposed, within the courthouse.

Interestingly, while delay was a constant criticism in the higher courts, all signs are that cases moved quickly through the lower court – almost always reaching trial before a judge within one day in the early part of the century (Gil, 1837). This challenges the popular view that caseload pressure in the courts may have given rise to plea bargaining. The fact that caseload increased steadily over the last half of the 19th century, while concessions attendant to bargaining fluctuated, further challenges the power of caseload as a cause.

SOCIAL CLASSIFICATION AND SORTING

What was taking place was a form of social classification and sorting. By relying on character witnesses, employment histories, family ties and criminal records, the court was clearly attempting to identify those who were hardworking family people who had simply made one misstep in an otherwise worthy life and who could be reclaimed as productive workers and responsible citizens. These the judges differentiated from marginals, transients and those with few ties who were more often sentenced to serve time at the house of industry or, later, the house of correction. In selecting judges for the police court, great care was paid

to their educative capacities across a broad range of social backgrounds.

In contrast to present day policy in the United States, where imprisonment is growing more and more widespread, the courts used the plea bargain initially almost like a form of symbolic surety – re-embedding the defendant, who is granted leniency, back into his or her world of work and family amidst that most powerful, for most persons, of all forms of social control – the web of watchful relationships of everyday life. Those who had testified symbolically staked their reputation on the defendant's future good behaviour and, thus, had an interest in seeing the commitment kept.

LEGITIMATING DEMOCRATIC INSTITUTIONS

As the American republic set off into the 19th century, self rule was an extraordinary experiment and, though optimistic, no one knew whether it would survive. One dilemma was that, as in any democracy, it became particularly difficult for law to enforce order by coercion because the very need to use force challenged and undercut the claim of the regime to represent the “popular sovereignty” of the citizen – the will of all. Thus, it became vital for the regime to win the consent of the people to its rule. In the new polity, governed no longer by monarchy but rather by self-rule, there was a realization that social order must rest, not on power or coercion, but instead on popular acceptance of sovereign commands and on a sense among the people of a duty to obey. A vision of political authority was needed for a world of popular rule.

What emerged was, at first appearance, a conception of modern political authority rooted in the “rule of law”. Yet, despite the basic modernity of the new republic, that authority, in practice, came to comprise a unique blend of rational-legal and traditional elements. The world's newest republic turned to some of mankind's oldest and “tried and true” approaches to bolster authority amidst the extraordinary adventure of self rule.

BLENDED FORM OF AUTHORITY: RATIONAL-LEGAL AND TRADITIONAL ELEMENTS

The result of this invocation of the rule of law and concomitant turn to episodic leniency was a move in the 19th century United States toward a vision of political authority that is, at first appearance, of a rational-legal sort – that is, authority whose legitimation is based in the enactment of rules in law and the specification of offices in law.

Yet, as judges focused on the problem of creating political authority anew, their reliance on tools of discretionary leniency, in addition to the powerful discourse of a “rule of law”, was quite shrewd. Despite the basic post-Enlightenment modernity of the new

“republic”, they were crafting authority of a unique blend of modern rational-legal and traditional elements. There was a strong sense that ideology unsupported by the stabilizing influence of participation in an integrative network of social roles could prove a fragile basis on which to build social order. The Jacobin excesses in France had been vividly seared into the collective American political imagination and rioting in Britain prior to the Reform Act of 1832 was well noted too.

Local politicians sensed that, along with laws, the subjects of political authority that is solely rule-based require the normative guidance that comes from a secure place in the web and routines of social structure. During the excesses of the 1790s following the French Revolution, they believed those who had taken to the streets did so because they had come detached from their places in the habitual role structures and related rounds of activity of everyday life. Because religion, social consensus and the deference accorded status were all eroding apace, the established view was that hierarchy and a sense of social position had to be sustained lest the American masses be turned out as radically unconstrained individuals whose penchant for violence and excess might equal the French. If political stability could be had in America, the interconnected networks of social roles that honeycomb society (ie, family, community and, especially, work) must, it was believed, play a vital part.

New uses of leniency in the courts and the “symbolic suretyship” that I have been describing were well-suited to accomplish that embeddedness as defendants were turned back to the community under the watchful eye of their intercessors. In doing so, the plea bargain and other new forms of leniency linked the rule-based authority of the courts with the traditional authority that had historically served as a cornerstone of the socializing web of communal membership of everyday life.

Thus, the model of authority that emerged was a unique mix of modern rational-legal and traditional authority. The former based its legitimation, as noted a moment ago, in the enactment of its rules and specification of its offices in law which empowered a “rule of law”. The latter, which relied on customary hierarchies and relationships (eg, parents, spouses, employers, ministers, elders) legitimated its claims on the behaviour of citizens through enduring regard for the sacredness of custom which endowed the relational web of membership of everyday life with a formidable capacity for social control. Amidst open turmoil of crime, riot and unrest, this was how the political leadership of the day faced the, if anything, deeper problem of translating the dreams of the framers of the constitution, in practical terms, into an enduring social and political order capable of wise political action. In that one man’s authority is another man’s hegemony, dissent nonetheless would continue to flourish. ☯

BIBLIOGRAPHY

- Alschuler, Alan (1979) “Plea Bargaining and Its History”, *Law and Society Review*, 13, 211–245.
- Bendix, Reinhard (1964) *Nation Building and Citizenship*. New York: Wiley.
- Chitty, J. (1816) *Criminal Law*. Philadelphia, PA: Isaac Riley.
- Council of Massachusetts Temperance Society (1834). Report.
- Forbath, Willaim E. (1991) *Law and the Shaping of the American Labor Movement*. Cambridge, MA: Harvard University Press.
- Gurr, Ted (1981) “Historical Trends in Violent Crime: A Critical Review of the Evidence,” in M. Tonry and N. Morris, eds., *Crime and Justice: An Annual Review of Research*, Vol. 3. Chicago, IL: University of Chicago Press.
- Handlin, Oscar (1979) *Boston’s Immigrants*. Cambridge, MA: Harvard University Press.
- Hay, Douglas et al. (1975) *Albion’s Fatal Tree: Crime and Society in Eighteenth Century England*. New York: Pantheon.
- Horwitz, Morton J. (1977) *The Transformation of American Law, 1780–1860*. Cambridge, MA: Harvard University Press.
- House Report, Commonwealth of Massachusetts Legislature, Number 4, January 1845.
- Jones, D. L. et al. (1993) *Discovering the Public Interest*. Canoga Park, CA: CCA Publications.
- Lane, Roger (1971) *Policing the City: Boston, 1822–1885*. New York: Atheneum.
- Langbein, John (1973) “The Origins of Public Prosecution at Common Law”, *The American Journal of Legal History*, 17, 313.
- Langbein, John H. (1978) “The Criminal Trial Before the Lawyers”, *The University of Chicago Law Review*, 45, 263–316.
- Lockridge, Kenneth A. (1981) *Settlement and Unsettlement in Early America: The Crisis of Political Legitimacy before the Revolution*. New York: Cambridge University Press.
- Mensch, Elizabeth (1982) “The History of Mainstream Legal Thought” in David Kairys, ed., *The Politics of Law*. New York: Pantheon.
- Moley, Raymond (1929) *Politics and Criminal Prosecution*. New York: Minton, Balch.
- Monkkonen, Eric H. (1996) Remarks presented at the Social Science History Association annual meetings in New Orleans, LA (October).
- Nedelsky, Jennifer (1990) *Private Property and the Limits of American Constitutionalism: The Madisonian Framework and Its Legacy*. Chicago: University of Chicago Press.

- Nelson, William E. (1981) *Dispute and Conflict Resolution in Plymouth County, Massachusetts, 1725–1825*. Chapel Hill, NC: University of North Carolina Press.
- Novak, William (1996) *The People's Welfare*. Chapel Hill: University of North Carolina Press.
- Poulantzas, Nicos A. (1975) *Political Power and Social Classes*. Atlantic Highlands, NJ: Humanities Press.
- Quincy, Josiah (1822) Speech on July 4, 1822, Manuscript Copy, Statehouse Archives, Commonwealth of Massachusetts.
- Sellers, Charles (1992) *The Market Revolution: Jacksonian America, 1815–1846*. New York: Oxford University Press.
- Skowronek, Stephen (1982) *Building a New American State*. New York: Cambridge University Press.
- Smith, Adam (1778) *The Wealth of Nations*. Lausanne: Societe Typographique.
- Thompson, E. P. (1975) *Whigs and Hunters: The Origins of the Black Act*. New York: Pantheon.
- Vogel, Mary E. (1988) *Courts of Trade: Social Conflict and the Emergence of Plea Bargaining in Boston, Massachusetts, 1830–1890*. Doctoral Dissertation, Harvard University.
- Vogel, Mary E. (1999) “The Social Origins of Plea Bargaining: Conflict and the Law in the Process of State Formation”, *Law and Society Review*, 33, 161.
- Vogel, Mary E. (2001) “Lawyering in an Age of Popular Politics: Plea Bargaining, Legal Practice and the Structure of the Boston Bar, 1800–1860” in Jerry Van Hoy, ed. (2001) *Legal Professions: Work, Structure and Organization*. New York: Elsevier.
- Warren, Charles (1931) *Jacobin and Junto*. Cambridge, MA: Harvard University Press.
- Weber, Max (1978) *Economy and Society*. Ed. G. Roth & C. Wittich. Berkeley, CA: University of California Press.
- Wiebe, Robert (1966) *The Search for Order*. New York: Hill and Wang.
- Wood, Gordon (1992) *The Radicalism of the American Revolution*. New York: Vintage.
- The author wishes to thank *Law and Society Review* and Elsevier Publishing for permission to include excerpts of material that previously appeared in Vol 33 (Spring 1999) and in “Lawyering in an Age of Popular Politics: Plea Bargaining, Legal Practice and the Structure of the Boston Bar, 1800–1860” in Jerry Van Hoy, ed (2001) *Legal Professions: Work, Structure and Organization*, London, Elsevier, respectively.

Dr Mary E Vogel

Visiting Fellow, 2001–2002 and then Associate Senior Research Fellow, Institute for Advanced Legal Studies, School for Advanced Study, University of London. Lecturer in Socio-Legal Studies and Criminology, University of Leicester. Dr Vogel received her Ph D from Harvard University and has taught at the University of Michigan, Ann Arbor, and the University of California at Santa Barbara. Her articles on plea bargaining received the Law and Society Association's Award for Best Article in 1999 and the Article Prize for 1999–2000 of the Law Section of the American Sociological Association. Mary Vogel was named a Bunting Fellow at Harvard University in 1992–1993. She is the author of *Coercion to Compromise: Plea Bargaining, the Courts and the Making of Political Authority* which is being published by Oxford University Press and forthcoming in 2005.