Mathematical Model of Plea, Perjury and Pendency

Why almost no one pleads guilty in India, and what would change it

Team Satyug · Volume 3 of 5 · November 2026

How to read this volume. The paper is in plain prose; the mathematics is in the technical appendix that follows it. Results come from two models, shown side by side: a *foundation model*, which follows each case through the enacted procedure, and a *whole-court simulator*, which adds the feedback between pleas and the court's queue. Every figure is what a model returns under stated assumptions. None is a forecast, and no Indian case-level data have yet been analysed. Readings of the statute are marked as questions for counsel.

Summary

About one conviction in a hundred in India comes through plea bargaining. In England and Wales and in Australia the figure is about nine in ten. This paper asks why, and what would change it.

The answer in brief:

1. **The system is in a loop.** Trials take years, so most accused are released on bail. On bail, waiting is cheap, and with time the case is often settled or the witnesses give up, ending in acquittal. So almost no one pleads guilty, every case needs a trial, and the queue grows. Each step is a reasonable response to the delay, and together they keep the delay in place. The plea law does not break the loop: Chapter XXIII of the BNSS gives an accused thirty days to apply, requires several people to agree, and does not tell him what sentence he will receive.

2. **A Practice Direction can open a working route.** A High Court direction that gives a known sentence on a guilty plea, at the reductions the National Sentencing Guidelines themselves provide (the Rule 14 route), raises completed pleas in the foundation model from about 0.3% to about 24% of eligible cases, which is about a third of convictions for the offences modelled. With trial within six months, the figure is about 36% of cases and half of convictions. If the Rule 14 route is lawful, which is a question for counsel, neither step needs new law.

3. **The full programme passes the international median.** With the Bill's notice, perjury and witness measures added, and trial within six months, the foundation model gives about 43% of eligible cases and 57% of convictions by plea for the offences it covers. The median across 52 jurisdictions is 43% of convictions. If Indian accused respond to discounts as participants in a foreign experiment did, the figures are 40% and 53%. Two different accounts of behaviour arrive at the same place. The six months is for fresh filings on a separate list. At present capacity that is affordable, but the old cases would then wait a decade or more.

4. **The whole-court simulator goes further**, to 74% of cases and 87% of convictions, It differs from the foundation model in several ways at once, among them that it lets pleas shorten the queue. Only a pilot can show which model is nearer the truth.

5. **Three things carry the result, and each has a cost.** The size of the reduction, the wait for trial, and the lawfulness of the Rule 14 route. Large reductions also draw in more accused whose cases are weak (about 11% of them plead, against a target of 5%). Across the assumed ranges the full programme passes the international median in seven runs in ten, but keeps

weak-case pleading within target in only three in ten. So bail must come before the plea, the plea must keep a real cost, and the ceiling on the earliest reduction needs to be settled.

1 The question

India has had plea bargaining on the statute book since 2006, first as Chapter XXIA of the Code of Criminal Procedure, 1973. Since 1 July 2024 it is Chapter XXIII of the Bharatiya Nagarik Suraksha Sanhita, 2023 (BNSS), and it covers most offences punishable with up to seven years. Yet it is hardly used. The National Crime Records Bureau's *Crime in India*, taking Indian Penal Code and special and local law cases together, records:

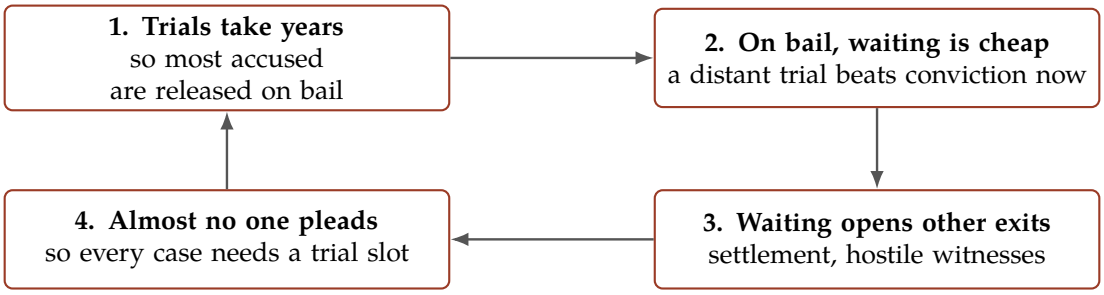
	2022	2023	2024
Cases disposed of by plea bargaining	19,893	42,318	15,735
As a share of convictions	1.01%	1.82%	0.69%
Conviction rate in completed trials	65.7%	66.6%	67.2%

Source: NCRB, Crime in India 2022–2024, Vol. III, Tables 18A.2 and 18A.4, all-India totals; the shares are Team Satyug's own compilation from those tables. The 2023 figure reflects unusually high reporting from Uttar Pradesh and Bihar. NCRB does not count ordinary guilty pleas separately. The 2022 and 2023 figures, and part of 2024, arose under the earlier Code, which had the same structure but no thirty-day limit.

Across 52 jurisdictions the median share of convictions by plea is 43%. The usual explanation for India's backlog is a shortage of judges. This volume takes a different view. It treats the near-absence of pleas as the result of the choices that the present rules reward. An accused who is guilty compares two things: a certain conviction now, and a trial years away that he may well win. If the second is cheaper, he waits. The task is to find which rules make waiting cheaper, and which changes would reverse that without pressing the innocent to plead.

2 The loop

The usual account of India's backlog is a shortage of judges. The account in this volume is that the system has settled into a loop in which every participant responds sensibly to delay, and each response makes a guilty plea less attractive.



1. Delay forces bail. A court cannot hold a person for the better part of a decade awaiting trial. Most accused are therefore released, in non-bailable cases too, and the law itself requires release once an undertrial has served a third or a half of the maximum (s. 479). In the model between one and four accused in ten are in custody when the choice to plead arises.

2. Bail makes waiting cheap. A guilty accused at liberty compares a certain conviction now with a trial years away that he may win. In the foundation model, with a one-third reduction on offer, about 3% of accused on bail plead.

3. Waiting opens the other exits. As the years pass the complainant tires and witnesses turn hostile. The case is compounded, withdrawn or quashed on settlement, in practice even where the offence is not compoundable, and each of these ends in acquittal. In 2022, settlement-type disposals of IPC cases outnumbered plea-bargained ones by about fifteen to one.

4. No pleas, so every case needs a trial. That lengthens the queue, which returns the system to the first step.

Who pleads in such a system. The loop produces the worst combination. Those with means wait and are often acquitted. Those who cannot obtain bail or furnish surety, usually the poorest, are the ones for whom a plea is the price of release: in the model about 36% of accused in custody plead at a one-third reduction, against 3% of those on bail. This is why the two safeguards in this volume point in different directions and are both needed. For the minority in custody, release must come before the plea. For the majority on bail, the plea must be made worth taking, by a known and sufficient reduction, a short wait and the closing of the acquittal exits. With all of these the model's plea rate among accused on bail rises from 3% to about 42%.

Why single measures disappoint. Closing the exits alone sends cases to trial and lengthens the queue. More judges alone shorten the wait slowly while the exits stay open. A reduction alone, at today's wait, moves few people on bail. And the loop does not unwind quickly once pleas begin: with one common queue the model's wait falls from about five years to about two and a half over twelve years. That is the case for a separate list for fresh filings.

The title names three things because they are parts of this one loop. *Pendency* makes delay cheap. *Perjury* that goes unpunished makes acquittal likely. Both make the *plea* a bad bargain for a guilty accused.

3 Why pleas fail today

Inside the loop, the model identifies six reasons why the plea law itself does not work. Each is drawn from the enacted text, and each can be tested separately.

1. Access. The accused must apply within thirty days of the framing of charge (s. 290(1)). The limit is new: the earlier Code had none, and pleas were rare under it too, so the window adds to the problem and does not explain it. Many do not know of the window, do not understand it, or have no lawyer in time. In the foundation model, access is the largest single barrier inside Chapter XXIII: it is the largest in 69% of the simulated draws.

2. Many signatures. Chapter XXIII produces a "mutually satisfactory disposition" worked out among the accused, the prosecutor, the investigating officer and the victim (s. 291). On Team Satyug's reading, which counsel must settle, any one of them can in effect stop the plea. The victim's signature is the second-largest barrier in the model.

3. No known sentence. The reduction on a plea is at the court's discretion within limits (s. 293). An accused cannot tell in advance what he will receive, so he cannot weigh it against the risk of trial.

4. Cheaper exits remain open. A case can also end by compounding, by withdrawal from prosecution, by quashing on settlement, or by acquittal after witnesses turn hostile. These stay open until judgment, long after the thirty-day plea window has closed. The price a victim will accept falls as the years pass. So the accused's best course is to let the plea window lapse and settle later for an acquittal.

5. Lying is free. The summary penalty for false evidence is three months or a small fine (s. 383), and the full route needs a written complaint by the court itself (ss. 215, 379). A witness who turns hostile faces almost no expected penalty, and so does the accused who procures it.

6. Delay. A trial that is years away is discounted heavily. In custody the calculation reverses: an undertrial may plead only to get out. Such a plea depends on custody, whatever the truth of the charge, and Section 8 treats it as a principal risk.

The choice in one example

The arithmetic below is an illustration with round numbers, not a model result. Take an offence for which the court would impose three years after trial. The accused is guilty and on bail.

	Chance of conviction	Wait for trial	Cost of contesting (years, expected)
Today	67%	9 years	0.8
Prompt trial, same evidence	67%	1 year	1.8
Prompt trial, truthful witnesses	95%	1 year	2.6
Plea, one-third off, served now			*2.0*

Expected cost = chance of conviction × three years, reduced by 10% for each year of waiting. Stigma, legal costs and time in custody are left out.

At today's wait, contesting costs the equivalent of ten months and the plea costs two years, so he contests. A prompt trial narrows the gap but does not close it. The plea becomes the cheaper course only when conviction at trial is close to certain. This is why procedure alone moves few accused on bail, and why the pricing of perjury is part of a plea reform and not a separate subject. For an accused in custody the sum is different: every month of waiting is served in jail, so almost any plea that ends custody looks cheaper. That is the danger Section 8 addresses.

4 The argument in one rule

This section stands on its own. It uses neither of the two simulation models described below and carries none of their figures. It sets out, in one rule and one calculation, when an accused person will rationally choose to plead guilty and when that choice reduces the work of the court. It then shows which terms of that rule the Guidelines supply.

The section is an argument from arithmetic. It is not a forecast of how many accused will plead, and every figure in it is an illustration. The statutory readings it relies on are stated as readings.

The choice before the accused

An accused who is told of a plea route compares two costs, each measured in years of imprisonment or its equivalent.

The cost of pleading is the sentence left to serve after set-off, the burden of a conviction record, and any fine or compensation:

$$P = (S_p - s_1)^+ + \rho K + m_p.$$

The cost of contesting is the expense of the trial, any further time in custody before the verdict, and the consequences of conviction, weighted by the chance of conviction p and discounted for the wait of T years:

$$C_{trial} = c_L + x + \delta^T p \left[(S_0 - s_1 - x)^+ + K + m_0\right].$$

Where the law offers another ending, such as compounding, its cost is compared as well. C is the cheapest course open to the accused other than the plea.

The rule. A plea is *attractive* when $P < C$. It is *worth attempting* only when, in addition, the route is reliable enough:

$$\Pi\,(C - P) \;>\; (1 - \Pi)\,\Delta_f + c_n,$$

where Π is the chance that an application results in a lawful disposition being put before the accused, Δ_f is what a failed attempt costs him and c_n is the cost of applying. Appendix A13 gives the short proof and the symbols.

The rule has three consequences. Each can be checked by hand.

1. **The plea terms must be known.** The accused cannot compare P with C unless he is told S_p before he decides.
2. **The route must be reliable.** A generous offer is rationally refused if Π is low. Attractiveness is necessary; it is not sufficient.
3. **Delay works against the plea for an accused on bail.** δ^T falls as the wait grows, so contesting becomes cheaper the slower the court. The reduction needed to make a plea attractive therefore rises with delay and falls with the strength of the evidence.

Why the present law produces almost no pleas

About 1% of convictions come through the Chapter XXIII procedure (between 0.7% and 1.8% in 2022 to 2024: NCRB, *Crime in India*, Tables 18A.2 and 18A.4, IPC and special and local laws). NCRB does not count ordinary guilty pleas separately; they are concentrated in petty fine cases and long-custody cases, which is exactly where the cost of contesting exceeds the sentence.

The rule explains this without supposing that accused persons are unwilling.

- P **is not known in advance.** On an ordinary plea of guilty the sentence lies in the court's discretion. Under Chapter XXIII of the BNSS the court "may" sentence at the fractions in section 293. Neither tells the accused, before he decides, what he will receive.
- Π **is low.** Chapter XXIII requires an application within thirty days of the framing of charge (s. 290(1)), an examination by the court (s. 290(4)), a mutually satisfactory disposition worked out at a meeting (s. 291) and a report signed by those who took part (s. 292). Each step can fail.
- C **is low.** For an accused on bail, a wait of years discounts the consequences of conviction. Where the offence can be compounded, composition has the effect of an acquittal (s. 359(8)) and is cheaper than any conviction.

All three point the same way. Low use is what the rule predicts of the present design.

What the Guidelines change

The Guidelines act on the first two terms directly.

- **They fix** P **in advance.** The reduction is published and graded, and it is taken off the sentence the court would otherwise impose under the Guidelines. The accused therefore knows S_p before he pleads, and S_p is always below S_0: the plea can never cost more than conviction after trial.
- **They raise** Π**.** A plea of guilty to the charge as framed (ss. 252, 264 and 275 BNSS) needs no one's consent. Under Rule 14 the accused receives a written indication and decides within a short, fixed time. The only step outside his control is the court's own acceptance of the plea.
- **They leave the court in control.** The sentence is still passed by the judge, within the statutory range, for reasons recorded. Compensation is fixed by the court on conviction and is not a term of the plea.

The third term, the wait, is not changed by the Guidelines alone. That is why they are issued together with listing measures, and why this section does not claim that a published reduction is sufficient in every case.

Illustration. An accused on bail faces a trial sentence of two years; the conviction record weighs as six months; contesting costs him the equivalent of about five weeks; he puts his chance of conviction at 60%; $\delta = 0.9$.

Situation	Cost of contesting	Cost of pleading	Does he plead?
Wait of five years; half off	0.99	1.50	No. Contesting is cheaper.
Wait of six months; half off	1.52	1.50	No. The plea is better by a hair, and would be worth attempting only if completion exceeded 96%.
Wait of six months; 60% off; completion 90%	1.52	1.30	**Yes.**

Situation	Cost of contesting	Cost of pleading	Does he plead?
The same, but completion only 20%	1.52	1.30	No. The route is too unreliable.

The smallest reduction at which the plea becomes attractive to this accused is:

Wait	Chance of conviction 40%	60%	80%
Six months	73%	49%	25%
Two years	80%	59%	39%
Five years	90%	76%	61%

Two things follow. A modest reduction suffices where the evidence is strong and the trial is prompt. And no reasonable reduction suffices where the trial is years away. A reduction that is published, graded by stage and paired with prompt listing is the design the arithmetic calls for.

Why the court's work falls

Let h_{alt} be the judicial time the case would otherwise have taken, h_{plea} the time of a completed plea and h_{fail} the time lost on a failed attempt. For each accused who applies, the time saved is

$$\Pi\,(h_{alt} - h_{plea}) - (1 - \Pi)\,h_{fail},$$

which is positive when $\Pi > h_{fail}/(h_{alt} - h_{plea} + h_{fail})$.

If a plea takes one-tenth of the time of a trial and a failed attempt one-twentieth, an attempt saves time whenever completion exceeds about 5%. At 90% completion each applicant saves about four-fifths of a trial. A route in which most attempts fail saves nothing. The reliability that makes the plea worth attempting is the same reliability that saves the court's time.

Why it is safe

The same rule shows where the dangers lie, and gives the court two tests.

Weak evidence. Against trial, a plea is attractive only to an accused whose chance of conviction exceeds a threshold p^* (Appendix A13). The deeper the reduction, the lower the threshold. This is the reason for a reduction that is bounded and graded and is not open-ended, and for the Guidelines' bar on reductions in the gravest band. The court can compute p^* for the terms on offer and satisfy itself, on the material before it, that the plea is supported.

Custody. Further detention raises the cost of contesting by more than it lowers the sentence, so an accused in custody may plead for release alone. The test is to decide release first (including under s. 479 BNSS), and then to ask whether the plea would still be made by the same accused at liberty, *with the time he has already served credited to him*. A plea that passes that test is not the product of custody.

These tests are additional to the court's duty to be satisfied that the plea is voluntary. They do not measure guilt, and nothing in this section estimates the number of wrongful pleas.

What this section does not claim

It does not forecast plea rates, disposal times or the clearance of arrears. It does not rely on the experience of other countries. It does not model bargaining: every payment is taken as fixed by the court. It gives the accused no advantage from delaying by a failed application. Its statutory readings, in particular the reading of sections 290 to 293 BNSS, are submitted for the Court's consideration and are not asserted as settled law.

5 Two models of the same question

Team Satyug has built two models. They share a core: an accused pleads when the plea costs him less than contesting. They differ in how much else they try to explain.

The foundation model. This model follows a case through the stages of the enacted procedure: notice, application, the judge's first check, the prosecutor, the investigating officer, the victim, and the final disposition. It asks how many cases survive every stage. It does not fix the unknown quantities at single values. Each one is given a range (an "assumption box"), the model is run several hundred times across the ranges, and the result is reported as a median with a 5th-to-95th percentile band. The chance of conviction at trial is held within a range centred on NCRB's figure (the model's conviction rate on contest is 67%, against NCRB's 66–67%). For the full programme it is allowed to rise by a stated amount, since the Bill's perjury and witness measures exist to raise it; that amount has no Indian estimate and is shown at several values. The model and its code have been independently reviewed, and part of its algebra has been machine-checked. It is set out in the technical appendix.

The whole-court simulator. This is a fuller model of the whole court. Besides the accused, it follows the complainant who may give up, the witness who may be bribed or threatened, and the court's queue, which sets the wait. Its central feature is feedback. If lying on the record is priced, fewer witnesses turn hostile; the chance of conviction rises; more accused plead; the queue shortens; and a shorter wait makes the plea more attractive still. It is calibrated to single values with stated bounds. It reproduces the United States and Italy when run on their inputs, but not England and Wales.

The first model answers: *what do the instruments do, case by case, with the queue taken as given?* The second answers: *what could the whole programme do if the feedback through the queue works as designed?* The pilot is designed to find where India lies between them.

6 What each instrument changes

Figures in this section are from the foundation model: completed pleas as a share of eligible cases, then as a share of convictions.

Practice Direction inside Chapter XXIII. A High Court direction can require plain-language notice, legal-aid counselling and a published guide to the likely sentence, all within the existing chapter. This helps access and predictability but leaves the signatures in place. *1.4% of cases; 2% of convictions.*

Practice Direction, Rule 14 route. Rule 14 of the proposed Practice Directions (Volume 2) uses the ordinary guilty plea, not Chapter XXIII. The court gives a written indication of sentence under the

National Sentencing Guidelines, valid for fifteen days. The victim is heard but does not sign. Lies on the record are routed to the court's own complaint. Whether a published reduction is lawful on this route is a question for counsel. The size of the reduction matters a great deal. With a flat one-third off, the model gives 12% of cases. With the reductions Volume 2 actually provides, which average about two-thirds on the assumed mix of cases, it gives *24% of cases; 34% of convictions.*

A separate list for fresh filings. A High Court can direct that new cases be heard on their own list, so that they do not wait behind the old stock. With trial within six months: *36% of cases; 50% of convictions.* Trial within a year gives almost the same.

The Bill. The Obstruction of Justice (Satyug) Bill (Volume 5, Phase 2) does what only a statute can: it prices perjury, protects witnesses, makes recorded statements usable and gives notice of the plea by statute. "The Bill" is reported in two ways, because the text and the programme are not the same thing. The *Bill text* (v6.7) alone, inside Chapter XXIII, sets ceilings on the statutory maximum and not a known sentence: *4.7% of cases; 7% of convictions.* The *full programme* is the Bill together with the Guidelines sentence, the Rule 14 route, the separate list and the evidence measures: *43% of cases; 57% of convictions.* If the same programme could run only through Chapter XXIII, with the prosecutor's signature and the statutory exclusions, it would give 19% of cases.

Compensation. Under the programme, compensation to the victim is ordered by the court on conviction, whether by plea or by trial. It is never a term of the plea, which keeps the plea from becoming a purchase. The full-programme and Bill figures include the order. The Practice Direction figures do not; Section 8 gives them with it.

Which barrier matters most inside Chapter XXIII

If every barrier inside Chapter XXIII were removed, completed pleas in the foundation model would rise by about 6.5 points. The model divides that gain among the barriers by removing them in every possible order and averaging (a Shapley division):

Barrier removed	Gain in completed pleas (points; median [5th, 95th])
Access (notice, understanding, counsel)	2.83 [0.51, 10.02]
Victim's signature	2.43 [0.47, 8.13]
Predictable sentence	0.38
Investigating officer's signature	0.37
Realistic benchmark for the reduction	0.15
All together	6.49 [1.23, 21.69]

Medians do not add up exactly to the median of the total. Access is the largest item in 69% of draws.

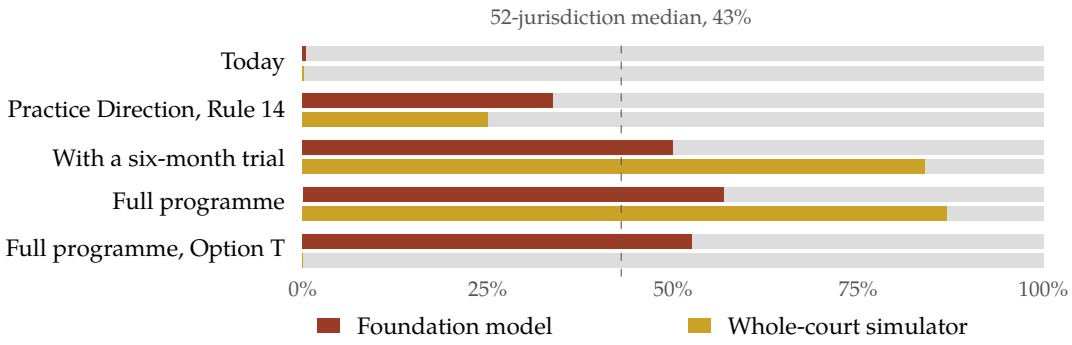
Two things follow. Notice and legal aid, which need no new law, are worth as much as the most contested legal question (the victim's signature). And even with every barrier gone, Chapter XXIII with a one-third reduction yields under 7%. The larger gains need a different route and a known, larger reduction, which is what Rule 14 supplies.

7 The numbers, side by side

	Foundation model		**Whole-court simulator**
	% of eligible cases; median [5th, 95th]	*% of convictions*	*% of all cases; % of convictions*
Today	0.28 [0.02, 1.74]	0.5	0.04; 0.2
Practice Direction inside Chapter XXIII	1.4 [0.2, 6.0]	2	not modelled
Practice Direction, Rule 14 route (the Pilot)	23.8 [3.9, 52.6]	34	9.3; 25 (bounds 0.6–93)
The same, with a six-month trial (separate list)	36.3 [7.2, 58.7]	50	48–74; 78–90
Bill text alone	4.7 [0.7, 17.8]	7	not modelled
Full programme	43.4 [7.5, 62.0]	57	73.8; 87 (bounds 46–98)
Full programme, foreign behavioural response (Option T)	40.4 [32.1, 51.0]	53	not modelled
Observed elsewhere	Share of convictions by plea: 52-jurisdiction median 43%; England and Wales 85–91%; Australia 91%		

The foundation model counts cases eligible for a plea at the framing of charge (offences up to seven years); the simulator counts all cases. Shares of convictions are the closer comparison, but not an exact one: the foundation model's share is among the offences it covers, and India's recorded figure counts only Chapter XXIII disposals where other countries count every guilty plea. The full programme row takes trial within six months and assumes the perjury and witness measures close half of the gap between today's conviction rate and certainty; from none to three-quarters, the result runs from 36% to 47% of cases. Neither model's figures are a forecast.

Share of convictions that come by plea: the two models



The simulator's separate-list bar is drawn at the middle of its 78–90% range. Where a model has no figure, no bar is drawn.

Where the two agree

- Today's rules produce almost no pleas. Both models reproduce this without being tuned to it. Run as the law stood before July 2024, without the thirty-day limit, the foundation model gives 1.9% of convictions by plea, against NCRB's 1.01% and 1.82% for 2022 and 2023. With the limit it gives 0.5%, against 0.69% for 2024, the first year partly under the new law.
- Changes inside Chapter XXIII alone do little, and so does the Bill's text without the Guidelines sentence.
- A known sentence on the ordinary guilty plea, by Practice Direction, is the step that matters first: a quarter to a third of convictions.
- Delay is the hinge. In the foundation model the full programme at today's wait gives 22% of cases; at six months, 43%.
- Under the full programme the plea share of convictions, for the offences modelled, is above the 52-jurisdiction median.

What carries the foundation model's result

In order of weight:

1. **The size of the reduction.** One-third off gives 12% under the Practice Direction; the Volume 2 reductions give 24%.
2. **The wait.** A six-month wait adds twelve points. The model takes the six months as given; the next subsection asks whether the courts can deliver it.
3. **Conviction risk.** The evidence measures add about eight points at the middle assumption. The result does not hang on this unmeasured quantity.
4. **The route.** All of the above assumes the Rule 14 route is lawful. Through Chapter XXIII alone the full programme gives 19%.

The result does not depend on the Volume 2 reductions being as large as they are. With a flat one-third off, the full programme still gives 37% of cases and 49% of convictions, above the international median, and weak-case pleading falls to 6%. On the lower conviction rate of IPC cases alone the evidence measures matter more: the full programme runs from 24% to 45% of cases across the range assumed. Taking both cautious assumptions together, IPC conviction rates and a flat one-third off, the full programme gives 29% of cases and 43% of convictions, with hours saved between -10% and $+47\%$. So the programme does not clear the international median, or save hours, under every combination of assumptions.

Can the courts deliver six months?

The headline figures for the separate list and the full programme are for fresh filings tried within six months. Two checks on that assumption, both taking present capacity as equal to the work now arriving each year:

- **A separate list is affordable, at the old cases' expense.** Under the full programme a fresh list held at six months needs about 64% of present capacity (46% to 97% across the assumptions), because pleas free more hours than the list uses. The old stock, served by what is left, then takes about thirteen years to clear (5 to 64). Without more judges or separate terms for old cases, the reform is quick for new cases and slow for old ones.
- **Without a separate list, the gain is slower and smaller.** If new and old cases share one queue, pleas shorten the wait and the shorter wait brings more pleas, but slowly. Under the full programme pleas rise from 22% of cases to 31% over twelve years, and from 29% to 40% of convictions, as the wait falls from about five years to about two and a half. In two runs in three the wait is still above one year after twelve years.

If the backlog is in fact growing today, capacity is below inflow and both results are optimistic.

Why the simulator is higher still

The simulator gives 87% of convictions where the foundation model gives 57%. The two models differ in several ways at once, so the gap cannot be assigned to any one of them. Feedback, measured inside the foundation model with everything else fixed, adds about nine points over twelve years, far less than the gap. The foundation model loses cases to missed notice, to advice not understood and to the court's own check. It lets about one eligible case in twenty end by compounding instead. And it has no feedback: in the simulator each plea shortens the queue for the next accused, and complainants who would have given up stay the course.

Option T replaces the foundation model's cost-based rule of behaviour with a large vignette experiment (Helm and Zottoli, 2026; 3,646 participants), assuming its results carry over to India. That is a strong assumption: the experiment's trial sentences were nine or eighteen years, against one to seven here. On the full programme it gives 40% of cases, against 43% from the cost-based rule. That two unrelated accounts of behaviour agree is a useful check on the level. It is not independent validation, since both rest on the same assumptions about the institutions. The agreement does not rest on a detail of the cost-based rule: it adds a chance of probation on top of the Volume 2 reduction, and without that it gives 41% of cases and 53% of convictions, almost exactly the Option T figures.

How to quote these numbers. For a court or a ministry: "On Team Satyug's foundation model, a Practice Direction giving a known sentence on a guilty plea would bring about a third of convictions by plea, and the full programme more than half, which is above the international median. A fuller model of the whole court gives higher figures. These are model results under stated assumptions, to be tested in a pilot." The simulator's Practice Direction figures also assume that every element of the Pilot can be adopted by Practice Direction, a point on which counsel's opinion is awaited.

8 Safeguards

A plea system fails if it convicts the innocent or if it extracts pleas by detention. The models were built to measure both.

Custody-induced pleas. The foundation model runs each case twice, once as the accused actually is and once as if he were on bail, and counts the pleas that occur only because of custody. With a small reduction this is the dominant problem: at one-third off, about 70% of pleas under the Rule

14 route are custody-induced, because few accused on bail find the plea worth taking. With the Volume 2 reductions and a short wait, accused on bail plead too, and the custody-induced share falls to about 6%. The number falls as well as the share: from 5.7 custody-induced pleas per 100 eligible cases to 2.1. The simulator reports the same concern on a different measure: coerced pleas of 11.7 per 10,000 under the Pilot and 0.5–2.1 under the Bill.

Pleas in weak cases. The larger reduction has its own cost. The model contains a group of accused against whom the case is weak. A weak case is not innocence, and the rate at which this group pleads is not an estimate of wrongful convictions. It is a warning signal. At one-third off, about 8% of them plead. With the Volume 2 reductions it is 11–12%, and 8% under Option T, against the Guidelines' own target of 5%. Taking custody out of the decision brings it to about 6%. The direction matches the Helm and Zottoli experiment, in which participants told to assume innocence accepted the largest reductions more often.

Compensation as a screen. In the table below, weak-case pleading falls from 16% with a separate list to 11% under the full programme. The reason is the compensation order. The court orders it for certain on a plea, but on a contest only if the accused is convicted. It therefore raises the cost of pleading most for those least likely to be convicted. Without the order, the full programme would give 48% of cases by plea and 20% of weak-case accused pleading; with it, 36% and 11% (both before the evidence measures). The order halves weak-case pleading at a cost of about twelve points of pleas. This is a property of the model's arithmetic, not a reason to set compensation: the amount must rest on the victim's loss, and its effect on real accused, especially poor ones, is untested. If courts under the Practice Direction also ordered compensation on every conviction, the Practice Direction figures would be 15% of cases and 22% of convictions, and 28% and 41% with a separate list, with weak-case pleading at about 8%.

The medians are not guarantees. Across the assumed ranges, under the full programme the share of weak-case accused who plead runs from 1% to 37%, the custody-induced share of pleas from 1% to 41%, and hours saved from 3% to 54%. The full programme passes the international median in 70% of runs, keeps weak-case pleading within the 5% target in 29%, and does both while saving hours and keeping custody-induced pleas under a tenth in only 4%. With custody taken out of the decision the last figure is 13%. These are shares of the assumed ranges, not probabilities; they show that success on pleas and success on safety do not usually arrive together, and that the pilot must measure both.

What a ceiling and bail before plea buy. Two levers bear on this: a ceiling on the reduction, and taking custody out of the decision by release on bond before the court indicates a sentence. A search over the ceiling, with and without custody in the decision, gives a clear ranking. The ceiling alone does little: at any ceiling from 20% to the drafted schedule, the full programme meets all four tests in 7% of runs or fewer. Taking custody out of the decision does most of the work: weak-case pleading falls from 11% to 6% at the drafted reductions. The two together do best. With custody out and a ceiling of 60%, the full programme gives 39% of cases and 52% of convictions by plea with 3.4% of weak-case accused pleading, and meets all four tests in 22% of runs; under the foreign behavioural response the figures are 34%, 45%, 3.7% and 46%. Against the drafted schedule this costs five to seven points of pleas. The ceiling here is on the reduction measured against the Guidelines sentence after trial; how it maps to the Stage 1 fractions in Volume 2 is a drafting question.

If the evidence measures also bite on weak cases. The model assumes that pricing perjury raises

conviction risk only where the accused is of the guilty type. If it raised conviction risk in weak cases by half as much, 19% of weak-case accused would plead, and 31% if by as much, against 11% as modelled. Measures that make conviction more likely across the board would make this risk worse, not better.

Regime (foundation model)	Pleas (% of eligible)	On bail (% of bailed)	Custody-induced (% of pleas)	Custody-induced (per 100 eligible)	Weak cases pleading (%)
Rule 14 route, one-third off	12.3	3.1	69.9	5.7	7.7
Rule 14 route, Volume 2 reductions	23.8	18.3	22.0	3.9	12.1
The same, with a six-month trial	36.3	34.7	5.9	1.8	16.4
Full programme	43.4	41.5	5.9	2.1	11.0
Full programme, custody shield	41.3	41.5	—	—	6.1
Full programme, Option T	40.4	38.5	4.8	1.9	8.4

Three safeguards follow, and all are in the Practice Directions:

- **Bail before plea.** Release on bond is considered before any indication of sentence is given, so that the plea is not the price of liberty.
- **A residual cost, and a ceiling on the earliest reduction.** The plea must always cost the accused something. The figures above make the ceiling on the Stage 1 reduction a live decision: a lower ceiling gives fewer pleas and fewer pleas in weak cases.
- **The judge's check.** The court satisfies itself of voluntariness and of a factual basis before accepting the plea. The Bill text as drafted removes the s. 290(4) voluntariness stage; the full programme restores it, and this paper assumes it restored.

Hours, not counts. A plea saves court time only if it replaces something that would have taken longer, and a failed application costs time. The foundation model therefore counts net judicial hours, which can be negative. Under the Rule 14 route with the Volume 2 reductions the saving is about 20% of today's hours, and under the full programme about 36%. A rise in pleas with no saving in hours is not a success.

Closing the other exits. Making non-compoundable offences truly non-compoundable sounds like an obvious companion measure. The model advises caution. If acquittal-by-settlement is closed and nothing else changes, most accused contest, court hours rise (by about a tenth in the model, even with the Rule 14 route open), and victims receive less. Closing one exit moves settlements to another: close withdrawal and quashing, and hostile witnesses rise; price hostility alone, and withdrawals rise. The exits can be closed usefully only together with perjury pricing, a working plea route beyond thirty days, and prompt trials. Quashing cannot be abolished by statute, since the High Courts' constitutional powers remain; it can only be channelled.

9 The pilot, and what would prove the model wrong

No number in this volume should be relied on until it is tested. The test is a pilot in selected courts under a Practice Direction.

What may be varied. Only administrative matters: the form and timing of notice, the order of

legal-aid counselling, the scheduling of meetings, and a leaflet stating the likely sentence. Rights, eligibility, sentences, bail and judicial decisions are never randomised.

What is measured. Applications and completed pleas as a share of eligible cases; plea convictions as a share of all convictions; judicial hours; time to disposal; signs of involuntary or wrongful pleas, such as withdrawal attempts and failed factual-basis checks; and, at each stage, the recorded reason a plea failed. Stage-wise reasons matter most. Without them, no study can tell whether pleas fail for want of access, for want of a signature or for want of a known sentence.

What would count against the model. Each prediction is conditional, and a null result counts only if the intervention was in fact delivered and the study was large enough to detect the effect.

- Where a participant's agreement is sometimes withheld, requiring it should lower completion among those who apply. If it does not, the signature account is wrong.
- Completion should be lower where the victim must agree than where the victim is only heard.
- Holding the plea sentence fixed, applications should rise where the expected sentence after trial is higher.
- Where access is assured, no later stage blocks the plea, and the plea is worth attempting for enough of the accused newly reached, applications should rise. If they stay near zero under those conditions, access is not the barrier.

Separately, as a test of the policy and not of the model: if pleas rise but hours do not fall, or the rise is concentrated among detained accused, the reform has failed whatever the count.

A narrower sentencing range was expected to raise applications in every case. The mathematics shows this holds only under conditions, so it is tested but not relied on.

Size. To detect a rise in completed pleas from 1% to 3% of eligible cases needs about 770 eligible cases in each arm, and between 1,500 and 8,400 once the grouping of cases by court is allowed for.

Before the pilot. Several predictions can be checked now from e-Courts case histories and NCRB tables: that compounding becomes more likely as a case ages while recorded compensation falls; that withdrawals and quashings in non-compoundable offences cluster after the thirty-day window; that hostile-witness acquittals rise with delay; and that accused in custody settle earlier.

10 Limits

- **No Indian case data have been analysed.** The only external check is against NCRB's national tables.
- **The shares of convictions are for the offences modelled.** If offences outside the model produced no pleas and carried a tenth to three-tenths of convictions, the full programme's 57% would be 40% to 51% of all convictions.
- **Every number depends on its assumptions.** The bands are wide: the full-programme figure runs from 8% to 62% of cases. Halving every assumed range leaves the median at 42% and narrows the band to 22%–56%. The results are most sensitive to three quantities that have never been measured in India: how much of the stigma of conviction a plea retains, the cost

of litigation, and the stigma itself. A band is the spread of results across the assumed ranges. It is not a confidence interval, and it does not give the probability that the true figure lies inside it. The Volume 2 reductions are applied on an assumed mix of cases and stages (six in ten pleas at the earliest stage), not on Indian data.

- **The results are checked, not replicated.** The figures were checked against the result files, and the code line by line, by an independent reviewer. They have not been replicated by a second party. The effect of the evidence measures on conviction risk is assumed, not estimated.
- **The legal readings are provisional.** The effective veto under s. 292, the scope of "may" in s. 293, eligibility timing in summons cases, the legality of a published reduction on the Rule 14 route, and the reading that Chapter XXIII consent applies only to charge agreements are all for counsel.
- **The hostile-witness channel is not yet reconciled with data.** The settlement-exits model puts hostile-witness exits at about a fifth of eligible cases today. Counted as NCRB counts them, that implies a conviction rate of about 45%, against the recorded 66–67%. Either the model's share is too large, or few recorded acquittals are of this kind, or the model counts twice the hostile acquittals that are already inside the trial acquittal rate. That level is not quoted in this paper.
- **Applications are less robust than completions.** In the foundation model the number of applications depends on an error term in the choice rule. This paper therefore reports completed pleas.
- **The cross-country evidence is correlational**: 52 observations and weak significance.
- **Custody is understated.** The model treats release at the statutory threshold (s. 479) as automatic.

The appendix (Section A11) lists each open question with what would close it.

11 Conclusion

India's near-zero plea rate is not a mystery and not only a shortage of judges. Under today's rules a guilty accused does better by waiting, and the rules that make it so can be named: a short window, several signatures, an unknown sentence, exits that stay open, lies that cost nothing, and delay.

A High Court can change the first part of this by Practice Direction. On the foundation model, a known sentence on the ordinary guilty plea, at the reductions the Guidelines provide, would bring about a quarter of eligible cases to an early end by plea, and a separate list more than a third, provided bail comes before the plea. The full programme takes pleas past half of convictions for the offences modelled, above the international median at the middle of the assumed ranges, though not across all of them. How much further it goes depends on whether shorter queues feed on themselves, as the simulator supposes. What it costs depends on where the ceiling on the earliest reduction is set. Both are for the pilot to measure.

Terms used

Eligible or at-risk case
A case in which the offence carries up to seven years and the accused could plead at the framing of charge.

Completed plea
A plea that passes every stage and ends in conviction and sentence. An application alone is not counted.

Assumption box
The set of ranges given to the quantities for which India has no estimate. Results are medians across the box.

Rule 14 route The ordinary guilty plea with a written indication of sentence under the Practice Directions, outside Chapter XXIII.

Full programme
The Bill together with the Guidelines sentence at the Volume 2 reductions, the Rule 14 route, the court's check, the separate list and the evidence measures.

Option T The foundation model with the accused's response taken from foreign experiments.

Custody-induced plea
A plea that would not have been made had the same accused been on bail.

Sources

National Crime Records Bureau, *Crime in India* 2022, 2023 and 2024, Vol. III, Tables 18A.2 and 18A.4. Bharatiya Nagarik Suraksha Sanhita, 2023. Paolini, Kantorowicz-Reznichenko and Voigt (2025) for the comparison of 52 jurisdictions. Helm and Zottoli (2026) for plea acceptance by discount and assigned guilt. Dobbie, Goldin and Yang (2018) for the effect of pre-trial detention. Full references follow the technical appendix.

Technical appendix

The foundation model: structure, assumptions and results

A1 Scope

This appendix states the foundation model used in the paper: what it assumes, what it computes and what it returns. It is a model of an accused person's decision to plead guilty and of the procedural stages a plea must pass, evaluated over stated ranges of the quantities for which India has no estimate. It is not an estimate from Indian case data and it does not forecast. The whole-court simulator is described in Section A9. Statutory references are to the Bharatiya Nagarik Suraksha Sanhita, 2023 (BNSS), checked against the Gazette text. Readings of the statute are marked as such and are for counsel to settle.

A2 The statutory framework

Chapter XXIII (ss. 289–300). The Chapter has been in force since 1 July 2024. It succeeds Chapter XXIA of the Code of Criminal Procedure, 1973 (ss. 265A–265L), which had the same structure without the thirty-day limit and still governs proceedings that were pending on that date (s. 531). The Chapter applies to offences punishable with up to seven years, excluding offences against a woman or a child and notified socio-economic offences (s. 289). The accused applies within thirty days of the framing of charge (s. 290(1)), with an affidavit of voluntariness and of no previous conviction for the same offence. The court examines the accused in camera and allows up to sixty days for a "mutually satisfactory disposition" (s. 290(4)). In police-report cases the participants are the Public Prosecutor, the investigating officer, the accused and the victim; in complaint cases, the accused and the victim (s. 291). The report of a disposition is signed by the presiding officer and all who participated (s. 292). On a disposition the court awards the agreed compensation, may release on probation, and otherwise may sentence to one-fourth of the punishment provided (one-sixth for a first-time offender), or to half of any minimum (one-fourth for a first-time offender) (s. 293). The judgment is final, subject to Articles 136, 226 and 227 (s. 295).

Compounding (s. 359). For listed offences the person named may compound, in some cases with the court's permission. Composition has the effect of an acquittal (s. 359(8)).

Ordinary guilty pleas. The trial chapters allow the court to record a plea of guilty and convict on it. Any concession then lies in ordinary sentencing discretion.

Readings adopted (for counsel). (i) A disposition that a participant will not sign is not mutually satisfactory, so the prosecutor, the investigating officer and a participating victim each hold an effective veto. (ii) A victim who is noticed but does not attend need not sign; the opposite reading roughly halves status-quo completions. (iii) Complaint cases have no prosecutor stage. (iv) A High Court may by Practice Direction regulate notice, listing, legal aid and the information given to the accused, and may indicate a Guidelines sentence on an ordinary guilty plea (Rule 14 of the proposed Directions); it cannot alter ss. 289–293.

A3 The model

A3.1 Population

Each simulated case is an offence with statutory maximum $S_{max} \in \{1, 2, 3, 5, 7\}$ years (weights 0.15, 0.25, 0.30, 0.15, 0.15). The accused is of a *guilty type* or a *low-p type*. The low-p type faces a conviction risk 0.4 times that of the guilty type. It is a proxy for the innocent and for weak cases; it is not innocence, which the model cannot observe. Each case draws a conviction risk on contest p, an expected trial sentence $S_0 = \lambda S_{max}$, a stigma of conviction K (in years-equivalent), a wait to verdict T, custody status, time already served s_1, transferable wealth, and whether there is an individual victim.

A3.2 The accused's choice

The cost of contesting is

$$C(T) = \delta^T \, \Omega \, p \left[v\big((S_0 - s_1 - x)^+\big) + K \right] + c_L + x,$$

where δ is the annual discount factor, $\Omega \in (0, 1]$ a weight for the prospect of appeal, $v(s) = s^\gamma$ the disutility of a sentence, c_L the cost of litigation, and x further time in custody before verdict, capped at the s. 479 thresholds (one-third of the maximum for a first-time offender, one-half otherwise). The cost of a completed plea is

$$P = \mathbb{E}\, v\big((S_p - s_1)^+\big) + \rho K + m_c,$$

where S_p is the plea sentence, $\rho \in [0.4, 1]$ the share of stigma that a plea conviction retains, and m_c any compensation ordered by the court. The gain from attempting a plea is

$$G = \tilde{\Pi} \, \max\{C - P, \, -\Delta_f\} - (1 - \tilde{\Pi}) \, \Delta_f - c_n,$$

with $\tilde{\Pi}$ the anticipated probability of completion, Δ_f the extra cost of contesting after a failed attempt (up to sixty days' delay under Chapter XXIII; fifteen days, the validity of the court's indication, on the Rule 14 route), and c_n the cost of attending. The accused applies if $G + \varepsilon > 0$, where ε is a small logistic error (scale 0.05 years-equivalent). Beliefs about completion are correct on average.

A3.3 Access and completion

A completed plea requires, in order: eligibility and effective notice; an application; and passage through each stage. The probability of access is $\alpha = e \, \nu \, \pi$ (eligible; noticed in time; understood). The completed-plea rate is

$$R = \Pr(E, N) \cdot \Pr(A \mid E, N) \cdot \Pr(\text{complete} \mid A, E, N).$$

Under Chapter XXIII the stages after application are the court's finding of voluntariness (κ_{J0}), the prosecutor (κ_P), the investigating officer (κ_{IO}), a participating victim, and the accused's own signature. The victim signs if the compensation offered meets a private reservation; the accused chooses the offer, limited by wealth. On the Rule 14 route the only stage is the court's acceptance of the plea (κ_{J0}). Because the same aggregate rate R can arise from low access, low willingness or failed completion, aggregate plea counts cannot distinguish these; stage-wise records are needed.

A3.4 The plea sentence

Chapter XXIII. $S_p = f \, S_{max}$ with $f = 1/6$ (first-time offender) or $1/4$, replaced by probation with probability q. The accused treats "may" in s. 293 as ambiguous: with weight ω he evaluates the worst case in which the fraction is exceeded. Because the fraction is taken on the statutory maximum while trial sentences are well below it ($\lambda < 1$), the statutory reduction can be small or negative against the expected trial sentence: with $q = 0$ the plea is never preferred on sentence alone when $f \geq p\lambda$.

Rule 14 route. $S_p = (1 - d) \, S_0$, known in advance, where d is the reduction that the Volume 2 rules give for the case's offence band, first-offender status, score and stage of plea. On the model's offence mix the mean reduction is 66%. This rests on an assumed mix of scores and stages (60% of pleas at the earliest stage, 30% at the second, 10% at the third), not on Indian data. A flat one-third reduction is run as a lower bound.

Bill text alone. Ceilings on the statutory maximum (one-fourth for a first-time offender, one-half otherwise, at the earliest stage), with the sentence taken as uniform below the ceiling and the ceiling as the worst case.

A3.5 Compensation, compounding and custody

Compensation. Where the regime includes it, the court orders compensation on conviction by either route: $m_c = \min(\text{victim's reservation}, \text{wealth})$, zero in victimless cases. It is certain on a plea and adds $\delta^T \Omega \, p \, m_c$ to the cost of contesting. It is never a term of the plea.

Compounding. In a compoundable case (20–50% of eligible cases) the accused who can satisfy the victim prefers compounding, which ends in acquittal, to any plea. Compounding is left open in every regime.

Custody-induced pleas. Each case is evaluated twice, as it is and with custody removed from the decision. A plea completed in the first and not in the second is custody-induced. This measures how far the model's pleas depend on custody. It is not a finding that any real plea is involuntary in law.

A3.6 Evidence effect and transported response

Evidence effect. The programme's perjury and witness measures are represented by a parameter $e \in [0,1]$: for guilty types $p' = p + e\,(1-p)$, so that a share e of the gap between today's conviction risk and certainty is closed. Low-p types are unchanged; that the measures would not also raise conviction risk in weak cases is an assumption, not a safety result. India has no estimate of e; it is run at 0, 0.25, 0.5 and 0.75.

Transported response. As an alternative to the cost-based rule, the probability of accepting an offer at reduction d is taken from the experiment of Helm and Zottoli (2026; 3,646 participants; acceptance by reduction and by assigned guilt), within its 95% bands, with the detention effect of Dobbie, Goldin and Yang (2018) of 10.8 to 12.0 points added for accused in custody. The experiment used vignettes with trial sentences of nine or eighteen years, against maxima of one to seven years here, and the detention effect is added to it as a uniform increment. This assumes that responses measured abroad, at another sentence scale, carry over to Indian accused. That assumption is not defended here; the transported response is reported as a second reading, not as the model.

A3.7 What is measured

Completed pleas as a share of eligible ("at-risk") cases. The plea share of convictions: completed pleas divided by completed pleas plus the expected trial convictions of those who contest, which is the nearest counterpart of NCRB's plea disposals divided by convictions and of the cross-country figures. It is a share among eligible cases only: a share for all offences would also need the conviction weight and plea share of the offences outside the model. The share of pleas that are custody-induced. The share of low-p types who plead. Net judicial hours saved: with a contested trial normalised to one, a completed plea costs 0.03–0.15, a compounded case 0.05–0.15, a failed attempt adds 0.03–0.10 to the contest, and up to 30% of gross savings are re-absorbed by induced filings. The saving can be negative.

A4 Formal results and proofs

The propositions below hold in the model of Section A3. Each has also been compiled in Lean 4 with no errors. Where the compiled statement is narrower than the one proved here, that is noted, and a compiled theorem is credited only as stated. Throughout, $\Pi = \tilde{\Pi}$ is the completion probability, C the cost of contesting, P the cost of a completed plea, Δ_f the cost of a failed attempt and c_n the cost of attending.

P1 (aggregation). $R = \Pr(E,N)\cdot\Pr(A \mid E,N)\cdot\Pr(\text{complete} \mid A,E,N)$. This differs from the product of the three population averages by covariance terms of either sign.

Proof. A plea is completed only if the accused has access, applies and completes, so the identity is iterated conditioning. For the sign, take two equally weighted types with (access, application, completion) probabilities $(1,1,1)$ and $(0,0,0)$: $R = \frac{1}{2}$, while the product of averages is $\frac{1}{8}$. With types $(1,1,0)$ and $(0,0,1)$: $R = 0 < \frac{1}{8}$. □

Corollary. Any triple with the same product gives the same R. Aggregate plea counts therefore cannot separate low access, low willingness and failed completion.

P2 (added stages). Here Π is the probability that every stage other than the accused's own last decision is passed, and an accused who is indifferent at that decision declines. (a) For a fixed set of applicants and fixed offers, a further required stage with pass probability $\kappa \in [0,1]$ multiplies completions by κ. (b) The gain G

is non-decreasing in Π, and strictly increasing if and only if $P < C + \Delta_f$. (c) An accused with $P \geq C + \Delta_f$ never completes. With $\Delta_f, c_n \geq 0$ his gain is $-\Delta_f - c_n \leq 0$, so he applies only when the error term is large enough. Such applications are noise in the choice rule; the model gives delay no value, so they are not evidence of delay-seeking.

Proof. (a) is the chain rule for the probability of passing every stage. For (b), write $G = \Pi\,[\max\{C - P, -\Delta_f\} + \Delta_f] - \Delta_f - c_n$, which is affine in Π with slope $\max\{C - P + \Delta_f,\, 0\}$. For (c), if $P > C + \Delta_f$ the accused strictly prefers to decline at the last stage, and at equality he declines by the convention stated; the slope is zero, so $G = -\Delta_f - c_n$ whatever Π is. □

So a further signature weakly lowers completions among those who apply (a stage that always passes changes nothing); κ is the pass probability given that the earlier stages were passed. It lowers applications as well only for accused for whom the plea is better than a failed attempt. In the simulation offers re-optimise when a signature is removed, so (a) is exact only for fixed offers.

P3 (signature as veto). Let the victim sign if and only if the compensation offered, m, is at least a private reservation r_V with distribution function F, and let the accused choose $m \in [0, w]$, where w is his wealth. (a) If $F(w) < 1$, completion has probability below one for every feasible offer. (b) If the chosen offer has the form $\min(w, m^*)$, completion is non-decreasing in w.

Proof. (a) The victim signs with probability $F(m) \leq F(w) < 1$, and the other stages can only lower the product. (b) $\min(w, m^*)$ is non-decreasing in w and F is non-decreasing. □

That optimal offers take the form in (b) is assumed, not proved. Since the prosecutor and the investigating officer cannot lawfully be paid, their pass probabilities do not respond to m.

P4 (predictability). The plea sentence enters the gain through $g(S_p) = v\big((S_p - s_1)^+\big)$, the disutility of the sentence net of time served, with v increasing and $v(0) = 0$. Compare two distributions of S_p with the same mean, the second a mean-preserving spread of the first, holding all else fixed. (a) If v is convex, and in particular if it is linear, g is convex and G is weakly lower under the spread. (b) If both distributions lie at or above s_1, then G is unchanged for linear v and weakly higher for concave v. (c) If v is strictly concave and the support crosses s_1, the sign is not determined. (d) If the accused evaluates the plea at the worst sentence in a set $[S_{lo}, S_{hi}]$, G is non-increasing in S_{hi} whatever happens to the mean, and strictly decreasing where $\Pi > 0$, $P < C + \Delta_f$, $S_{hi} > s_1$ and v is strictly increasing.

Proof. G is non-increasing in P, and S_p enters P only through $\mathbb{E}\, g(S_p)$. (a) $s \mapsto (s - s_1)^+$ is convex and non-decreasing, and an increasing convex function of it is convex; a mean-preserving spread weakly raises the expectation of a convex function. With linear v and $s_1 = 1$, a certain sentence of one year adds nothing to time served, while a sentence of zero or two years with equal probability has the same mean and adds half a year in expectation. (b) On $[s_1, \infty)$, $g(s) = v(s - s_1)$, which is linear or concave as v is; apply Jensen's inequality. (c) Take $v(s) = s^{0.8}$ and $s_1 = 1$. Spreading a certain one-year sentence to zero or two years with equal probability raises expected disutility from 0 to 0.5. Spreading "zero or four years with equal probability" to "zero with probability $\frac{1}{2}$, three or five years with probability $\frac{1}{4}$ each" keeps the mean at two and lowers expected disutility from 1.204 to 1.193. (d) The plea is valued at $g(S_{hi})$, which is non-decreasing in S_{hi}; by P2 the gain is flat once $P \geq C + \Delta_f$, which gives the conditions for strictness. For example, with $s_1 = 0$ a plea of one year with probability $7/9$ and ten years with probability $2/9$ has mean 3; with $C = 5$ and certain completion an expected-value accused applies and a worst-case accused does not. □

So a narrower range raises applications under convexity, linearity with time served, or aversion to ambiguity, and not in general under concavity. (Parts (a), (b) and (d) and the first example in (c) are compiled for symmetric two-point spreads.)

P5 (benchmark and application threshold). Take an accused on bail with no time served ($x = s_1 = 0$), linear v, no compensation order and no ambiguity over the sentence, and ignore the error term. Part (b) holds for any P and C that include every applicable component. Let the plea sentence be $(1 - q) f S_{max}$ in expectation and the expected trial sentence λS_{max}. (a) A completed plea is preferred to contesting at wait

T if and only if $P < C(T)$, and at no wait $T \geq 0$ if and only if

$$(1-q)fS_{max} + \rho K \;\geq\; \Omega\, p\,(\lambda S_{max} + K) + c_L.$$

With $K = 0$, $\Omega = 1$ and $c_L = 0$ this is $(1-q)f \geq p\lambda$. (b) For $0 < \Pi \leq 1$ the accused applies ($G > 0$) if and only if

$$C(T) \;>\; P + \frac{1-\Pi}{\Pi}\,\Delta_f + \frac{c_n}{\Pi},$$

and for $\Pi = 0$ he does not apply. (c) The reduction measured against the expected trial sentence is $1 - (1-q)f/\lambda$, which is negative when $\lambda < (1-q)f$.

Proof. (a) $C(T) = \delta^T \Omega\, p\,(\lambda S_{max} + K) + c_L$ is decreasing in T and largest at $T = 0$, so $P < C(T)$ for some T if and only if $P < C(0)$. (b) If $G > 0$ then $\max\{C - P, -\Delta_f\} = C - P$, since on the other branch $G = -\Delta_f - c_n \leq 0$. Then $G = \Pi(C - P) - (1-\Pi)\Delta_f - c_n$, and dividing by Π gives the threshold; the converse is the same algebra. At $\Pi = 0$, $G = -\Delta_f - c_n \leq 0$. (c) is arithmetic. □

Part (b) is the condition that governs applications: a low chance of completion raises the bar well above P. With $C(0) = 2$, $P = 1$, $\Pi = 0.1$ and $\Delta_f = c_n = 0.1$ the completed plea is much the better outcome, yet the threshold is 2.9 and the accused does not apply. With $\lambda = 0.2$ and $q = 0$, the statutory "three-quarters off" ($f = 1/4$) is a 25% premium over the expected trial sentence. With the error term these are statements about the deterministic part of the choice; the probability of applying is never exactly zero. (Part (a) compiled for $q = 0$; part (b) compiled as stated.)

P6 (compounding). If the victim would accept the same transfer m by either route, and the costs of attending and of any permission are no lower for a plea, compounding costs the accused strictly less than a plea whenever $\mathbb{E}\, v\big((S_p - s_1)^+\big) + \rho K > 0$.

Proof. Compounding ends in acquittal and costs $m + c$. A plea costs $m + \mathbb{E}\, v\big((S_p - s_1)^+\big) + \rho K + c'$ with $c' \geq c$. □

P7 (delay). (a) If the trial is discounted by $\beta\delta^T$ for $T > 0$ with $\beta < 1$ (present bias), then for $0 < T_2 \leq T_1$ cutting the wait from T_1 to T_2 raises the discount factor by less than cutting it to zero. (b) If discounting is hyperbolic, $D_h(T) = 1/(1+kT)$, matched to δ at one year, then $D_h(T) \geq \delta^T$ for $T \geq 1$ and $D_h(T) \leq \delta^T$ for $0 \leq T \leq 1$.

Proof. (a) $\beta\delta^{T_2} - \beta\delta^{T_1} < 1 - \beta\delta^{T_1}$ because $\beta\delta^{T_2} < 1$. (b) Matching gives $\delta = 1/(1+k)$, so $\delta^T = (1+k)^{-T}$. By Bernoulli's inequality $(1+k)^T \geq 1 + kT$ for $T \geq 1$ and $(1+k)^T \leq 1 + kT$ for $0 \leq T \leq 1$. □

So shortening a nine-year wait to two years moves a present-biased accused less than the exponential model predicts, and the exponential model overstates how heavily long waits are discounted.

P8 (hours). Per eligible case, net hours saved by a policy are

$$S = \Pr(\text{conv})\big(\mathbb{E}[h^{cf} \mid \text{conv}] - h_{plea}\big) - \Pr(\text{fail})\, h_{fail} - h_{ind},$$

where "conv" are the cases converted to completed pleas, h^{cf} the hours those cases would otherwise have taken, "fail" the attempts that fail, and h_{ind} the hours of induced filings. Pleas can rise while $S < 0$.

Proof. Partition cases into converted, failed and unaffected, and sum the change in hours over each. For the second claim take $\Pr(\text{conv}) = 0.1$ with $\mathbb{E}[h^{cf} \mid \text{conv}] = 0.3$ and $h_{plea} = 0.1$, and $\Pr(\text{fail}) = 0.3$ with $h_{fail} = 0.1$: $S = 0.02 - 0.03 < 0$. □

Pleas save work when the path they replace would have taken more hours than the plea and its overhead. That is clearest when the case would have gone to full trial; where it would have ended by compounding, withdrawal or early acquittal, the saving is small or negative.

P9 (conditional low-use trap). Let the backlog in hours follow $B_{t+1} = F(B_t) = \max\{0,\; B_t + g(B_t)\}$, where $g(B) = W(T_{min} + B/H) - H$ is continuous, W is the inflow of work at wait T and $H > 0$ is capacity. Let

$0 < B_1 < B_2$ and suppose $g < 0$ on $[0, B_1)$, $g > 0$ on (B_1, B_2), $g < 0$ on (B_2, ∞), and $B + g(B)$ is non-decreasing. Then orbits starting in $[0, B_1)$ converge to 0, orbits starting in (B_1, ∞) converge to B_2, and B_1 is an unstable threshold.

Proof. On $(0, B_1)$, $F(B) < B$, so the orbit decreases; it is bounded below by zero and converges to a fixed point, and the only fixed point in $[0, B_1)$ is 0. On (B_1, B_2), $F(B) > B$; since F is non-decreasing and $F(B_2) = B_2$ by continuity, $F(B) \leq B_2$, so the orbit increases to a fixed point, which is B_2. Above B_2, $F(B) < B$ and $F(B) \geq F(B_2) = B_2$, so the orbit decreases to B_2. □

The sign pattern requires inflow to rise with the wait somewhere, which happens if pleas fall with delay faster than attrition removes work. Computed from the model, the plea rate does fall with the wait, but not fast enough: inflow never rises with the wait and no trap arises. The trap is a possibility, not a finding about India. Capacity shocks would test it.

Machine checking. The propositions are compiled in Lean 4 as 60 theorems, eleven of which state P2, P4 and P5 in the forms printed here (the gain with the accused's own exit; the sentence net of time served, for two-point spreads; the application threshold). The settlement-timing results of Section A8 and the adjournment-cost results add 21. All 81 compile with no errors. Several are algebraic identities or simplified forms, as noted; the count is not evidence for the substantive claims. An index of theorems by proposition is in the source package.

A5 Regimes

Regime	What is assumed
Today	Chapter XXIII as enacted; notice 5–50%, understanding 30–80%.
Practice Direction inside Chapter XXIII	Notice and legal aid at framing of charge (60–95% each). Signatures, window and fractions unchanged.
Practice Direction, Rule 14 route	Ordinary guilty plea; no prosecutor, officer or victim signature; the court's check; known sentence at the Volume 2 reduction; no s. 289 exclusions; no compensation element.
— with a separate list	Trial within six months, taken as given. Whether a separate list delivers it within present capacity is not modelled.
Bill text alone	Chapter XXIII with s. 292 repealed (victim heard; prosecutor remains); statutory notice (80–95%); ceilings on the statutory maximum; voluntariness stage as drafted (deleted); compensation by the court on conviction.
Full programme	Rule 14 route; Volume 2 reductions; statutory notice; the court's check; compensation by the court on conviction by either route; trial within six months; evidence effect e.
Full programme, transported response	The same institutions, with acceptance taken from the experiment.

Not represented in any regime: disclosure of the prosecution's evidence before plea; relief from stigma for the earliest pleas to minor offences (which Volume 2 provides, and which would raise both pleas and low-p pleading); the folding of ordinary guilty pleas into Chapter XXIII by the Bill; any feedback from pleas to the wait.

A6 The assumption box

Each run draws every parameter independently and uniformly from its range, then simulates 6,000 cases. Results are the median and the 5th and 95th percentiles over 400 runs. These percentiles describe the spread of results over the assumed ranges and weights. They are not confidence intervals and give no probability that the true Indian figure lies within them; the same holds for statements such as "largest in 69% of runs". Common random numbers are used across regimes.

Parameter	Range
Wait from charge to verdict if contested (mean, years)	2–8
Annual discount factor δ	0.85–0.95
Conviction risk on contest, guilty type (mean)	0.58–0.95
Share of low-p types	0.10–0.30
Expected trial sentence as a share of the maximum, λ	0.05–0.40
Stigma of conviction K (mean, years-equivalent)	0.2–2.0
Stigma retained after a plea, ρ	0.4–1.0
Litigation cost of contest c_L (years-equivalent)	0.05–0.60
Share in custody at the decision	0.10–0.40
Share excluded by s. 289	0.15–0.40
First-time offenders	0.50–0.80
Court finds the plea voluntary, κ_{J0}	0.85–1.00
Prosecutor signs, κ_P	0.30–0.90
Investigating officer signs, κ_{IO}	0.60–1.00
Victim reachable and participates	0.50–0.95
Victim's reservation (median); accused's wealth (median)	0.10–1.50; 0.10–1.00
Cases with no individual victim (police-report cases)	0.10–0.40
Probation on a plea, q	0.10–0.50
Ambiguity over "may": excess; weight on worst case	0–1; 0–0.6
Curvature of sentence disutility, γ	0.7–1.3
Compoundable share of eligible cases	0.20–0.50
Hours of an acquittal path relative to a full trial	0.30–0.80

The conviction-risk range is set so that the model's conviction rate on contest brackets NCRB's completed-trial rate for IPC and special-law cases together. It is the only range set against data.

A7 Results

A7.1 Check against NCRB

The model's conviction rate on contest today is 67.31 [51.81, 83.24]%, against NCRB's 65.7, 66.6 and 67.2% (2022–2024). Its plea share of convictions today is 0.46 [0.03, 2.94]%, against NCRB's 1.01, 1.82 and 0.69%. The second figure was not targeted. The model is of the BNSS procedure, while the 2022 and 2023 disposals, and part of 2024, arose under the earlier Code, which had no thirty-day limit; Section A7.6 gives the matched run.

A7.2 Main results

	Completed pleas, % of eligible cases	Plea share of convictions, %	On bail, %	Custody-induced, % of pleas	Low-p types pleading, %	Net hours saved, %
Today: Chapter XXIII as enacted	0.3 [0.0, 1.7]	0.5 [0.0, 2.9]	0.0	—	—	0.0
Practice Direction inside Chapter XXIII	1.4 [0.2, 6.0]	2.0 [0.4, 9.0]	0.2	—	—	0.1
Practice Direction, Rule 14 route	23.8 [3.9, 52.6]	33.8 [6.5, 66.1]	18.3	22.0	12.1	19.9
— with a separate list (six-month trial)	36.3 [7.2, 58.7]	50.0 [11.7, 72.0]	34.7	5.9	16.4	31.4
Bill text alone	4.7 [0.7, 17.8]	7.2 [1.3, 27.6]	0.6	86.5	—	2.8
Full programme ($e = 0.5$)	43.4 [7.5, 62.0]	56.8 [10.6, 74.7]	41.5	5.9	11.0	36.2
Full programme, transported response	40.4 [32.1, 51.0]	52.6 [44.1, 62.1]	38.5	4.8	8.4	33.6

Median [*5th, 95th percentile*]. *"On bail" is the completed-plea rate among accused on bail. A dash means not computed for that regime.*

A7.3 Sensitivities

	Completed pleas, % of eligible cases	Plea share of convictions, %	On bail, %	Custody-induced, % of pleas	Low-p types pleading, %	Net hours saved, %
Rule 14 route with a flat one-third reduction	12.3 [2.4, 42.7]	19.0 [4.1, 56.9]	3.1	69.9	7.7	10.2
Rule 14 route, Volume 2 reductions, no probation chance	21.4 [3.8, 48.2]	30.9 [6.2, 62.8]	15.4	26.9	11.1	17.9
Rule 14 route, Volume 2 reductions, one-year trial	35.4 [7.0, 57.9]	48.7 [11.4, 71.5]	33.4	7.6	16.2	30.8
Rule 14 route, with the compensation order	14.9 [2.8, 39.8]	22.0 [4.3, 53.9]	9.0	38.6	7.8	11.9
— the same, with a separate list	27.8 [3.8, 50.0]	40.9 [6.8, 64.5]	25.2	9.7	8.5	24.6
Full programme, $e = 0$	35.8 [5.2, 60.0]	51.3 [7.9, 73.7]	32.2	9.5	11.0	30.9
Full programme, $e = 0.25$	39.5 [6.2, 61.2]	53.9 [9.3, 74.2]	36.9	7.7	11.0	33.5
Full programme, $e = 0.75$	46.6 [9.2, 62.9]	59.1 [12.7, 74.8]	44.7	4.2	11.0	37.6
— $e = 0$, no compensation order	47.6 [9.6, 67.4]	64.1 [15.9, 80.5]	44.7	5.8	20.0	40.5
— $e = 0.5$, custody taken out of the decision	41.3 [4.5, 60.8]	55.1 [6.5, 73.8]	41.5	—	6.1	34.4
— $e = 0.5$, one-year trial	41.3 [7.7, 61.9]	53.8 [11.1, 73.8]	38.4	8.5	11.9	33.9
— $e = 0.5$, today's wait	22.2 [4.4, 54.1]	29.1 [5.8, 67.2]	15.9	30.7	9.8	15.0
— $e = 0.5$, flat one-third reduction	36.7 [3.5, 59.7]	49.4 [5.3, 73.4]	33.2	10.5	6.3	29.5
— $e = 0.5$, no probation chance	40.7 [6.3, 60.1]	53.0 [9.4, 72.7]	37.9	7.7	9.3	32.9
— $e = 0.5$, Chapter XXIII route only	18.7 [2.9, 34.8]	24.4 [4.2, 44.8]	17.7	6.0	4.0	11.8
Transported response, Practice Direction access, today's wait	31.9 [22.2, 45.9]	45.0 [32.6, 60.5]	30.6	4.6	6.4	27.5
Full programme, transported response, custody out	38.5 [30.6, 48.6]	50.4 [42.0, 59.7]	38.5	—	6.4	32.1

Reading.

1. *The two behavioural rules agree on the full programme.* The cost-based rule adds a chance of probation to the Volume 2 reduction, which may count probation twice for the earliest pleas; without it the result is closer still to the transported response.
2. *What carries the result:* the size of the reduction, then the wait, then the evidence effect. The result moves little across the range of e. It moves a great deal with the wait and with the route: confined to Chapter XXIII, with the prosecutor's signature, the s. 289 exclusions and the sixty-day delay, the full programme gives less than half as many pleas.
3. *Custody.* With a small reduction most pleas are custody-induced, because few accused on bail find the plea worth taking. With the Volume 2 reductions and a short wait, accused on bail plead and the custody-induced share falls to about one plea in seventeen.
4. *Low-p types.* Large reductions draw them in, above the Guidelines' target of 5%. Taking custody out of the decision roughly halves the rate. So does the compensation order: it changes the difference between the cost of contesting and of pleading by $-(1 - \delta^T \Omega p)\, m_c$, which is largest in size where conviction risk is lowest. This is a comparative static of the model, not a ground for setting compensation. The Practice Direction regimes carry no compensation order; with one, they give fewer pleas

and fewer low-p pleas (rows above). A flat one-third reduction brings the rate to about 6% at a cost of about seven points of pleas.

5. *Compounding* diverts 3.4% of eligible cases under the Practice Direction and 4.8% under the full programme, where the shorter wait and higher conviction risk make contest costlier.

A7.4 IPC cases alone

If the conviction-risk range is set on NCRB's IPC series alone (completed-trial conviction rate 53–54%; model 52.38 [37.18, 69.05]%), conviction risk starts lower and the evidence effect matters more.

	Pleas, % of eligible cases	Share of convictions, %
Today	0.2 [0.0, 1.3]	0.5 [0.0, 2.9]
Practice Direction, Rule 14 route	17.5 [3.3, 52.8]	31.4 [6.5, 67.7]
— with a separate list	27.8 [3.8, 56.8]	47.5 [8.2, 74.5]
Full programme, $e = 0$	24.4 [2.7, 53.3]	43.9 [5.3, 74.3]
Full programme, $e = 0.5$	37.8 [5.4, 60.0]	53.5 [8.3, 74.9]
Full programme, $e = 0.75$	44.7 [9.0, 61.4]	58.4 [11.8, 75.9]
Full programme, transported response	40.6 [31.3, 51.6]	55.3 [45.9, 65.4]

A7.5 Barriers inside Chapter XXIII

Removing all five barriers inside Chapter XXIII, with a one-third reduction on the expected trial sentence, raises completed pleas by 6.49 [1.23, 21.69] points. The Shapley division of that gain, in points: access 2.83 [0.51, 10.02]; victim's signature 2.43 [0.47, 8.13]; predictable sentence 0.38 [0.02, 2.17]; investigating officer's signature 0.37 [0.02, 2.08]; realistic benchmark 0.15 [-1.29, 2.43]. Access is the largest item in 69% of runs.

A7.6 The statutory baseline and the s. 293 fractions

The earlier Code had no thirty-day limit. Removing the timing constraint on notice (every eligible accused may apply whenever he learns of the Chapter; comprehension and all else unchanged) gives the largest effect the limit can have in this model. Separately, the s. 293 fractions are taken as binding, with no fear that they will be exceeded.

	Pleas, % of eligible cases	Share of convictions, %
Today, BNSS (thirty-day limit)	0.28 [0.02, 1.74]	0.46 [0.03, 2.94]
Earlier Code: no thirty-day limit	1.20 [0.18, 5.25]	1.85 [0.31, 8.36]
BNSS, s. 293 fractions binding, no feared excess	0.37 [0.03, 2.02]	0.58 [0.05, 3.42]
Earlier Code, fractions binding	1.42 [0.30, 6.00]	2.18 [0.48, 9.79]
Practice Direction inside Chapter XXIII	1.36 [0.20, 6.03]	2.03 [0.37, 8.96]
— fractions binding	1.62 [0.28, 7.25]	2.55 [0.51, 10.24]

NCRB's plea share of convictions was 1.01% in 2022 and 1.82% in 2023, both under the earlier Code, and 0.69% in 2024, when the BNSS applied to new cases from 1 July. The run without the limit is the matched comparison for 2022 and 2023. The limit lowers the model's status-quo plea rate to about a quarter. Treating the fractions as binding changes little: uncertainty about the fraction is not what holds pleas back. NCRB does not publish disposals by procedural regime, so the 2024 comparison remains approximate.

A7.7 Custody-induced pleas as rates

	Per 100 eligible cases	Per 100 in custody	% of pleas
Rule 14 route, one-third reduction	5.7 [1.7, 13.4]	25.3 [8.1, 44.4]	69.9 [5.8, 100.0]
Rule 14 route, Volume 2 reductions	3.9 [1.1, 9.8]	18.0 [4.9, 31.5]	22.0 [2.3, 94.8]
— with a six-month trial	1.8 [0.5, 5.5]	8.0 [2.3, 18.2]	5.9 [1.0, 37.9]
Full programme	2.1 [0.6, 6.1]	9.2 [3.3, 19.9]	5.9 [1.2, 41.2]
Full programme, transported response	1.9 [0.9, 3.3]	7.8 [5.9, 9.6]	4.8 [2.2, 7.8]

Each figure is computed within a run and then summarised; the columns are not products of one another's medians. The fall is not only in the share: the number of custody-induced pleas per eligible case falls as well.

A7.8 Ranges, and efficiency and safety together

The main table gives medians for the safety and workload measures. Their 5th to 95th percentile ranges across the assumed box:

	Low-p pleading, %	Custody-induced, %	Hours saved, %
Practice Direction, Rule 14 route	2.8 to 41.9	2.3 to 94.8	2.9 to 44.1
— with a six-month trial	1.5 to 45.8	1.0 to 37.9	7.1 to 49.4
Full programme	0.7 to 37.5	1.2 to 41.2	2.9 to 54.4
— custody taken out of the decision	0.1 to 32.1	—	-0.3 to 53.8
— flat one-third reduction	0.3 to 27.5	1.7 to 86.3	-2.3 to 53.0
Full programme, transported response	4.8 to 12.4	2.2 to 7.8	25.4 to 43.7

Because every regime is evaluated on the same runs, the share of runs in which a regime meets several tests at once can be counted directly: (i) plea share of convictions at least 43%; (ii) at most 5% of low-p types pleading; (iii) hours saved net; (iv) at most 10% of pleas custody-induced.

Share of runs, %	(i)	(ii)	(iii)	(iv)	All four
Practice Direction, Rule 14 route	36	18	100	32	0
— with a six-month trial	64	18	100	68	0
Full programme	70	29	96	71	4
— custody taken out of the decision	66	45	95	100	13
— flat one-third reduction	58	44	93	48	7
Full programme, transported response	97	7	100	100	6
— custody taken out of the decision	92	28	100	100	22

These are shares of the assumed ranges and weights, not probabilities for India. They show that the medians of the main table are not reached together in the same runs: where pleas are many, low-p pleading is usually above target. On the IPC box with a one-third reduction the full programme gives 29.4% of cases, 42.8% of convictions, 4.8% of low-p types pleading, and net hours saved of 21.0% [-10.1, 47.3].

Evidence measures and weak cases. The published regimes leave conviction risk in low-p cases unchanged. If the measures also closed a share e_L of the acquittal gap in those cases, the share of low-p types pleading under the full programme would be: $e_L = 0$: 11.0%; $e_L = 0.25$: 18.5%; $e_L = 0.5$: 30.8%.

A7.9 The ceiling on the reduction

Each case keeps its own Volume 2 reduction, capped at a ceiling: $d = \min(d_{\text{Vol. 2}}, \text{ceiling})$; 100 is the schedule as drafted. Each design is run with custody in the accused's decision and with it taken out. Columns: completed pleas, % of eligible cases; plea share of convictions, %; low-p types pleading, % (medians); then the share of runs, %, meeting test (i), test (ii), and all four tests of Section A7.8.

	Ceiling, %	Pleas	Conv.	Low-p	(i)	(ii)	All four
Practice Direction, six-month trial	33	30.6	43.5	9.5	51	32	2
	50	33.1	46.9	11.6	58	26	1
	60	34.5	48.0	13.0	61	23	0
	100	36.3	50.0	16.4	64	18	0
— custody out of the decision	33	27.4	40.6	3.4	44	57	7
	50	30.4	43.5	5.2	52	49	6
	60	32.2	45.2	6.2	55	43	5
	100	34.5	47.4	10.7	60	32	2
Full programme	33	35.9	48.3	6.1	57	44	7
	50	40.0	53.6	7.5	63	38	6
	60	41.9	54.7	8.5	66	34	5
	100	43.4	56.8	11.0	70	29	4
— custody out of the decision	33	32.0	44.1	1.4	51	71	23
	50	36.9	50.0	2.5	58	63	23
	60	38.7	52.1	3.4	62	59	22
	100	41.3	55.1	6.1	66	45	13
Full programme, transported response	33	28.0	37.6	4.2	16	72	4
	50	33.8	44.3	5.0	62	50	19
	60	36.0	47.2	5.7	80	34	19
	100	40.4	52.6	8.4	97	7	6
— custody out of the decision	33	25.9	35.0	2.1	10	99	9
	50	31.7	42.3	3.0	43	92	36
	60	33.9	45.1	3.7	66	79	46
	100	38.5	50.4	6.4	92	28	22

A ceiling without the custody measure does not bring the median low-p rate to the target at any level above 20%. With custody out of the decision, ceilings of 50% to 60% keep the median conviction share above the benchmark under the cost-based rule and near or above it under the transported response, with low-p pleading between 2.5% and 3.7%. On the IPC box the ranking is the same (full programme, custody out, ceiling 60%: 31.9% of cases, 46.8% of convictions, 2.1% of low-p types pleading, all four tests in 21% of runs). "Custody out of the decision" is an idealisation of bail before plea: it assumes release is in fact granted and removes custody from the accused's calculation entirely.

A7.10 Can the courts deliver the shorter wait?

The main results take a six-month trial as given. Two checks ask what present capacity allows. Both assume that today's wait is a steady state, so that capacity equals the work arriving each year. If the backlog is in fact growing, capacity is lower and both checks are optimistic.

One queue, with feedback. There is no separate list. The backlog starts at today's level, the wait is six months plus backlog divided by capacity, pleas respond to the wait, and each year the backlog changes by the hours the regime's cases need less capacity. "Off" holds the wait at today's level; "on" is the state after twelve years.

	Pleas, % of cases, off	Pleas, % of cases, on	Share of convictions, %, off	Share of convictions, %, on	Wait, years (median)	Runs never under one year, %
Practice Direction, Rule 14 route	23.9 [3.9, 53.4]	32.7 [4.1, 58.2]	34.0 [6.3, 66.6]	44.1 [6.5, 71.1]	5.0 → 2.0	62
Full programme, $e = 0$	18.7 [3.8, 50.5]	25.3 [3.9, 59.6]	27.6 [6.0, 65.8]	37.1 [6.0, 73.1]	5.0 → 2.6	66
Full programme, $e = 0.5$	22.1 [4.1, 54.6]	31.4 [4.2, 61.6]	28.8 [5.3, 67.4]	40.3 [5.2, 74.5]	5.0 → 2.5	65

Feedback is real in this model but slow. Over twelve years it adds about nine points of pleas under the full programme and roughly halves the wait; in about two runs in three the wait does not fall below one year. This is a like-for-like measure of feedback, with the population and parameters held fixed. It is much smaller than the gap between this model and the simulator.

A separate list. A list of fresh filings held at six months needs a share of present capacity equal to the regime's hours per case at that wait. The old stock is then cleared by what remains.

	Fresh list, % of present capacity	Years to clear the old stock
Practice Direction, Rule 14 route	68.6 [50.6, 92.9]	14.4 [5.3, 66.3]
Full programme, $e = 0.5$	63.9 [45.6, 97.1]	12.7 [4.7, 63.7]

The six-month list is feasible within present capacity in nearly all runs, because pleas free more hours than the list consumes. Its cost falls on the old cases, which wait a decade or more unless capacity is added or the old stock is offered its own terms. The six-month figures in Section A7.2 are therefore figures for fresh filings.

A7.11 What drives the bands

Rank correlations between each parameter and completed pleas, across runs (the five largest): Practice Direction, Rule 14 route: stigma of conviction (-0.49); stigma retained after a plea (-0.46); litigation cost of contest (+0.40); discount factor (+0.29); wait to verdict (-0.26). Full programme: stigma retained after a plea (-0.59); litigation cost of contest (+0.41); stigma of conviction (-0.33); discount factor (+0.24); conviction risk (+0.23). The stigma that a plea conviction retains, the cost of litigation and the stigma of conviction matter most; none is observed in India. With every range halved about its midpoint:

Pleas, % of eligible cases	Full ranges	Ranges halved
Practice Direction, Rule 14 route	23.8 [3.9, 52.6]	22.3 [9.1, 37.9]
— with a six-month trial	36.3 [7.2, 58.7]	36.5 [20.5, 47.9]
Full programme	43.4 [7.5, 62.0]	42.3 [22.4, 55.7]
Full programme, transported response	40.4 [32.1, 51.0]	40.9 [32.8, 49.1]

The medians barely move and the bands narrow in proportion, so the medians are not an artefact of the width of the ranges. Parameters are drawn independently; dependence between them has not been tested.

A7.12 From eligible cases to all convictions

Let q_E be the plea share of convictions among eligible cases and q_N the share among offences outside the model, with w the eligible cases' share of all convictions. Then the share for all offences is $q = w\,q_E + (1 - w)\,q_N$. Neither w nor q_N is known. If offences outside the model produced no pleas at all, a full-programme q_E of 57% would be 51% of all convictions at $w = 0.9$ and 40% at $w = 0.7$. Since the ordinary guilty plea is open for every offence, q_N would not be zero under the Rule 14 route.

A8 Settlement, timing and the acquittal exits

A companion model adds the exits that end a case in acquittal: compounding, withdrawal from prosecution (ss. 280, 360), quashing on settlement (s. 528) and acquittal after witnesses turn hostile. The victim's reservation falls as the case ages; the plea window closes at thirty days; the exits stay open to judgment. Three results follow.

1. *Delay is the accused's best strategy.* He lets the plea window lapse and settles later, for less, for an acquittal.

2. *The exits are substitutes.* Closing withdrawal and quashing moves settlements to hostile witnesses; pricing hostility alone moves them to withdrawal. Only closing all of them shrinks acquittal-by-settlement substantially.
3. *Closing the exits does not by itself create pleas.* Most would-be settlers contest, hours rise and victims receive less, unless a conviction route is open beyond thirty days and trials are prompt.

The level of the hostile-witness exit in this model is not reconciled with NCRB's conviction rate and is not quoted (Section A11). The following can be tested now from case histories: the compounding hazard rises with case age and recorded compensation falls; withdrawals and quashings in non-compoundable offences cluster after the plea window; hostile-witness acquittals rise with delay; accused in custody settle earlier.

A9 The whole-court simulator

The simulator is a separate, heterogeneous-agent model of the court as a whole. It follows the accused, the complainant who may give up, the witness who may be bribed or threatened, and co-accused; the court's queue is the aggregate of their choices and sets the wait. It embeds the quintile sentencing engine of the National Sentencing Guidelines and is closed by a twelve-year docket. It is calibrated to single values, with bounds obtained by moving every assumed constant by its half-range. Run on foreign inputs, its decision rule reproduces the United States and Italy; it does not reproduce England and Wales. Its figures in this volume are:

	Pleas, % of all cases	Pleas, % of convictions
Today	0.04	0.2
Pilot by Practice Direction, today's wait	9.3 (bounds 0.6–93)	25
Pilot with a separate list	48–74	78–90
Bill, at the six-month wait its docket produces	73.8 (bounds 46–98)	87

The two models differ at once in population, in losses to notice, comprehension and the court's check, in behaviour and in queue feedback, so the gap between them is not attributed to any one of these. The simulator is interactive at `satyug-sim.pages.dev`.

A10 Pilot protocol

Primary estimand. The intention-to-treat effect over all eligible accused at first notice.

What may be randomised. Administrative matters only: plain-language notice; the order and timing of legal-aid counselling; scheduling; a leaflet stating the likely sentence. Never rights, eligibility, sentences, bail, judicial decisions, victim participation or access to legal aid.

Designs. A stepped-wedge roll-out across courts in random order; comparisons at the s. 479 custody thresholds; bench-assignment comparisons only where assignment is as good as random. The seven-year line coincides with other thresholds and is treated as a comparison of offence classes.

Outcomes. Applications and completed pleas over eligible cases; plea convictions over all convictions; judicial hours from hearing histories and time-use samples; contested-trial rate; duration; indicators of involuntary or wrongful pleas (withdrawal attempts, challenges, factual-basis failures, an audit sample); and the recorded reason for failure at each stage.

What would count against the model. Each prediction is conditional on the intervention having been delivered, on a study large enough to detect the effect, and on comparable groups. (i) Where a participant's agreement is sometimes withheld, requiring it lowers completion among those who apply; allowance is needed for any change in who applies. (ii) Completion is lower where the victim must agree than where the victim is heard. (iii) Holding the plea sentence fixed, applications rise where the expected sentence after trial is higher. (iv) Where access is assured, no later stage blocks the plea and the plea is worth attempting for enough of those

newly reached, applications rise. A failure of any of these counts against the mechanism named. Separately, as a test of the policy and not of the model, a rise in pleas with no saving in hours, or concentrated among detained accused, does not count as success.

Size. Detecting a rise in completed pleas from 1% to 3% of eligible cases (80% power, 5% two-sided) needs about 768 eligible cases per arm, and about 1,500 to 8,400 with clustering by court.

A11 Limits and open questions

1. *No Indian case-level data.* The quantities that matter most have no Indian estimate: the ratio of sentences to maxima, conviction risk by offence, victim participation, prosecutor agreement, custody status at charge, the compoundable share, hours by path. Stage-wise court records would supply most of them.
2. *The evidence effect is assumed.* Data on acquittals after hostile testimony, and conviction rates before and after the complaint routing of Rule 14, would estimate it.
3. *The Volume 2 reductions are applied on an assumed mix of scores and stages.* An Indian mix would change the mean reduction. The low-p pleading rates make the ceiling on the earliest reduction a decision to be taken with these figures in view.
4. *The route.* The results depend on a known sentence on an ordinary guilty plea being lawful by Practice Direction.
5. *The hostile-witness exit.* In the companion model this exit is about a fifth of eligible cases today; counted as NCRB counts acquittals, that implies a conviction rate near 45%, against 66–67%. Either the exit is too large, or few recorded acquittals are of this kind, or it is counted twice. It awaits Indian data on hostile acquittals.
6. *Applications.* The number of applications is sensitive to the error term in the choice rule. Completed pleas are reported.
7. *Simplifications.* Stage outcomes are independent given the parameters. Release at the s. 479 threshold is automatic, which understates custody. Sentence risk at trial is ignored while sentence risk on a plea is not. Compounding ignores the court-permission requirement. Utility is in years-equivalent, so compensation and wealth are in the accused's units.
8. *The statutory baseline.* The run without the thirty-day limit (Section A7.6) matches the 2022 and 2023 figures in order of magnitude. A closer comparison needs disposals by procedural regime, which NCRB does not publish.
9. *The six-month wait* is an input to the main results. Section A7.10 shows it is affordable for fresh filings at present capacity only if that capacity matches today's inflow, and that the old stock then waits a decade or more. The gap between this model and the simulator is not explained by feedback as this model measures it.
10. *Cross-country evidence* is correlational: 52 observations, and India's figure counts only Chapter XXIII disposals where common-law figures count every guilty plea.
11. *Review.* The model, code and figures were checked by an independent reviewer. They have not been replicated by a second party.

A12 Reproducibility

All results use 400 runs of 6,000 cases with a fixed seed. The full derivations, the record of independent review, the model code, result files, regression tests and Lean sources are in the Team Satyug working paper and source package of 12 November 2026, available on request from `help@teamsatyug.in`.

A13 The simple model: symbols and proofs

Symbol	Meaning
S_p, S_0	sentence on a plea; sentence on conviction after trial
s_1, x	time already served; further custody before verdict if he contests
K, ρ	burden of a conviction record; share of it a plea conviction carries
m_p, m_0	fine or compensation on a plea; on conviction after trial
p, T, δ	chance of conviction; wait in years; yearly discount factor
c_L, c_n, Δ_f	cost of contesting; of applying; of a failed attempt (none negative)
Π	chance that an application yields a lawful disposition for him to accept
$(a)^+$	a if positive, otherwise zero

Rule. The gain from applying is $G = \Pi(C - P) - (1 - \Pi)\Delta_f - c_n$. He applies when $G > 0$. For $\Pi > 0$ this is the same as $P < C$ together with

$$\Pi > \frac{\Delta_f + c_n}{C - P + \Delta_f}.$$

Proof. $G = \Pi(C - P + \Delta_f) - \Delta_f - c_n$. If $P \geq C$ then $G \leq 0$. If $P < C$ the bracket is positive; divide. □

Reduction needed. For an accused on bail with no time served, and $S_p = (1 - r)S_0$, the plea is attractive when

$$r > 1 - \frac{c_L + \delta^T p\,(S_0 + K + m_0) - \rho K - m_p}{S_0}.$$

The right-hand side rises with T and falls with p.

Evidence threshold. Against trial, $P < C_{trial}$ exactly when

$$p > p^* = \frac{P - c_L - x}{\delta^T\left[(S_0 - s_1 - x)^+ + K + m_0\right]}.$$

p^* falls as P falls, and is negative when $P < c_L + x$.

Custody. One further day of detention adds one day to the cost of contesting and removes at most one day from the sentence, which is weighted by $\delta^T p \leq 1$. So C_{trial} never falls as x rises. The release test sets $x = 0$ and leaves s_1 unchanged.

Work. A completed plea replaces h_{alt} by h_{plea}; a failed attempt adds h_{fail}. Rearranging gives the threshold stated under "Why the court's work falls'' in Section 4.

The rule and every figure in Section 4 are reproduced by a short open program (`plea_core.py`, version plea-core 2026-11-14.2, with 27 tests) and can be tried in the accompanying calculator.

References

Bharatiya Nagarik Suraksha Sanhita, 2023 (Gazette of India Extraordinary, 25 December 2023), ss. 215, 280, 289–300, 346, 359, 360, 379, 383, 395, 479, 528, 529.

Dobbie, W., Goldin, J., and Yang, C. (2018), "The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges", *American Economic Review* 108(2): 201–240.

Helm, R. K., and Zottoli, T. M. (2026), "Extending the Fuzzy-Trace Theory Informed Model of Guilty Plea Decision Making: Accounting for Factual Guilt and Innocence", *Law and Human Behavior* 50(2): 205–224.

National Crime Records Bureau, *Crime in India* 2022, 2023, 2024, Vol. III, Tables 18A.2 and 18A.4.

Paolini, G., Kantorowicz-Reznichenko, E., and Voigt, S. (2025), "Plea Bargaining Procedures Worldwide: Drivers of Introduction and Use", *Journal of Empirical Legal Studies* 22(1): 27–75.

Team Satyug (2026), *National Sentencing Guidelines*, Volumes 1 and 2; *Obstruction of Justice* (*Satyug*) *Bill and Research Note*, Volume 5.