Wesleyan Economic Working Papers

http://repec.wesleyan.edu/
Nº: 2018-005

Plea Bargaining in South Africa: An Economic Perspective

Richard Adelstein

September 2018

WESLEYAN UNIVERSITY

Department of Economics
Public Affairs Center
238 Church Street
Middletown, CT 06459-007

Tel: (860) 685-2340
Fax: (860) 685-2301
http://www.wesleyan.edu/econ

Plea Bargaining in South Africa:

An Economic Perspective

Richard Adelstein

Woodhouse/Sysco Professor of Economics

Wesleyan University

Middletown, CT 06459 USA

radelstein@wesleyan.edu

September 2018

Prepared for the *Constitutional Court Review*

PLEA BARGAINING IN SOUTH AFRICA:

AN ECONOMIC PERSPECTIVE

I TWO TRACKS TO AGREEMENT

A *negotiated guilty plea*, or *plea bargain*, takes place when a defendant agrees to forego his right to a full criminal trial by pleading guilty to some offense in exchange for a favorable concession by the prosecution or, less frequently, the court. Where prosecutors exercise substantial discretion in the selection of charges, the range of possible concessions is wide, particularly where the crime is complex or involves several perpetrators. To secure a guilty plea in the circumstances of the case, prosecutors might, among other things, agree to reduce or drop the charges faced by the defendant, a co-defendant or some third party, protect the defendant from imprisonment or other onerous sanctions, limit the evidence or facts that might be led or revealed to the court, or stay their hand in some other way that might induce, or pressure, defendants to surrender their right to trial.[1] But in practice, in the many and varied forms of plea bargaining that have emerged in recent decades in both adversarial and inquisitorial systems around the world,[2] the concession ultimately sought by almost every defendant is a reduction in the fine or prison sentence he will receive upon conviction, punishments whose relative severity can meaningfully be represented as objective numbers. In these cases, a defendant pleads guilty to some offense in exchange for a measurably more lenient sentence than he expects to receive after conviction at trial. Famously, for the better part of a century plea bargains of this kind have accounted for at least 90 percent of all criminal convictions in the state and federal

---

[1] Compare E Steyn 'Plea-bargaining in South Africa: Current Concerns and Future Prospects' (2007) 20 *South African Journal of Criminal Justice* 206, 210-211.

[2] See generally JI Turner *Plea Bargaining Across Borders: Criminal Procedure* (2009) and the papers collected in JS Hodgson ed. *Comparative Perspectives on Plea Bargaining: A Conference for Young Scholar*s (2013), available at http://www2.warwick.ac.uk/fac/soc/law/research/centres/cjc/conferences/outandabout/adversarial/youngscholars.

courts of the United States,[3] and to focus attention on the essential aspects of the exchange, I assume here that every plea bargain takes this straightforward form.

Guilty pleas have been negotiated in South Africa for several decades.[4] After the passage of the Criminal Procedure Act 51 of 1977, what came to be called 'informal' plea bargains appear to have been routinely concluded under s 112 of the Act, which governs guilty pleas at summary trials. Section 112(1)(a) provides that, where a defendant pleads guilty to an offense that does not merit imprisonment or a large fine, and the prosecutor has accepted the plea, the presiding judge may convict the defendant of the offense to which he has pleaded and impose any lawful sentence, without the taking of any evidence beyond the guilty plea itself. For more serious offenses involving imprisonment or large fines, s 112(1)(b) still allows the judge to convict and sentence a defendant without trial, provided that she questions the defendant in court to establish a factual basis for the plea, an admission by the accused of facts sufficient to prove the charge to which he has pleaded guilty. Section 112(2) provides for the submission of a written statement of admission in lieu of the oral colloquy, though the judge remains free to question the defendant in court in addition to receiving the statement. Nothing in the section explicitly contemplates a plea bargain, and while the prosecutor may recommend a specific sentence to the court, the judge retains full discretion either to reject the plea under s 113 and force the parties to a full trial or accept the guilty plea, convict the defendant, and impose any lawful sentence for the offense to which the plea has been entered. The defendant does not hear the sentence until the plea has been entered and the factual basis established, at which point he may no longer withdraw the plea and demand a full trial.[5]

---

[3] See, eg, P Bekker 'Plea Bargaining in the United States of America and South Africa' (1996) 29 *Comparative and International Law Journal of Southern Africa* 168, 168-172; AW Alschuler 'Plea Bargaining and Its History' (1979) 79 *Columbia Law Review* 1, 26-32.

[4] D van Zyl Smit & NM Isakow 'The Decision on How to Plead: A Study of Plea Negotiations in Supreme Court Criminal Matters' (1986) 10 *South African Journal of Criminal Law and Criminology* 3; CT Clarke 'Message in a Bottle for Unknowing Defenders: Strategic Plea Negotiations Persist in South African Criminal Courts' (1999) 32 *Comparative and International Law Journal of Southern Africa* 141.

[5] Criminal Procedure Act 51 of 1977, ss 112 and 113; W de Villiers 'Plea and Sentence Agreements in Terms of Section 105A of the Criminal Procedure Act: A Step Forward?' (2004) 37 *De Jure* 244, 253; MB Rodgers 'The Development and Operation of Negotiated Justice in the South African Criminal Justice System' (2010) 23 *South African Journal of Criminal Justice* 239, 255-256.

Despite its apparent indifference to the practice, once the final procedural piece is in place, s 112 becomes, as Uijs, AJ, put it in 1999, 'virtually tailor-made' for plea bargains.[6] This was provided in 1985 by *State v Ngubane*, in which the defendant was charged with murder and, after discussions with the prosecutor, pleaded guilty under s 112 to the lesser offense of culpable homicide. The presiding judge nonetheless convicted Ngubane of murder with extenuating circumstances. On appeal the court overturned the conviction, reading ss 112 and 113 together to conclude that once the defendant has entered a guilty plea that the prosecutor agrees to accept, the prosecutor's acceptance

> limits the ambit of the *lis* between the state and the accused in accordance with the accused's plea. . . . That the *lis* is restricted appears from ss 112 and 113. The proceedings under the former are restricted to the offence "to which he has pleaded guilty" and the latter must be read within that frame.[7]

This means that, irrespective of what the defendant has actually done and what charge might accurately describe it, once the prosecutor and the defendant have agreed on the charge to which the latter will plead guilty, and the defendant admits facts sufficient to support a conviction on that charge, the presiding judge is bound either to reject the plea or sentence the defendant in accord with the statutorily authorised punishments for the offense to which he has pleaded guilty. Even if Ngubane had actually committed murder, that is, once the prosecutor agreed to accept his plea to culpable homicide and Ngubane admitted 'facts' sufficient to support this lesser charge, however true or false those admissions might be, he could no longer be sentenced to more than the maximum sentence for culpable homicide on his plea. In this way, the prosecutor can control the maximum sentence to which the defendant will be liable by agreeing to accept a guilty plea for an offense less serious than what he believes the defendant actually did. If the difference between the penalty for murder and that for culpable homicide is, say, fifteen years in prison, a prosecutor can offer a defendant he

---

[6] *North Western Dense Concrete CC v Director of Public Prosecutions (Western Cape)* 1999 (2) SACR 669 (C) at 677c.

[7] *State v Ngubane* 1985 3 SA 677 (A) at 683. Cf. Clarke (note 4 above) at 164-165.

thinks is actually guilty of murder the chance to plead guilty to culpable homicide. Save me the expense and uncertainty of a trial for murder, he says to the defendant, and I'll save you fifteen years.[8]

Given the heavy caseloads faced by criminal courts in South Africa,[9] it's not surprising that bargains of this sort quickly became common. In *North Western Dense Concrete CC v Director of Public Prosecutions (Western Cape)*, a case to which we will return, the high court in 1999 bowed to the inevitable, not just acknowledging the pervasiveness of plea bargaining in South African courts but, where the DPP had reinstituted charges against the applicant despite a plea agreement providing that they would not be brought, enforcing the terms of the bargain and ordering the charges dismissed. Though he was not 'filled with joy' to say so, Judge Uijs recognised that plea bargains had become an 'entrenched, accepted and acceptable part' of South African law, taking place daily at every level of criminal justice. He analogised them to civil settlements, which also adjust the *lis* to the needs of the parties and whose validity is unquestioned, and suggested that the entire system of criminal justice might break down if the courts refused to enforce individual bargains or struck down the practice altogether.[10]

Despite the high court's resigned legitimisation of the practice, uncertainty as to the legality of plea bargains remained. On its face, s 112 is about guilty pleas, not negotiations and agreements over sentencing, and it enables the bargaining system itself to remain underground, hidden from public view; indeed, had one of the parties not broken the agreement in *North Western Dense Concrete*, the case and the bargain at its core would never have been subject to judicial scrutiny at all.[11] Soon after this decision, as part of a larger inquiry into South African criminal procedure, the South African Law Commission (as it was then called) issued a report recommending that the Criminal Procedure

---

[8] Compare Clarke (note 4 above) at 160-162.
[9] Steyn (note 1 above) at 207; de Villiers (note 5 above) at 245; Clarke (note 4 above) at 142.
[10] *North Western Dense Concrete CC* (note 6 above) at 683f, 674e, 676f and 678c.
[11] Rodgers (note 5 above) at 242-243.

Act be amended to provide explicitly for plea bargains, exchanges of guilty pleas for leniency in sentencing, so they could be regularised and regulated, and helpfully offered a statute embodying its proposals for legislators to consider. With some modifications, this recommendation was promptly followed, and the Act amended by a new s 105A, governing plea and sentence agreements.[12]

Section 105A is a lengthy statute, and regulates its subject in minute detail, ostensibly to protect defendants from overzealous prosecutors who might deprive them of their right to a full trial. In many respects, it's a model statute. It allows only prosecutors who are authorised by the National Director of Public Prosecutions to do so to negotiate agreements with defendants, who must be legally represented. Before entering into an agreement, the prosecutor must consult both the investigating officer in the case and the complainant, advising the latter on the possibility of compensation as part of the agreement. Any agreement on plea and sentence must be reached before the defendant is asked to plead, and the specific terms of the agreement, the charge to which the defendant will plead, the facts to be admitted by the defendant in support of the plea, and the sentence to be imposed by the court after the plea, must all be made explicit in a detailed written document, signed by both parties, that affirms that the defendant's agreement was given voluntarily and in full knowledge of the rights that would be surrendered with the guilty plea. The agreement is then put before the presiding judge, who may hear evidence and take statements from the defendant or others before passing sentence. Once all these procedural requirements have been met, the judge, nominally still in full possession of the authority to sentence, considers the sentence proposed in the agreement. If she finds the sentence to be 'just,' she informs both sides of this finding, convicts the defendant on the guilty plea, and imposes the sentence provided in the agreement. If she finds the sentence 'unjust,' either too harsh or too lenient, she must inform the two sides of what sentence she would consider just, and offer them the chance to revise the bargain on these new terms. If either side objects

[12] South African Law Commission Fourth Interim Report, Project 73, 'Simplification of Criminal Procedure (Sentence Agreements)' (2001); Rodgers (note 5 above) at 241; de Villiers (note 5 above) at 244; Criminal Procedure Second Amendment Act 62 of 2001, s 105A.

to the judge's terms, they may withdraw from the agreement, after which the defendant is entitled to a new trial before a different presiding judge.[13]

In a perfect world, as discussed below, s 105A might operate to guarantee all defendants a procedure for the inducement of guilty pleas that ensures that every plea is made fully voluntarily, in the defendant's own interests, and with full knowledge of the sentence to be imposed upon the plea and the specific consequences that flow from the surrender of the right to trial. But it's easy to see how the imperfect world of criminal justice might frustrate this purpose. For one, s 105A applies only to defendants able to afford legal representation, or to whom representation is given. Sections 35(2)(c) and 35(3)(f-g) of the Constitution of the Republic of South Africa, 1996, guarantee legal representation at state expense to defendants who are detained and those being tried 'if substantial injustice would otherwise result.' But fulfilling this promise even for these defendants has proven difficult, and it is not clear that it extends to defendants not under detention in the period between arrest and trial, when most plea bargaining takes place. Since 1996, South Africa has struggled to provide representation to poor defendants in the face of very high rates of crime, a lack of trained attorneys, and inadequate funding and poor administration of the state's legal aid services.[14] The great majority of defendants are thus unable to avail themselves of the protections of the formal procedure, and must rely, as they always have, on the informal procedure of s 112. On the other side, the time- and labor-consuming requirements of formal authorisation from supervisors and extensive consultation with victims and the police often deter harried prosecutors facing unmanageable caseloads from employing the s 105A procedure when s 112 proceedings will do equally well. While there is evidence that the s 105A procedures are being employed in some courts, the effect of these conditions, as Steyn observes, has been to confine bargaining in the great majority of cases to the informal, shadowy

[13] Criminal Procedure Second Amendment Act 62 of 2001, s 105A; de Villiers (note 5 above) at 245-250; Rodgers (note 5 above) at 244-255.

[14] CJ Ogletree Jr 'The Challenge of Providing "Legal Representation" in the United States, South Africa, and China' (2001) 7 *Washington University Journal of Law & Policy* 47, 49-54.

realm of s 112 rather than allowing them to flow to the formal, rule-governed domain of s 105A, frustrating precisely the reforms the new law was designed to implement.[15] The same pressures of time and material cost that propelled the replacement of costly trials by cheaper plea bargains in the first place continue to bear heavily on South African criminal justice, driving the substitution of cheaper forms of plea bargaining for more costly ones.

The result is a two-tiered system of plea bargaining in South Africa, two alternative tracks to negotiated sentences on which passengers travel first- or second-class. The first-class track is governed by s 105A, with its close supervision of prosecutors by administrative superiors, its attention to the position of the police and the interests of victims, and the care it takes to ensure both that defendants are fully informed of the precise terms of the agreement and the rights they will surrender by pleading guilty and that their consent to the agreement is given freely. But travel on this track is very expensive, not just for defendants, few of whom can afford the legal representation that s 105A makes the ticket for first-class travel, but for the state as well, which must make good the prosecutorial time and effort required to certify a formal bargain. So the great majority of South African defendants must travel second-class, on the informal track governed by s 112, where prosecutors operate with far less oversight than on the first-class track, the information defendants have at the moment they must decide how to plead is significantly less complete and reliable, judges retain a freer hand in sentencing in cases where an agreement has been reached, and pleas remain binding even where the information on which a defendant has relied turns out to be faulty.

My purpose here is to examine the operation of this system by considering it in light of some very simple economic logic. If there is any legal institution that might usefully be viewed from an economic perspective, one might think, it would be plea bargaining, the forthright exchange of valu-

[15] M Watney 'Judicial Scrutiny of Plea and Sentence Agreements' (2006) 2006 *Journal of South African Law* 224, 224-225; Steyn (note 1 above) at 215-217. On the use of s 105A procedures, see also M Kerscher *Plea Bargaining in South Africa and Germany* (2013), unpublished LL.M. thesis, Faculty of Law, Stellenbosch University at 10-11.

able concessions between parties who each stand to gain from the trade. I hope to suggest more precisely how this might be true, but not by assaulting the problem with complex theory or mounds of statistical data. Instead, I propose a simple, intuitive, empirically plausible depiction, a *model*, of how plea bargaining works that allows us to see more clearly how the behavior of prosecutors, defendants and judges and the existence of various conditions in the bargaining environment affect the system's outcomes, the individual agreements that result in guilty pleas. It identifies the factors that move both prosecutors and defendants to seek bargains, shows how these factors act on their decisions, and examines the problems bargainers face when the information upon which they must base their decisions is unreliable or unavailable. And it illuminates the differences between formal plea bargains under s 105A and informal bargains under s 112, showing how imperfections in the latter environment can lead to dysfunctional outcomes that are less likely to occur in the case of s 105A bargains. To the extent it succeeds at all this, the model may be a valuable conceptual tool for South African scholars and policymakers, a fruitful way to understand plea bargaining and the problems it solves for the criminal process, and the new problems it creates at the same time.

But it is important to be clear as to what analysis of this sort can and cannot do. Though it does enable us to ask more, and more salient, questions about how the plea bargaining system operates, how to improve its outcomes, whether on balance it is a good or a bad thing, and what its constitutional implications might be, it cannot answer many of these questions at all, and can answer others only in a conditional or qualified way that leaves ample room for disagreement on political, moral or jurisprudential grounds. Economics, like every social science, has both a *positive* and a *normative* dimension, questions about *what is* and questions about *what ought to be*. The former suggests that economics is a science, a way to understand how the social world actually works, how relationships between individuals, groups and institutions are in fact conducted, and how social outcomes that we can see are produced or created. In practicing positive economics, economists are like physicists, trying to provide answers to questions about how things really are in terms that all rational

people are bound to accept, just as they're bound to accept that the earth is round, whether they like it or not. But physicists are typically not concerned with the rights of electrons or asteroids, or whether it's a good thing that force and mass are related as they are, or whether we should try to change the laws of thermodynamics because they lead to unjust results. All physics is positive science; it does not ask, or answer, normative questions.

The subject matter of economics, in contrast, is living people and their relations to one another, and in this realm of perpetual political and moral controversy, normative concerns, ancient, ultimately unresolvable questions about what justice demands, whether the way things are is the way they ought to be, and what the object of public policy, or social life itself, should be, irrepressibly jostle with positive questions for our attention. Positive economics can inform normative debate, by showing how and with what consequences various normative positions and the policies that flow from them might play out in practice, but positive economics alone can never answer or determine normative questions. How plea bargaining works, how its outcomes are produced, and how various conditions affect its operation are positive questions, ones about which economic reasoning may have much of value to say. What it would mean to make the plea bargaining system work 'better,' or whether it should exist at all, are normative questions, and in addressing them, economists have no claim to greater wisdom or expertise than anyone else. But if it cannot answer them, economic reasoning may serve to sharpen these questions, to show just what is at stake and how normative or constitutional concerns can be identified and addressed, and with what consequences. I hope here to show the value of an economic perspective on plea bargaining in both these contexts.

In part II, I start from the assumption that plea bargaining is a good thing, a desirable, or at least uncontroversial institutional innovation that raises no normative or constitutional issues. This makes it meaningful, after sketching how the system actually works, to ask how to make it work better, a question sharply posed by the two-track South African system. With the help of a general overview of the plea bargaining system and a simple algebraic depiction of the defendant's problem in

deciding whether to plead guilty or insist on a trial, I argue that bargaining under s 112 is fraught with problems stemming from uncertainty. Because defendants do not, or do not always, have accurate or reliable information about the factors that most strongly influence their decision to plead guilty or insist on a trial, the plea bargaining system is likely to fail in a specific way: plea bargains that would have been made had defendants had the information they needed will not be made, and plea bargains that would not have been made had defendants had this information will be made, because they don't. 'Making the system work better' thus means facilitating the former type of bargain and impeding the latter, a problem of institutional design and performance. If a criminal process commits itself, reluctantly or enthusiastically, to plea bargaining as a legitimate means of resolving criminal cases, it ought to prefer a bargaining system that works well at identifying and completing genuinely consensual agreements to one that works poorly. I argue that the procedure of s 105A, cumbersome as it is, addresses most of the informational problems that arise under s 112, so that, cost considerations aside, the system's operation would be improved – more bargains that should take place, and fewer that should not, will take place – if all South African plea bargains were formal, governed by s 105A.

The superiority of s 105A bargaining to its informal counterpart on various grounds has been widely observed, and while the courts have as yet had nothing to say on the question, several scholars have noted that this disparity raises a serious problem under s 9 of the Constitution, which guarantees equality of rights: the majority of defendants, who are poor, cannot take advantage of the formal protections and guarantees of s 105A because they, unlike wealthier defendants, generally have no legal representation.[16] In the final part of the essay, like these scholars, I relax the assumption of

[16] See, eg, de Villiers (note 5 above) at 253-255; Steyn (note 1 above) at 217-218; Rodgers (note 5 above) at 261; A Botman, *An Evaluation of the Benefit of Plea and Sentence Agreements to the Unrepresented Accused* (2016), unpublished LL.M. thesis, Faculty of Law, University of the Western Cape at 82-91 and *passim*. Rather different jurisprudential objections to plea bargaining are discussed by M Bennun 'Negotiated Pleas: Policy and Purposes' (2007) 20 *South African Journal of Criminal Justice* 17, some of which are considered below.

part II and ask the unanswerable, whether plea bargaining is in fact a good thing. But I raise a different normative problem than these predecessors, a second potential constitutional objection that applies to all plea bargaining, formal and informal alike: that the system, even (or especially) when it is working well, often makes the exercise of the constitutional right to trial a very costly enterprise for the defendant. In a series of cases culminating in the Supreme Court's vindication of the bargaining system through the 1970s, the American federal courts explicitly asked whether 'putting a price' on the exercise of the right to trial, as plea bargaining was alleged to do, amounted to a violation of that right, and concluded that, generally, it did not. I use the model of part II to clarify when and how a plea bargain can be said to put a price on the right to trial, and briefly suggest how the issue of making the exercise of the right to trial costly might be framed in terms of the Constitution. I hope in both these ways to advance the important emerging debate over plea bargaining in South Africa.

II MAKING PLEA BARGAINING WORK BETTER

A How Plea Bargaining Works

Challenging Judge Uijs's analogy of plea bargaining to civil settlements in *North Western Dense Concrete*, Mervyn Bennun called the differences between the two so sharp that 'the comparison verges on the cynical.' But it was, he went on, 'useful' in pointing out the resemblance of plea bargaining to 'the purchase of a rug at a Lebanese bazaar.'

> To move away from the outcome of a criminal trial as being unpleasantly redolent of the marketplace, a plea and sentence agreement negotiated under s 105A should result in the same verdict and sentence which would have followed had the matter gone to a full trial, and the court should satisfy itself that this is the case. . . . [A] negotiated agreement under s 105A [surely] should propose the sentence which would have been imposed after a full trial.[17]

Perhaps Bennun, a sharp critic of the practice, proposed this interpretation knowing that it would, if adopted, largely eliminate plea bargains under s 105A and render them available only under s 112 where, he thought, leniency might still be justified if the presiding judge were satisfied that the

[17] Bennun (note 16 above) at 29.

defendant's plea showed genuine remorse.[18] For its effect would surely be precisely that, to make s 105A plea bargains pointless, and thus make them disappear.[19]

The crux of every consensual plea bargain is the sentencing discount given in exchange for the guilty plea. As elaborated below, if there is no difference between the sentence the defendant would receive after conviction at trial and the sentence proposed in the agreement, there is little reason for even a defendant with a very slim chance of acquittal to accept an agreement, convict himself without putting the prosecution to the proof, and surrender his right to take the small chance that he'll somehow not be convicted. Even if the plea is motivated, as some are, not by fear of greater punishment after trial but by the adverse consequences of bringing the facts of the matter to light in a public trial, s 105A's requirement that the facts be stipulated in writing in the agreement largely removes this advantage of guilty pleas.

As Bennun suggests, the exchange of sentencing discounts for guilty pleas isn't pretty, but if the system of plea bargaining is to work at all, this is the way it has to work. Everywhere it has taken hold, plea bargaining has been an institutional response to the needs of the state, not the interests of defendants. When prosecutors and administrative judges face caseloads so heavy that only a small sliver of cases can possibly be afforded a full trial given the resources at hand, an environment in which the American criminal process has labored since at least 1900[20] and, as Steyn suggests,[21] has characterised South Africa for at least several decades as well, prosecutors have little choice but to somehow persuade large numbers of defendants to surrender their right to a trial and agree to convict themselves, obviating the need to mount a trial in those cases and allowing the state, as the provider of criminal justice, to put its extremely scarce prosecutorial and judicial resources to better use

---

[18] *Ibid.*

[19] Compare Watney (note 15 above), at 226.

[20] See, eg, LM Friedman *Crime and Punishment in American History* (1993) at 235-323; JH Langbein 'On the Myth of Written Constitutions: The Disappearance of Criminal Jury Trial' (1992) 15 *Harvard Journal of Law and Public Policy* 119; Alschuler (note 3 above).

[21] Steyn (note 1 above) at 207-208, 218.

elsewhere. Absent the public expenditure necessary to provide a full trial for every defendant, the alternative, as Judge Uijs recognised, is the collapse of the system itself. Because so few defendants are so remorseful or so unwilling to endure the publicity of a trial to agree to plead guilty without some material inducement, something they want and deem sufficient to compensate for the loss of rights and the possibility of acquittal their plea entails, prosecutors must offer something in return for the guilty pleas they need to keep the criminal process afloat. In plea bargaining systems everywhere, sentencing discounts are the only answer.

One way bargains might be achieved, *sentence bargaining*, would have the defendant or his legal representative bargain directly with the sentencer, the trial judge. But concerns about judicial neutrality and dignity in a process most observers (and participants) see as demeaning keep judges in almost every adversarial system, including South Africa and the United States, from intervening directly in plea negotiations. The alternative, *charge bargaining*, is almost universally employed because it is so well adapted to a peculiar feature of the adversary system, one that prosecutors facing severely overcrowded dockets exploit to the fullest. In inquisitorial systems, like those of continental Europe, prosecutors are generally bound by a *rule of compulsory prosecution*, which requires them to try every case on the highest possible charge the evidence will support, nothing more or less, so as to produce a judgment that, to the extent possible, accurately reflects the actual facts and moral weight of the case. If a defendant, like Ngubane, is actually guilty of murder, the prosecutor must charge him specifically with and try him for murder.

But subject to a provision for private prosecution, at common law and under s 6 of the Criminal Procedure Act South African prosecutors have almost complete discretion to decide whether or not to prosecute any defendant and, if so, for what crimes.[22] At the same time, the substantive criminal law establishes a large repertory of crimes covering a broad range of behaviors, most modified by degrees, many overlapping or governing activities typically undertaken as part of a single criminal

[22] Criminal Procedure Act 51 of 1977, s 6; Steyn (note 1 above) at 206-207.

transaction.[23] This gives prosecutors a great deal of sentencing power. Subject to whatever constitutional or administrative prohibitions of vindictive overcharging might exist, a prosecutor can typically charge a suspect with several independent crimes committed in the same course of action, exposing him to a potential punishment much greater than that prescribed for the principal crime in the case. Most importantly, as in *Ngubane*, since there's no rule of compulsory prosecution, the charge the prosecutor selects need not accurately describe what she believes the suspect has actually done. So she can charge a suspect with any crime or combination of crimes on the list he *has* committed, or with some lesser offense he *hasn't* actually committed, that is required to force the judge to sentence the defendant to the punishment the prosecutor has promised in the bargain.

Plea bargaining in this form is made *possible* by two aspects of adversarial procedure that are absent in inquisitorial systems. First, defendants have the right to plead guilty and abort trials before they occur. Section 35 of the Constitution guarantees every accused the right to a fair trial, but it does not command anyone to exercise it. Defendants may have reasons other than, or in addition to, an offer of leniency to avoid a trial, and pleading guilty allows them to do it. A guilty plea is a conviction, an agreement by the defendant to submit to punishment for the offense to which he has pleaded guilty, and it ends the criminal case. Typically, the only consequence of the plea to the defendant is the sentence to which he has agreed. Inquisitorial defendants cannot plead guilty or abort their trials, because the point of inquisitorial trials is to provide an accurate account of what actually happened and establish an appropriate punishment for what the perpetrator has done. But adversarial trials, adorned as they are with complex rules of evidence and procedural rights, like the right against self-incrimination, designed to protect defendants from the consequences of the truth, are not primarily about discovering the truth of the matter. They are, ideally, about finding guilty defendants criminally liable for something that allows appropriate punishments to be imposed on

[23] See, eg, G Kemp (ed) *Criminal Law in South Africa*, 2nd edition (2015) *passim*.

them. So the right to plead guilty, the power to subject the prosecutor to a costly trial or make it unnecessary, becomes a valuable bargaining chip for defendants in negotiations. Defendants have something to give that prosecutors want.

The second enabling feature of the adversarial system is that prosecutors, because of their broad charging discretion, have something of value to give them in return. By selecting charges and combining them with sentencing recommendations that, at least in the United States, they know judges, who operate under the same pressures they do, will respect, prosecutors can calibrate punishments as finely as necessary to win guilty pleas from defendants with penalties they consider in each case to be acceptable discounts for the defendant's saving the government the cost and uncertainty of trying them. Save us this much money, the prosecutor says, and we'll save you this much time. Again, this requires that defendants, in the last analysis, be the ones who bear the state's costs of convicting them at trial, in the form of heavier punishments upon conviction at trial, so as to preserve their incentive to bargain. Plea bargaining is only possible where defendants who do insist on a full trial are consistently sentenced upon conviction with more severe punishments than they would have been had they accepted a plea agreement, so prosecutors can make surrendering the right to the trial attractive to all defendants. Only then will a proffered sentencing discount be seen by defendants as credible.[24]

The institutions of adversarial procedure make plea bargaining possible. What makes it *necessary* in the systems that have come to rely on it is a combination of two factors common to most of them: the sheer volume of crime and correspondingly large caseloads these chronically underfunded systems must confront, and the elaboration of the trial process itself, the complexity of many modern statutes and the difficulties of proof they pose and the evolution of procedural safeguards for defendants and rules of evidence that increase the length, cost and uncertainty of full trials. The proportion of convictions in a system that result from plea bargains, that famous 90 percent in the

[24] Compare R Adelstein *The Exchange Order: Property and Liability as an Economic System* (2017) at 206-214.

United States, and the extent of the sentencing discounts that must be offered to secure the needed rate of guilty pleas, are measures of the pressure exerted by these factors on the system. In South Africa, where trial procedures are simpler than in the United States but the incidence of crime is comparable, and like the United States, prosecutors and courts do not have nearly the resources to meet the demands placed on them by the volume of crime, these indicators are harder to determine. But if Judge Uijs is right, and South African criminal justice, like its American counterpart, could not withstand the abolition of plea bargaining, the practice must be very widespread, and the discounts correspondingly deep.

B The Defendant's Problem

The prosecutor may control the charge and, to a large extent, the sentence, but because there can be no bargain without his consent, if guilty pleas are entered knowingly and voluntarily, the defendant controls the plea bargain. Prosecutors for whom inducing a guilty plea in a given case is not especially urgent, perhaps because they have some special reason to take the case to trial, or because the volume of crime generally does not press too heavily on their resources, may offer only low prices for a plea, small sentence discounts relative to the sentence the defendant can expect to receive after conviction at trial. Where, as in the United States, the need to induce pleas is acute because there are far too many cases to try with the available resources, prosecutors may offer very high prices for guilty pleas, deep discounts that make the bargained sentence a mere fraction of the expected sentence after a trial. But whether the proffered discount is high or low, it is the defendant who must agree to accept it. Only he can make the trial unnecessary.

So it's useful to examine the defendant's problem in a plea bargain – should he accept the sentence proposed by the prosecutor, or retain his right to trial by refusing the bargain and pleading not guilty – more closely. On the one hand, the bargain represents, we'll assume for now, a sentence the defendant can count on, that he knows for certain. On the other, going to trial presents the prospect

of two possible outcomes, an acquittal, which for simplicity we'll represent as a punishment of zero, and a conviction, which brings a punishment that, again for now, we'll assume the defendant knows, along with some perceived likelihood of one or the other. Suppose, for example, I'm accused of aggravated robbery, for which, if I'm convicted at trial, I'm certain to receive a ten-year sentence. My defense is that I didn't do it, which may or may not be true. But there's evidence against me: my gun was used in the robbery, and my fingerprints were found on it afterward. All things considered, my attorney tells me, the chance of my being convicted at trial is 70 percent. So the *expected punishment* I face in going to trial, defined as the sentence I'd receive on conviction multiplied by the likelihood that I'll in fact be convicted, is $0.7 \times 10 = 7$ years. This doesn't mean that I'll be sentenced to seven years, since the only possible outcomes of my trial are ten years or none. It means that if, hypothetically, my trial could be conducted identically a million times before a million different impartial judges, and my attorney is right that 70 percent of these trials would result in a punishment of ten years and 30 percent in a punishment of zero, the 'average' punishment I'd receive in the million trials would be seven years per trial. Going to trial is thus like being subjected to a punishment lottery that, 'on average,' costs me seven years per trial.

Suppose the prosecutor tells me that if I plead guilty to simple robbery, he'll guarantee a sentence of seven years. Then one could say that the 'objective values' of the plea bargain and the punishment lottery are the same: take the deal and get seven years for certain, or enter the lottery and expect to pay seven years for an 'average' trial. Whether I take the deal or not depends on how I feel about the risk posed by the lottery. If, for whatever reason, I'm especially terrified by the thought of going to prison, I might look at the lottery pessimistically and focus on the awful prospect of ten years behind bars, so that seven years for certain seems a better outcome for me now than the prospect of serving even more if I'm convicted at trial. If so, I'm *risk averse*, and will take the deal. Because the 70 percent chance of a sentence as long as ten years seems worse to me than a certain sen-

tence of just seven, I'd even be willing to accept an even lower price for my plea, a bargained sentence a little *greater* than seven years, just to be relieved of the possibility of serving ten. But if prison doesn't hold all that much terror for me, perhaps because I've been there before and think of it as part of criminal life, or because I can pursue valuable relationships inside (or outside) the prison while I'm there, I'll enter the punishment lottery by refusing the bargain and going to trial. Because I look optimistically at the trial as a chance to walk free rather than pessimistically as a chance to serve ten years, I *prefer the risk*. To get me to relinquish my chance for an acquittal, despite the odds against it, the prosecutor will have to offer a deeper discount, a higher price, for my guilty plea, a certain sentence of *less* than seven years.[25]

It is thus not enough to depict the defendant as concerned solely with choosing the option that yields the smallest possible 'objective' sentence, in which case every defendant faced with this situation would accept a bargain at a day under seven years or less and refuse it at a day over seven years or more. Different defendants value a year in prison differently. A year in prison, or each year of a sentence of more than a year, imposes a *cost*, a quantum of distress or suffering, on defendants that differs from defendant to defendant and, for any given defendant, from year to year, differences that produce different attitudes toward the risks of the punishment lottery. So the relevant question for any defendant contemplating a prosecutor's offer is not whether, given the likelihood of conviction at trial, the sentence the prosecutor offers, *B*, is smaller than the expected sentence at the trial, the trial sentence *T* multiplied by the likelihood of conviction *p*, that is, whether $B < p \times T$. It's whether the *cost* imposed on the defendant by the bargained sentence *B* is less than the *expected cost* of going to trial and risking the sentence *T*, his 'average cost per trial.' If $C(X)$, the unique, individual *cost function* that characterises each defendant and distinguishes him from the others, is a mathematical representation of how a given objective punishment *X* is transformed into the defendant's subjective experience of distress or suffering at being forced to endure *X*, and we assume, reasonably,

[25] *Ibid.* at 216-217.

that *C* rises with *X*, that is, more punishment is worse than less, then the question posed by the plea offer is whether

$$C(B) < p \times C(T).$$

If so, the defendant will agree to plead guilty and accept the sentence *B*. If not, he will reject the agreement and enter the punishment lottery.

This formulation suggests three distinct ways that a prosecutor can induce a defendant to plead guilty, that is, make the left side of the inequality smaller, or the right side larger, for that defendant. Other things equal, the trial alternative can be made less attractive relative to the bargain by increasing either *T*, the sentence on conviction at trial, or *p*, the likelihood the defendant will be convicted, or both. Or the bargain can be made more attractive relative to the trial by lowering *B*, the sentence the defendant will serve if he pleads guilty. These in turn correspond to the three principal bargaining tactics employed by prosecutors in adversarial systems everywhere, with each raising the question of constitutional propriety in a different way. Most commonly, prosecutors attempt to induce guilty pleas simply by deepening the sentencing discount ('sweetening the deal'), paying high prices for pleas by offering bargains at low values of *B*. This is the form taken by the archetypical plea bargain, so deciding on the propriety of this tactic is tantamount to deciding on the propriety of the practice itself. But even if this propriety is conceded in general, if prosecutors are so desperate for pleas that they offer extremely low values of *B* relative to *T*, say a three-month sentence for a weapons offense for a defendant facing the possibility of ten years for aggravated robbery after a trial, the very sweetness of the deal might induce even a truly innocent defendant who thinks there's even a small chance that he will be convicted at trial to plead guilty, whether he prefers risk or is averse to it.[26]

[26] R Adelstein 'The Negotiated Guilty Plea: A Framework for Analysis' 53 *New York University Law Review* (1978) 783, 820-823.

Alternatively, prosecutors can increase the likelihood of conviction by devoting more resources to investigation or, as is often done in cases involving more than one defendant, securing the testimony of one defendant against another by offering the defector a chance to plead guilty to a lesser offense in exchange for significantly increasing the likelihood that his co-defendant will be convicted at trial, and thus helping persuade the co-defendant to plead as well. It is precisely this situation that economists describe as the *prisoner's dilemma*. Given the propriety of bargaining itself, this common tactic creates no problem of convicting the innocent. This is because in any adversarial system, the only authoritative institution that exists for discovering, or pronouncing, anyone guilty or not guilty is the criminal process itself. If the evidence is strong enough to convince a judge or jury beyond a reasonable doubt that a defendant is guilty of a crime, he *is* guilty, and subject to punishment, even if he really didn't do it. It is the conviction that creates the legal fact of guilt, irrespective of what the defendant or anyone else may believe.[27] Subjecting defendants to the prisoner's dilemma thus cannot be unfair to the co-defendants who become more likely to be convicted, because finding defendants guilty or not as a legal fact, not finding authoritatively exactly who did or did not do what, is what adversary trials are about, and increasing the likelihood that a judge or jury will find a defendant guilty is what prosecutors are supposed to do.

The third strategy, increasing *T*, is more problematic. In the United States, this is typically achieved by 'overcharging,' piling a host of distinct charges on a defendant who has committed several offenses in the course of a single criminal transaction. In *Brady v United States* (1970),[28] in which the Supreme Court first gave its approval to plea bargaining in general and the first two tactics in particular, the Court stressed that the defendant's participation in a plea bargain must be truly volun-

[27] Compare *Herrera v Collins* 560 US 390 (1993) (petitioner convicted of capital murder is not entitled to *habeas corpus* relief simply because, years after his conviction, new evidence tending to prove his innocence is discovered.)
[28] *Brady v United States* 397 US 742, 750 (1970); Adelstein (note 26 above) at 823-827.

tary, and hinted that egregious overcharging might so 'overbear the will of the defendant' as to effectively coerce him into pleading guilty to lesser charges. But eight years later, in *Bordenkircher v Hayes*,[29] the Court was forced to face the harsh realities of American plea bargaining, conducted in an environment of high crime, inadequate resources to prosecute it and, by international standards, very severe prison sentences for almost every serious offense. Hayes, who had two prior felony convictions, was charged with passing a bad cheque for less than $90, an offense punishable by imprisonment of two to ten years. The prosecutor offered Hayes a five-year sentence for his guilty plea, and to pressure him to accept, told him that if Hayes were convicted at trial on the cheque charge, he would invoke the state's habitual offender act, which provided for an automatic sentence of life imprisonment after a third felony conviction. Hayes, fully aware of this threat, nonetheless refused the bargain, was convicted at trial and, when the prosecutor invoked the habitual offender act as threatened, duly sentenced to life imprisonment.

But because it had to in order to preserve the plea bargaining system and, with it, the entire American criminal process, the Court held that the threat and eventual invocation of the habitual offender act was an acceptable prosecutorial bargaining tactic. If an injustice was done to Hayes in this case, it suggested, it was not the prosecutor's threat that did it, but the fact that the state's habitual offender act, which was not challenged in the case and which all sides agreed Hayes had violated, permitted a life sentence in these circumstances. If plea bargaining is to work, defendants must know that the sentencing discount is real. The only way any defendant can be induced to plead guilty in exchange for a proffered discount is to turn those, like Hayes, who do insist on a trial into admonitory examples for the rest, by following through on the threat to sentence them more harshly when they are convicted than they would have been had they pleaded guilty. Prosecutors who need to induce pleas, the Court reluctantly recognised, must be permitted to credibly threaten recalcitrant defendants in this way and, as in *Bordenkircher*, to use existing statutes that defendants have in fact

[29] *Bordenkircher v Hayes* 434 US 357 (1978); Adelstein (note 24 above) at 227-229.

violated to do so.

> Indeed, acceptance of the basic legitimacy of plea bargaining necessarily implies rejection of any notion that a guilty plea is involuntary in the constitutional sense simply because it is the end result of the bargaining process . . . . [By] tolerating and encouraging the negotiation of pleas this Court has necessarily accepted as constitutionally legitimate the simple reality that the prosecutor's interest at the bargaining table is to persuade the defendant to forego his right to plead not guilty.[30]

C Information and Uncertainty

Defendants, and only they, know their own cost functions and the attitudes toward punishment these reflect. But in order to decide how to plead, a defendant also needs information about three variables, *B*, *p* and *T*. These are 'objective' in the sense that they can be expressed numerically, but the information the defendant has about them may be imperfect, forcing him to make largely subjective estimates on the basis of the information he has about what the true values of the objective variables are. In many situations, defendants do have about as much of the information as it is possible for them to have. If the prosecutor is clear as to what value of *B* he is offering and the defendant knows both that the prosecutor won't renege on the offer in court and that the sentencing judge will respect the agreement and impose the sentence *B*; if the defendant can rely on a well-informed guess (all that is possible in any case), formed either on the basis of his own experience or the advice of counsel, as to the likelihood that he will be convicted; and if he can be confident as to what sentence he would receive if he were convicted at trial, then the decision the defendant makes as to plea can be said to be fully informed, and absent coercion, a question we'll soon revisit, a reflection of his best interests as he sees them in a painful predicament. But if not, apart from whatever normative questions of fairness or constitutionality might arise from allowing defendants to plead guilty in the absence of reliable information about *B*, *p* and *T*, where defendants must in fact decide whether to plead guilty in substantial ignorance of these variables, the plea bargaining system itself is likely to perform poorly in the role the criminal process has assigned to it.

---

[30] *Bordenkircher v Hayes* (note 29 above) at 361-364.

How do defendants acquire this information and assess its reliability? Some, as suggested above, will be able to rely on personal experience. Having encountered the criminal process before, they may be aware of the sentencing practices of trial judges, have a good sense of what the likelihood is that they will be convicted on the available evidence, and perhaps even be able to judge whether the prosecutor is offering them the best deal they can get or if a bit of resistance might result in a better one. But whether they are experienced or not, personal knowledge, from whatever source it might be gleaned, and the representations and promises of the prosecutor are all the information that most defendants in s 112 proceedings have. For any defendant, the best source of reliable information during the bargaining process about all the relevant variables, and the person best positioned to wrangle the best bargain possible from the prosecutor, is a competent defense attorney, and this is a luxury most s 112 defendants do not enjoy. Without an attorney, defendants, typically at a moment filled with stress and anxiety, must evaluate not just the likelihood of conviction at trial and the sentence that will be imposed thereafter, but the precise terms and legal consequences of the proffered discount and its value relative to the trial sentence, by themselves. Compared to competently represented defendants, this puts them at a substantial disadvantage in identifying and concluding agreements that actually serve their interests. And should they err in estimating the key variables and agree to a bargain they come to regret, once entered and accepted by the judge, the deal becomes final, just as it does for represented defendants. As Msimang, J, noted in refusing a request to reopen a s 105A bargain after the defendant learned he'd agreed to less favorable deal than he thought he had, in *State v. Armugga*:

> In the process of bargaining, numerous assumptions are made and mistakes are bound to happen. Provided that a party is found to have acted freely and voluntarily, in his or her sound and sober senses and without having been unduly influenced when conducting a plea bargaining agreement, the fact that the assumptions turn out to be false, does not entitle such a party to resile from the agreement.[31]

[31] *State v Armugga* 2005 2 SACR 259 (N) at 265c-d.

In addition to an attorney to advise them about *T* and *p*, s 105A defendants also have the advantage of having the agreement in all its particulars reduced to writing, so that the true value of the prosecutor's offer *B*, the magnitude of the discount it represents and the full legal consequences it entails, are made as clear as possible to them. If, having agreed to plead guilty in return for the sentence specified in the agreement, the defendant is surprised at the judge's refusal to impose it in court, he may withdraw the plea and demand a full trial. His s 112 counterpart, in contrast, represented or not, will have no opportunity to withdraw his plea even if, having agreed to an informal deal with the prosecutor and pleaded guilty in reliance on it, he finds the judge unwilling to honor the agreement and determined to impose a sentence greater than *B*. Even where the information the defendant needs exists in reliable form, where, that is, prosecutors offer clear terms with no intention of reneging and judges are prepared to sentence as the agreement specifies, the procedures of s 105A are clearly superior to those of s 112 in providing defendants with the information they need to ensure that their decisions represent, as far as practicable in a challenging environment, their own interests as they see them. Where information exists but is unreliable, not only compounded difficulties for individual defendants faced with decisions, whether they have legal representation or not, but serious dysfunction in the plea bargaining system itself are the likely results.

Because every criminal case is different, and because at the moment defendants must plead, the trial lies in an unknowable future, it is hard to see how an experienced defense attorney's estimate of *p*, the likelihood of a particular defendant's conviction, can be improved by institutional design. And the punishments in the statute books themselves, supplemented perhaps by informal knowledge of the sentencing habits of particular judges, are generally sufficient to give defendants a useful predictor of *T*. But misinformation, not simply error, about the variable *B*, the precise value of the sentencing discount the defendant is being offered for his guilty plea, is often the result of specific actions, intentional and unintentional, by prosecutors and judges that can effectively be po-

liced and remedied by statute and judicial oversight, and which left unaddressed will erode the system's ability to identify and facilitate mutually beneficial bargains.

Recall that, given *B*, *p* and *T*, a defendant will plead guilty if the subjective cost to him of the bargained sentence is less than the expected cost of going to trial, that is, $C(B) < p \times C(T)$. Suppose, however, the possibility exists either that the prosecutor will renege on his offer when it comes time for him to perform, say by making a different sentencing recommendation to the court than that provided in the agreement, or that the judge, who ultimately remains in possession of the sentencing power, will not abide by the terms of the agreement and sentence the defendant more harshly than the agreement calls for. In either case, if the defendant thinks there's a chance that, having been promised a sentence *B*, he will actually receive a more severe punishment and not be able to extract himself from the bargain, he will have to replace *B* in his calculations with the higher sentence $(B + h)$, where *h* represents the expected increase in the sentence created by the possibility that the prosecutor will renege on his promise or the judge will impose a sentence greater than *B*. If there are enough broken promises or unexpected judicial nullifications of bargains so that *h* is significant, then whenever

$$C(B) < p \times C(T) < C(B + h),$$

that is, whenever the cost to the defendant of bargain at *B* is smaller than the expected cost of a trial, but the cost of a bargain at $(B + h)$ is greater than the expected cost of the trial, a plea bargain that would have served not just the interests of the immediate parties to it but the larger interest of the criminal process in diverting trials to less costly procedures as well, and that would have been made without this uncertainty, won't be made because of it. The existence of even a few unkept bargains will be enough to surround every subsequent negotiation with an uncertainty that keeps at least some otherwise mutually beneficial bargains from being concluded. Expensive trials that no one wants will have to be mounted, and the already burdensome costs of criminal justice will rise still

further. Just as bad money drives good money from circulation, and for the same reason, bad bargains will drive out good ones, to everyone's detriment.[32]

In 1971, in *Santobello v. New York*,[33] the United States Supreme Court addressed the first of these sources of misinformation, the failure of the prosecutor to keep the promises made in the agreement. Santobello, charged with two gambling offenses, agreed to plead guilty to one in return for the prosecutor's promise not to recommend a prison sentence. But between the time the agreement was concluded and Santobello was to be sentenced, a new prosecutor had taken over the case and, unaware of the promise his predecessor had made, recommended that Santobello be sentenced to a year in prison on his guilty plea, and the judge complied. Santobello protested, but was not allowed to withdraw his plea. On appeal, the Court, speaking through Burger, CJ, found for Santobello. It began by explicitly acknowledging the inability of the American criminal process to survive without plea bargaining, something the Court in *Brady* had refused to do the previous year, grounding its approval of bargaining instead on the 'mutuality of advantage' it offers to both prosecution and defense.[34] Plea bargaining, the Court now said, is 'an essential component of the administration of justice. Properly administered, it is to be encouraged.' But proper administration requires 'fairness in securing agreement,' and in ensuring it the Court applied a basic principle of ordinary contract law to hold that, despite the fact that the breach of the agreement was inadvertent and the one-year sentence itself not unjustified, Santobello must receive specific performance of the agreement or be permitted to withdraw his plea. '[W]hen a plea rests in any significant degree on a promise or agreement of the prosecutor, so that it can be said to be part of the inducement or consideration, such promise must be fulfilled.'[35]

---

[32] Adelstein (note 26 above) at 814-816.

[33] *Santobello v New York* 404 US 257 (1971).

[34] *Ibid.* at 260; *Brady v United States* (note 28 above) at 752.

[35] *Santobello v New York* (note 33 above) at 260, 261, 262. Note that reaching the 'fair' outcome in this case also had the effect of 'making plea bargaining work better.' This confluence of fairness and efficient operation in the bargaining system is discussed further in R Adelstein 'Institutional Function and Evolution in the Criminal Process' (1981) 76 *Northwestern University Law Review* 1, 71-79.

For decades before 1970, the Supreme Court turned a blind eye to the unpleasant business of plea bargaining until, its hand forced by some careless language in a 1968 opinion not involving a plea bargain,[36] it discussed the procedure in detail and began regulating it in *Brady* and *Santobello*. In South Africa too, plea bargaining under s 112, increasingly acknowledged by scholars and practitioners,[37] remained in the judicial shadows until 1999, when *North Western Dense Concrete CC v Director of Public Prosecutions* was decided by the high court on facts much like those in *Santobello*. The applicant and its employee were charged with culpable homicide, and the prosecutor agreed to drop the charge against the applicant if the employee pleaded guilty, which he did in a s 112 proceeding. But after another party had petitioned the respondent for a *nolle prosequi* in the case, the DPP reconsidered its position and reinstituted the charge against the applicant, in violation of the earlier plea agreement.[38]

Judge Uijs's opinion closely parallels those of the Supreme Court in *Brady* and *Santobello*. He began by acknowledging the long history of informal plea bargaining under s 112 in South Africa, a provision 'tailor-made' for plea bargains, and lamented the judiciary's unwillingness to confront it openly.[39] The court conceded, as the Supreme Court had in *Santobello*, that bargaining under s 112 was an integral component of South African criminal justice, one so entrenched that eliminating it was likely to cause the breakdown of the entire system[40] and, like the *Santobello* Court, appealed to ordinary contract law to draw the necessary conclusion. Because it would be unfair to allow the respondent to enjoy the benefits of the agreement without meeting the obligations it created for him, the court granted a permanent stay of prosecution to the applicants, effectively ordering specific performance of the contract. It held that a guilty plea under s 112, if accepted by the prosecutor, binds

---

[36] *United States v Jackson* 390 US 570 (1968). See Adelstein (note 24 above) at 218-220.
[37] Compare Bekker (note 3 above) at 218; Steyn (note 1 above) at 206; Kerscher (note 15 above) at 30.
[38] *North Western Dense Concrete CC* (note 6 above) at 671-672.
[39] *Ibid.* at 672-677.
[40] *Ibid.* at 683f, 676f and 678c.

both the prosecutor and the court on the facts stipulated in the agreement, and requires the prosecutor to honor the terms on which it has been reached.[41] As in *Santobello*, requiring prosecutors to fulfill the agreements on which defendants have relied in pleading guilty serves not just the interest of fairness to defendants, but goes far toward preserving the system's ability to facilitate the completion of agreements that represent the defendant's informed judgment of his own interests in the case.

The second source of uncertainty for defendants is whether the sentencing judge will abide by the agreement and impose the sentence it calls for. If the defendant fears that the sentence he receives will be greater than he bargained for, he will have to add the increment *h* to his assessment of the bargain's terms, reproducing the problem caused by breach of the agreement by the prosecutor. But the independence of sentencing judges makes securing their cooperation in the completion of plea bargains by sentencing in accord with them a delicate matter. One approach, adopted by the American federal courts for plea agreements couched in carefully specified language, is to explicitly withdraw the sentencing power from the judge in these cases and require the bargained sentence to be imposed.[42] Another, more implicit approach employed successfully by most of the states is not to direct the judge to impose the sentence specified in the agreement but to rely instead on administrative judges, whose interest in relieving the enormous pressure on the criminal process is closely aligned with the prosecutor's, to ensure that sentencing judges do what they have to do to keep the system afloat.

South Africa's approach differs across the two tracks to agreement. Because judges retain the full authority to impose any lawful sentence after a guilty plea in a s 112 proceeding, informal bargaining in that context can only be charge bargaining, with the parties' agreement on the facts and the charge the principal constraint on the judge's choice of sentence within the statutory range. If a defendant believes there is a significant likelihood that the judge will not impose the sentence he and

---

[41] *Ibid.* at 670f-g, 677b. See also South African Law Commission Fourth Interim Report (note 12 above) at 3.16.

[42] *Federal Rules of Criminal Procedure*, Rule 11(c)(3).

the prosecutor hope she will, this uncertainty will distort his choice as we have discussed, and his reluctance to bargain will only be increased by his inability to withdraw his plea should he discover after he's entered it that the sentence is greater than he bargained for. But in severely overburdened criminal courts, as in the United States, to the extent that trial judges respond to the pressures that heavy caseloads put on the judiciary and the prosecution alike, they too may tacitly agree to honor the bargains put before them, to keep the cases moving.

Section 105A seems to recognise this problem, and takes steps to alleviate it. The written agreement that manifests the formal bargain must not only 'state fully the terms of the agreement, the substantial facts of the matter, all other facts relevant to the sentence agreement and any admissions made by the accused,' it must also propose a specific punishment that, in the view of the parties, would constitute 'a just sentence to be imposed by the court.'[43] After the judge questions the defendant to determine that he understands the proceedings and the agreement and establish a factual basis for the plea,[44] she then considers the sentence agreed upon by the parties. If she finds it 'just,' she must then convict the defendant and sentence as the agreement requires. If she finds it 'unjust,' she must inform the parties of what alternative sentence she would consider just, and give them an opportunity to revise the agreement on the basis of this sentence and resubmit it to her for sentencing. If either side declines to accept these terms, the plea is withdrawn and the defendant entitled to a new trial *de novo* before a different judge.[45]

Steyn points out that the statute's use of the word *just* to describe an acceptable sentence differs from the more traditional usage in this context, *appropriate*, and suggests that this is meant to signal to judges that they should approach the sentencing of defendants in s 105A proceedings differently than they do the sentencing of defendants convicted after a full trial.[46] Two important cases

[43] Criminal Procedure Act 51 of 1977, ss 105A(2)(b), 105A(1)(a)(ii)(aa).
[44] Criminal Procedure Act 51 of 1977, ss 105A(5-6).
[45] Criminal Procedure Act 51 of 1977, ss 105A(7-9).
[46] Steyn (note 1 above) at 213-215.

support this view. In *State v Sassin*,[47] the high court considered the sentence imposed in a s 105A bargain, and interpreted 'just' to imply not that the sentencing judge is required to agree with the proposed sentence – indeed, Majiedt, J, suggested that, had he been the judge in the case, he would have imposed a longer sentence – but only that she take proper account of the 'triad' of traditional determinants of appropriate sentencing enunciated in *State v Zinn*,[48] the seriousness of the offense, the interests of society and the personal circumstances of defendant.[49] In this way, Watney notes, *Sassin* encourages judges to take account of 'the same factors a sentencing court would normally consider but with the proviso that any sentence that could possibly be considered appropriate under the circumstances would suffice.'[50]

Two years after *Sassin*, in *State v Esterhuizen*,[51] another s105A case addressing the 'justness' of the proposed sentence, the high court revisited the question. The sentencing judge's role, wrote Els, J, was to consider the 'well-known triad,' but also to take

> a broad overview of the facts admitted . . . together with the proposed sentence to be imposed, all with a view to establishing whether the sentence agreed upon and its effective content bear an adequate enough relationship to the crimes committed taking into account all of the agreed facts, both aggravating and mitigating, so that it can be said that justice has been served.[52]

Like Majiedt, J, Judge Els noted that, as a sentencing judge, he 'would probably have imposed a much heavier sentence under the circumstances,' but 'it must be so that the court, in considering the "justness" or "unjustness" of a sentence agreement cannot simply decide for itself *in vacuo* what sentence it would have imposed for crimes to which the accused is pleading guilty.' Even, or perhaps especially, in a case that is difficult to prove and the likelihood of conviction corresponding lower, a plea bargain may well result in a lower sentence than might be imposed after a successful trial, but he insisted that '[t]his does not mean justice has not been achieved.'[53]

---

[47] *State v Sassin* 2003 4 All SA 506 (NC).
[48] *State v Zinn* 1969 2 SA 537 (A).
[49] *State v Sassin* (note 47 above) at 15.5-15.8, 18.1.
[50] Watney (note 15 above) at 227.
[51] *State v Esterhuizen* 2005 1 SACR 490 (T).
[52] *Ibid.* at 495a, b.
[53] *Ibid.* at 494c, 494b, 494h. Compare Watney (note 15 above) at 228-229.

The high court seems to be treading a very narrow path, preserving nominal judicial discretion in sentencing and attention to the traditional factors of sentencing even in the presence of a s 105A agreement, at the same time that it encourages judges to agree to negotiated sentences that may be far less severe than those that would result after a conviction at trial. It seems to be inviting, or asking, as subtly and delicately as it can, sentencing judges to swallow hard and impose whatever sentence the parties may specify in a plea agreement, so long as it bears a reasonable relationship to the crime the defendant has admitted to committing. This, as Judge Els put it, is the 'price' that must be paid for the system's commitment to plea bargaining.[54]

III IS PLEA BARGAINING A GOOD THING?

Plea bargaining, formal or informal, involving defendants with legal representation or without it, remains deeply controversial, even in the United States, for four principal reasons. The first is that it produces inaccurate outcomes, not once in a while, or even inevitably, but *systematically*. Prosecutors have good reasons for allowing offenders to be punished for crimes less serious than what they have actually done. It lets them successfully conclude a larger portion of cases with convictions, and deploy their scarce human and material resources to inflict some punishment, even a small one, on as many offenders as possible, all of which is surely in the interests of the criminal process, if not necessarily justice. But when huge caseloads put their resources under great stress, so the price in punishment they are willing to pay for a guilty plea is very high, defendants plead guilty to crimes that bear only faint resemblance to, and suffer punishments far smaller than the law prescribes for, what they have actually done. And because the essence of the bargain is the sentencing discount, the bias is always in the same direction – *every* bargain results in a sentence smaller than what the defendant could expect after a trial on an accurate charge. When discounts are deep, punishments vastly understate the moral weight of the offenses that were actually committed. In the United States, something

[54] *Ibid.* at 494h.

like this characterises 90 percent of all criminal convictions. When this systematic effect is added to the uncertainty already attached to the law's inability to capture and convict every offender, the expected punishments that prospective offenders actually face for their crimes become very much lower than the moral harm done by their crimes, and to the extent that punishment acts to deter crime, many more crimes will be committed than would be the case were the punishments authorised by statute for the crime actually committed applied in every case.

A second objection is that plea bargains remove the doing of criminal justice from public view. Charges and sentences are negotiated, often hastily and under pressure, by prosecutors and defendants, or by prosecutors and defense attorneys, who must then negotiate with their clients, far from public view. Even in a s 105A bargain, the only evidence ever brought to light is what the parties choose to incorporate into the written agreement and the defendant admits in colloquy with the judge, and the only members of the public permitted even indirect access to the proceedings are the victims, whose views prosecutors are required to solicit.[55] In the United States, moreover, where public defenders do a great deal of the bargaining for impecunious defendants, defendants may be pressured to accept agreements by their attorneys, hoping to maintain cooperative relationships with prosecutors at the expense of the particular client whose fate they are deciding. It's not a pretty sight, even if we could see it.

Which is a third objection. Criminal trials have a kind of majesty, as the full power of the state is deployed against a lone defendant purposely equipped by the law with rights to resist it effectively. Conducted with due regard to these rights, they help make ordinary people into citizens by inviting them to observe how the government treats suspected offenders, what standards it applies to them and itself, and when and how it makes mistakes. Staying its hand against defendants until fair procedures convict them shows the modern state at its best, elevated in its citizens' eyes by the respect it

[55] Compare the similar criticism of Bennun (note 16 above) at 31-41.

shows them and the law. Plea bargains, as Bennun[56] and many others have pointed out, have none of this. Compared to the moral weightiness of a full trial, they seem shabby, even dirty, like dispensing rough-and-ready justice in a dusty bazaar. Judged against the normative and constitutional standards manifested in a modern criminal trial, the undignified, unlovely haggling of plea bargaining is hard to defend on any ground beyond necessity. Not even its defenders like it much.[57]

But the most serious objection to plea bargaining, if it is accurate, is related to the first: that the system itself routinely operates to coerce defendants into pleading guilty. It doesn't put a gun to their heads, but in the United States it does, very often, present them with sentencing discounts so large that even truly innocent defendants who think there is some chance that they will be erroneously convicted at trial will plead guilty to crimes they did not commit. Even in systems that devolve substantial discretionary authority on police and prosecutors to refuse to arrest or release suspects where there is little likelihood of conviction, so that it is unusual for anyone to reach the point of pleading without there being substantial evidence that he is guilty of some crime, it does happen that defendants against whom there is strong evidence are in fact innocent. Cases where there is good but imperfect evidence and the defendant insists, correctly or not, that he didn't do it are the hardest for prosecutors to win, so they may be exactly the ones in which a prosecutor who wants a defendant he thinks guilty to be convicted and suffer some small punishment will offer him a very deep discount to induce a plea to some punishable offense.[58]

Does plea bargaining impose a price on exercising the right to trial and create a danger that innocent defendants will falsely plead guilty? Our simple model of the defendant's problem suggests a

---

[56] *Ibid.* at 29.

[57] Cf Adelstein (note 24 above) at 214-216.

[58] A question addressed by Els, J, in *State v Esterhuizen* (note 51 above) at 494: 'It must . . . be clear that in the give and take of negotiations, an accused person may tender in the negotiation to plead guilty to a charge of which that accused person is guilty, but in respect of which the State may have had considerable difficulty in achieving a conviction. In return for the concession of a plea of guilty to a charge difficult to prove, it must be so that the Legislature has envisaged that the bargaining mechanism would bring home a result which satisfies the interests of justice.'

useful way to approach this question. As noted in part II, apart from serious factual errors or dispositive new evidence revealed after a conviction, the only authoritative institution we have for determining whether any defendant is guilty or not guilty is the criminal process itself. This means that, for any defendant who is offered a plea bargain, the relevant question is not what he actually did, but how likely he is to be convicted at trial of the charge against him. Now recall our earlier example: I'm accused of aggravated robbery, which I may or may not have done. The probability I will convicted at trial is 70 percent, and if I am, I will be sentenced to ten years in prison. So the expected trial sentence is $.7 \times 10 = 7$ years; were my trial to be repeated a million times, I'd be convicted 70 percent of the time and serve ten years each time, and acquitted 30 percent of the time and not be punished at all.

As we saw, if the prosecutor offers me a sentence of seven years in exchange for my guilty plea, the 'objective values' of the bargain and the punishment lottery are the same: accept the bargain and get seven years for certain, or enter the lottery and pay seven years for an 'average' trial. Whether or not I take the deal depends on my attitude toward the risk of the lottery. If I'm risk averse, I'll take the deal, and would continue to do so even if the prosecutor offered a smaller discount than three years, that is, a sentence somewhat greater than seven years; if I prefer risk, I'll refuse the deal, and continue to do so even if he offers bigger discount, a sentence somewhat less than seven years. But in neither case am I forced to 'pay a price' for exercising my right to trial. As far as anyone knows before the trial, the 'objective' punishments in the two alternatives are the same, and whether I take the deal or not depends on my preferences, just as whether I buy an orange for R5 depends on how I feel about that orange and those five rands. Indeed, were the prosecutor offering a lower price for my plea, say a nine-year sentence for it rather than seven, the price might be too low for even a risk averse me to surrender my lottery ticket, and I would take my chances at trial. It is only when the price the prosecutor is willing to pay for the plea, the extent of the sentencing discount, is greater than the objective value of the plea that a price is put on exercising the right to trial.

Suppose the prosecutor offers a sentence of five years. Then the price he is willing to pay for the plea, five years less than the trial sentence, is greater than the three-year discount needed to make the certain punishment of the bargain and the expected punishment in the lottery have equivalent objective value. Accepting the bargain means five years for me, when the expected sentence after trial is seven, so exercising my right to trial means paying an objective price of two years. Whether I surrender my right and accept the bargain or pay the two-year price for exercising it depends, again, on my preferences. If I'm risk averse, I'll certainly take the deal, since I would have accepted a seven-year sentence to avoid this risk but will actually have to pay only five. This might even get a risk-preferring gambler to plead guilty – it's hard to turn down a five-year sentence discount for a lottery ticket that is only worth three. But in either case, refusing the deal means paying an objective price of two years for exercising the right to a trial.

Now, as prosecutors in large American cities must do every day, turn up the heat. Suppose a hard-pressed prosecutor offers me nine months for my plea. This may be an offer I can't refuse, a bargain that, given my predicament, is much too good to pass up no matter how I feel about the risk of trial. In light of the possible outcomes, the price I would now have to pay to exercise my right to trial is very high. If I'm risk averse, the minimum discount I would accept for my right to trial is somewhat less than three years, and the prosecutor is offering me more than nine years to surrender it. That, as an economist, or a defense attorney, might say, is a lot of value to pass up in an uncertain world. Even if I know that I am not guilty, and even if I am a gambler and prison holds less terror for me than it does for risk averters, if there is a 70 percent chance that I will have to serve ten years if I go to trial and I can avoid this risk by serving just nine months, I am very likely to accept the deal and plead guilty. The overwhelming need of American prosecutors to secure guilty pleas has, for at least a hundred years, meant that terms like these are offered to most American defendants, and thus that most are being made to pay a substantial price for exercising their right to trial.[59]

[59] Adelstein (note 24 above) at 216-218.

Despite all this, or because of it, for much of the twentieth century, as plea bargaining sank deep roots in American criminal liability, the Supreme Court simply averted its eyes. It refused to acknowledge the existence of plea bargains or consider their constitutionality, and regulated guilty pleas only by requiring that they be made voluntarily by defendants informed of their con- se-quences. But by 1969, the lower federal courts had been confronting the propriety of plea bargaining within this framework for several years, and in that year, *Scott v United States*[60] was decided in the court of appeals for the District of Columbia. Scott was convicted at trial of robbery and sentenced to an indeterminate sentence of up to fifteen years. At the moment of sentencing, the trial judge re-marked that he thought Scott had deliberately lied during the trial, and said, 'If you had pleaded guilty to this offense, I might have been more lenient with you.' On appeal, Bazelon, J, noted that '[t]he stark import of this comment is that the defendant paid a price for demanding a trial [and] the pricetag thus placed on the right to a fair trial . . . would, on first impression, seem clearly impermis-sible.'[61]

But Judge Bazelon recognized that this was too simple, for the relevant comparison is not be-tween the punishment a defendant receives after conviction at trial and the discounted sentence a bargain might have offered, but between the discounted sentence and the expected sentence at trial. The decision to plead must be made before the results of the trial are known, so the relevant ques-tion is whether, at the moment the plea decision must be made, the promise of leniency puts an im-permissible price on the right to trial. Offering a detailed analysis very much like our own, he charac-terised a proffered sentence equal to the expected sentence after trial as merely reflecting 'the uncer-tainties of litigation' and thus imposing no price on the right to trial, but concluded that discounts that went beyond this did indeed impose such a price for the purpose of deterring the defendant's exercise of his right to trial, and to this extent made the defendant 'a pawn sacrificed to induce other

[60] *Scott v United States* 419 F.2d 264 (D.C. Cir. 1969).
[61] *Ibid.* at 267, 269. Compare the very similar facts in *Regina v Atkinson* 1978 2 All ER 460, 461-463.

defendants to plead guilty.'[62] But even this, he argued, would not necessarily render the guilty plea involuntary or improper. That would require a discount so deep as to create a significant risk that innocent defendants will plead guilty in response, and to enable prosecutors to determine when the discounts they were offering passed this threshold of impermissibility, the courts or the legislature had an obligation to develop standards to guide them.[63]

This is in fact the approach ultimately adopted in England. After 1994, Parliament greatly lessened the need for explicit bargaining itself by establishing a fixed price for guilty pleas in every case: a one-third reduction in the statutorily prescribed sentence if the plea is made at the first reasonable opportunity; a one-quarter reduction once the trial date has been set; and a one-tenth reduction if the plea comes after the trial begins.[64] But in *Brady* and *Bordenkircher*, as we have seen, the Supreme Court created no such guidelines and put few limits on the discounts a prosecutor may offer to secure a guilty plea, even where these are so great as to impose a significant price on a defendant's exercise of the right to trial. And in *Santobello*, the Court made explicit that the reason it was tolerating such a regime was that the alternative, a full criminal trial for every defendant, was far beyond the material means of the criminal process. In effect, it applied a limitations analysis to the right to trial, in the absence of an explicit limitations clause in the American Constitution, and found that it was reasonable, and imperative to the continued operation of the criminal process itself, to limit the right to trial by making it costly for defendants to exercise it.

This suggests the kind of claim that might be made in South Africa to challenge the constitutionality of plea bargaining, formally under s 105A and informally under s 112, in terms of s 35 of the Constitution, and the kind of response to the claim that might be framed under the limitations

---

[62] *Ibid.* at 277, 272

[63] *Ibid.* at 272-278. Cf Adelstein (note 26 above) at 817-823.

[64] UK Sentencing Council 2017 'Reduction in Sentence for a Guilty Plea: Definitive Guideline' at 5, 11. Available at: https://www.sentencingcouncil.org.uk/wp-content/uploads/Reduction-in-Sentence-for-Guilty-plea-Definitive-Guide_FINAL_WEB.pdf

clause, s 36. Making and defending such a constitutional claim would surely require not just imaginative lawyering, but a good deal of empirical research on the environmental and material conditions of criminal justice in South Africa, the incidence of guilty pleas of both kinds in the system, and the sentences being imposed after both guilty pleas and convictions at trial. All of these are tasks for another time. I hope here to have simply opened new avenues of study and debate about a problem of great importance to the law in South Africa, and around the world.