Den dynamiske rettsstaten

Forsinkelse, tillit og stabilitet i styringen av rettshåndhevelsen

En systemteoretisk, spillteoretisk og algoritmisk modell for kriminalitetskontroll

Kjell Tore Guttormsen*

November 2026 · versjon 2 · konseptnotat

Sammendrag

Tanken bak den dynamiske rettsstaten er enkel: Innsatsen for å håndheve lover og regler bør løpende stå i forhold til hvor mye, og hvor alvorlig, de faktisk brytes. Øker bruddene, bør innsatsen øke. Går de ned, kan ressursene flyttes dit behovet er størst. Straffebud og strafferammer ligger fast, mens regelverket for prioritering, prosess og valg av reaksjon innenfor lovens rammer må gi den fleksibiliteten som gjør dynamisk håndhevelse mulig. Artikkelen gir prinsippet et presist innhold og undersøker hva som skal til for å gjennomføre det stabilt, legitimt og i hele straffesakskjeden. Flere henleggelser av kapasitetsgrunner i politi og påtale, målbrudd på saksbehandlingstid i ankeinstansen og en voksende soningskø tyder på at politi, påtalemyndighet, domstoler, kriminalomsorg og lovgiver må ses og styres som ett system. Gjennomføringen er et styringsproblem med tilbakekobling, der dagens tiltak påvirker morgendagens kriminalitet og dermed framtidige prioriteringer. Alle tiltak — innsats, reaksjonsformer, kapasitet, prosessuell kvalitet og lovbestemt handlingsrom — forstås som bevisste responser på det observerte bildet, formulert som én beslutningsregel, $T \leftarrow f(K, \mathcal{O})$, fra kriminalitetsbildet og observasjonene til en samlet tiltaksvektor. Rammeverket forener rettsøkonomi, reguleringsteori, spillteori, teori om prosessuell rettferdighet og skrankeoptimering: et alvorlighetsvektet mål på faktisk kriminalitet, et makroøkonomisk drevet pådrag, en tidsforsinket styringsregel, strategisk forskyvning mellom lovbruddskategorier og en tillitsdynamikk der både for lite og uforholdsmessig mye håndhevelse svekker legitimiteten. Vi viser analytisk at stabiliteten avhenger av forholdet mellom politisk reaktivitet, tidsforsinkelse og samfunnets egen dempning, og utleder en tommelfingerregel for hvor kraftig en proporsjonal respons kan være ved en gitt forsinkelse. Simuleringer viser hvordan en kraftig reaksjon på forsinkede tall gir vedvarende svingninger mellom under- og overhåndhevelse, hvordan en avdempet regel fjerner dem og verner tilliten, og hvordan prognosebasert styring i tillegg fjerner det varige avviket. Lovgivningen er sløyfens langsomste ledd; handlingsrommet bør bygges inn på forhånd gjennom rammelover, forsøkshjemler og solnedgangsklausuler innenfor Grunnlovens og EMKs faste grenser. Med utgangspunkt i *Theory of Constraints* modelleres straffesakskjeden som et kønettverk koblet til dynamikken, med domstolen som gjennomregnet eksempel på et ledd som binder: økt politiinnsats ved en mettet domstol omsettes i restanser, frafall og lavere tillit, underordning må kombineres med kapasitetsutvidelse, og en mettet domstol kan maskere en ustabil sløyfe; resultatene står seg med politimetning og prioriterte domstolskøer, mens selvforsterkende rekruttering kan gi vippepunkter som krever tidlig og tilstrekkelig respons. Vi foreslår en syklisk beslutningsmetode som forener responsprinsippet med *Theory of Constraints* og reguleringsteorien: observer og estimer, sammenlign alvorlighetsvektet med målene, identifiser den bindende skranken, velg type tiltak etter skranken, dimensjoner avdempet og med prognose, kontroller de rettsstatlige skrankene, og gjenta. Når også underordning, kapasitet og kvalitet gjøres til responser på observasjoner, gir desentraliserte regler der hver aktør svarer på sitt eget signal, i den stiliserte simuleringen en kapasitet om lag 30 prosent over behovet og en dempet

* Idéen, problemstillingen og konseptet «den dynamiske rettsstaten» er forfatterens, og forfatteren har styrt arbeidet og tatt de faglige valgene. Teksten, den matematiske formaliseringen, simuleringene og figurene er utarbeidet med omfattende bruk av kunstig intelligens (Claude, Anthropic), også til fagfelle- og kvalitetsgjennomganger. Forfatteren har ansvaret for innholdet.

kapasitetssvingning over seks til sju år, fordi innstrømningen til en overbelastet domstol er et signal flaskehalsen selv har blåst opp; en koordinert regel med felles skrankebilde og prognose gir nesten samme utfall som vedtak fattet med full kjennskap til likevekten, med mindre kapasitet og uten svingninger. Til slutt beskrives kalibrering og bruk som beslutningsstøtte i budsjett- og prioriteringssyklusen, og de styringsmessige og etiske skrankene for slik bruk. Simuleringene bygger på stiliserte parametere og illustrerer mekanismer; de er ikke prognoser for norske forhold. Artikkelen er et konseptnotat og en invitasjon til videre forskning: modellen er ikke kalibrert på norske data, hvilket ledd i kjeden som faktisk binder, er et empirisk spørsmål, og de samfunnsøkonomiske størrelsene er ikke tallfestet. Kjente begrensninger og en veiplan står i avsnitt 10.

Nøkkelord: rettshåndhevelse; avskrekking; reguleringsteori; tidsforsinkede differensialligninger; kriminalitetsforskyvning; legitimitet; modellprediktiv styring; straffesakskjeden; *Theory of Constraints*; lovgivningsmessig handlingsrom.
JEL-koder: K42, H50, D73, C61, C73.

1 Innledning og problemstilling

Tanken bak den dynamiske rettsstaten er enkel: Innsatsen for å håndheve lover og regler bør løpende stå i forhold til hvor mye, og hvor alvorlig, de faktisk brytes. Øker bruddene, bør innsatsen øke. Går de ned, kan ressursene flyttes dit behovet er størst. Dette er artikkelens utgangspunkt og dens sentrale budskap. Resten av artikkelen handler om hva som skal til for å gjennomføre prinsippet stabilt, legitimt og i hele straffesakskjeden.

Artikkelens utgangspunkt

Definisjon 1 (Den dynamiske rettsstaten)**.** Innsatsen for å håndheve lover og regler bør løpende stå i forhold til hvor mye, og hvor alvorlig, de faktisk brytes. Øker bruddene, bør innsatsen øke. Går de ned, kan ressursene flyttes dit behovet er størst.

Presisering. Straffebudene og strafferammene ligger fast: hva som er straffbart, og hvor strengt det kan straffes. Det som tilpasses, er prioriteringer, innsats, organisering, prosess og valget av reaksjon innenfor lovens rammer, etter kriterier som er fastsatt på forhånd, og regelverket må gi den fleksibiliteten som gjør dette mulig.

Responsprinsipp. Alle slike tilpasninger — uansett type og omfang, fra politiets daglige prioriteringer til kapasitet og lovbestemt handlingsrom — er bevisste svar på det *observerte* bildet av bruddene og av tilstanden i straffesakskjeden, ikke engangsvedtak løsrevet fra utviklingen (avsnitt 4.1).

I sin enkleste form kan prinsippet skrives som at myndighetenes tiltak er en funksjon av kriminalitetsbildet, $T = f(K)$. Men tiltakene kan bare bygge på det som faktisk observeres, og de settes og justeres løpende. Prinsippet uttrykkes derfor mer presist som

$$T \leftarrow f(K, \mathcal{O}), \tag{1}$$

der K er kriminalitetsbelastningen, $\mathcal{O}$ er observasjonene som foreligger — registrerte lovbrudd, offer- og innbyggerundersøkelser, køer og saksbehandlingstid i straffesakskjeden, tillit og makroindikatorer — og pilen markerer at tiltakene T settes og oppdateres som et bevisst svar på dem. Her står T for alle tiltak, uansett type og omfang. Fordi K selv ikke observeres, inngår den i praksis gjennom estimatet $\hat{K}$ som dannes fra $\mathcal{O}$ (avsnittene 2.4 og 4.1). Artikkelen gir de tre leddene i definisjon 1 et presist innhold. «Hvor mye» måles ved aktivitetsnivået N_i, det faktiske antallet lovbrudd i hver kategori i; «hvor alvorlig» ved vektene $A_i S_i^{\theta}$ for alvorlighet og skadeomfang; sammen gir de den vektede kriminalitetsbelastningen $K_{\text{obj}} = \sum_i N_i A_i S_i^{\theta}$ i ligning (2). «Løpende stå i forhold til» betyr at innsatsen T settes etter en styringsregel som svarer på avviket mellom det målte kriminalitetspresset og et samfunnsmessig akseptabelt

nivå (ligning (11)). Politiets innsats er bare ett av virkemidlene: reaksjonsformer, kapasitet i domstolene, prosessuell kvalitet og lovgivers handlingsrom er tiltak av samme art med lengre ledetid, og hvert av dem er et svar på det som observeres, ikke på det som skjer. Sammenhengen (1) formaliseres i avsnitt 4.1 som en beslutningsregel fra observasjoner til en samlet tiltaksvektor, og avsnitt 7.7 gir en metode for å komme fram til tiltakene. Prinsippet er modellens styringsprinsipp, ikke et tillegg til den.

Rettsstatlig avgrensning. Prinsippet må avgrenses mot det som ligger fast i en rettsstat. Hva som er straffbart, og hvilke strafferammer som gjelder, fastsettes av lovgiver og skal være kjent på forhånd (Grunnloven §§ 96 og 97); disse størrelsene justeres ikke etter kriminalitetsbildet. Likhet for loven (Grunnloven § 98: «Alle er like for loven. Intet menneske må utsettes for usaklig eller uforholdsmessig forskjellsbehandling») og kravet om forutberegnelighet innebærer at en dynamisk innsats må begrunnes i saklige forskjeller i omfang og alvorlighet, ikke i hvem som rammes. Innenfor disse grensene er prioritering av håndhevelsesinnsatsen ikke noe nytt: Riksadvokaten fastsetter allerede årlig mål og prioriteringer for straffesaksbehandlingen. Det prinsippet tilfører, er at prioriteringen gjøres systematisk, etterprøvbar og koblet til den faktiske utviklingen i lovbruddene, med en målt størrelse å styre etter, en kjent reaksjonsstyrke og en dokumentert begrunnelse for endringer. Spenningen mellom dynamikk og forutberegnelighet drøftes nærmere i avsnitt 7.1 og avsnitt 10.

Utviklingen i straffesakskjeden. Spørsmålet er ikke bare teoretisk. Ifølge Statistisk sentralbyrå (2026) ble 36,8 prosent av de etterforskede lovbruddene oppklart i 2025, mot nær 50 prosent på 2010-tallet. Om lag 30 300 lovbrudd ble henlagt på grunn av manglende saksbehandlingskapasitet, mer enn dobbelt så mange som i 2021, og antallet uoppklarte lovbrudd var 195 000, en økning på 3,7 prosent.

Tallene må leses med varsomhet. Fallet i oppklaringsprosent er ikke i seg selv et mål på kapasitetspress: SSB peker på at andelen avhenger sterkt av hvilke lovbrudd som inngår, og at endringer i kriminalitetsbildet har bidratt vesentlig. Antallet narkotikalovbrudd, som har svært høy oppklaringsandel, er redusert med nesten 30 000 på ti år, mens uoppklarte bedragerier, tyverier og skadeverk har økt, og antallet forelegg og tiltaler er redusert med rundt 40 prosent i samme periode. Dette er måleproblemet i avsnitt 2.4 i praksis: registrerte tall for kontrollavhengige lovbrudd følger innsatsen. SSB tar også forbehold om at lovendringer fra 1. juli 2022 og nye retningslinjer ser ut til å ha påvirket henleggelsespraksisen. Det tydeligste tegnet på kapasitetspress er kapasitetshenleggelsene, og de skjer i politi og påtale.

Bildet nedstrøms er sammensatt. Tingrettene lå i 2025 nær Stortingets mål: 3,1 måneder i gjennomsnitt for meddomsrettssaker mot et mål på 3, og 0,3 måneder for enedommersaker, men restansen i meddomsrettssaker økte med 15 prosent, og lagmannsrettene brukte 6,2 måneder på bevisanker mot et mål på 3 (Domstoladministrasjonen, 2026). Dommerforeningen melder at mange domstoler opplever økende sakstilfang og saksbehandlingstid (Den norske dommerforening, 2025). I straffegjennomføringen sto 657 ubetingede dommer i soningskø ved utgangen av oktober 2026, mot 284 ett år tidligere (Kriminalomsorgen, 2026).

Artikkelen tar ikke stilling til hvor mye hver forklaring betyr. Det tallene viser samlet, er at presset oppstår i flere ledd, at det flytter seg, og at leddene henger sammen: det som ikke blir etterforsket, når aldri domstolen, og det som blir pådømt, må kunne gjennomføres. Hvilket ledd som til enhver tid er den bindende skranken, er et empirisk spørsmål, og i dag peker tallene mot politi og påtale og mot straffegjennomføringen snarere enn mot tingrettene. Slike utfordringer lar seg derfor ikke løse ved å styrke ett ledd om gangen. Politi, påtalemyndighet, domstoler, kriminalomsorg og lovgiver må ses og styres som ett system, med en felles måling av bruddene og en samordnet dimensjonering av kapasiteten. Denne helhetlige, systemiske tilnærmingen er forutsetningen for at innsatsen skal kunne følge bruddene slik definisjon 1 krever, og den ligger til grunn for resten av artikkelen.

Hva prinsippet krever. Prinsippet er lett å slutte seg til og vanskelig å gjennomføre, fordi sammenhengen (1) i praksis verken er statisk, lineær eller isolert fra resten av samfunnet. Politiet og rettsvesenet arbeider i en lukket *tilbakekoblingssløyfe*: dagens tiltak påvirker morgendagens kriminalitet, som igjen styrer framtidige ressursprioriteringer. Når denne sløyfen utsettes for institusjonelle *tidsforsinkelser* — fra et lovbrudd begås til statistikken foreligger, budsjettet er vedtatt og patruljene er satt inn — oppstår et klassisk styringsproblem: *pendelsvingninger* mellom under- og overhåndhevelse, der satsinger avløses av kutt og kutt av nye satsinger uten at dette er tilsiktet. Samtidig er verken lovbrytere eller lovlydige borgere passive. Lovbrytere tilpasser seg strategisk, blant annet ved å forskyve aktivitet til områder med lavere oppdagelsesrisiko. Befolkningens opplevde trygghet og tillit til politiet og rettsstaten — løpende målt i offentlige spørreundersøkelser — virker både som en normativ skranke og som en operativ multiplikator for politiets effektivitet, fordi tillit påvirker viljen til å anmelde og samarbeide. Videre drives grunnstrømmen av kriminalitet av makroøkonomiske forhold som arbeidsledighet, ulikhet og kjøpekraft. Endelig virker ikke politiets innsats i et tomrom: den må omsettes i reaksjoner gjennom påtalemyndighet, domstoler og straffegjennomføring, innenfor et lovverk som selv endres langsomt. Dersom lovverket ikke gir noe handlingsrom og et ledd nedstrøms er en flaskehals, blir dynamisk styring av politiinnsatsen i beste fall virkningsløs og i verste fall en kilde til restanser og lengre forsinkelser.

Av dette følger fire forutsetninger som må være oppfylt for at innsatsen skal kunne følge bruddene slik definisjon 1 krever, og som strukturerer artikkelen:

(1) *Måle bruddene riktig* — ikke bare det som blir anmeldt eller registrert (avsnitt 2.4).

(2) *Reagere avdempet nok* til å unngå overreaksjoner, fordi tallene kommer med forsinkelse (avsnittene 4 og 6).

(3) *La hele straffesakskjeden følge med* — fra etterforskning til dom og gjennomføring (avsnittene 7.2–7.6).

(4) *Gi lovverket et avgrenset handlingsrom på forhånd*, innenfor Grunnlovens og menneskerettighetenes faste grenser (avsnitt 7.1).

For å analysere disse forutsetningene utvikler artikkelen et samlet matematisk rammeverk som forener fem teoretiske tradisjoner:

(i) **Rettsøkonomi** (*law and economics*): rasjonell avskrekking og minimering av samlede samfunnskostnader (Becker, 1968; Polinsky og Shavell, 2000).

(ii) **Regulerings- og systemteori** (*control theory, system dynamics*): tilbakekobling, PID-regulering og tidsforsinkede differensialligninger (Minorsky, 1922; Wiener, 1948; Forrester, 1961; Åström og Murray, 2021), som forklarer hvordan treghet i rettsapparatet kan gi ustabile sykluser og overreaksjoner.

(iii) **Spillteori**: strategisk samspill mellom lovbrytere og rettshåndhevere (Tsebelis, 1989), inkludert forskyvning mellom lovbruddskategorier (Guerette og Bowers, 2009).

(iv) **Prosessuell rettferdighet og legitimitet** (*procedural justice*): hvordan tillit og opplevd rettferdighet påvirker viljen til å anmelde, samarbeide med politiet og etterleve loven (Tyler, 1990; Sunshine og Tyler, 2003).

(v) **Skrankeoptimering og algoritmisk styring**: rettshåndhevelse som et dynamisk skrankeoppfyllelsesproblem (CSP) og som modellprediktiv styring (MPC) under harde konstitusjonelle, budsjettmessige og kapasitetsmessige skranker (Rossi mfl., 2006; Rawlings mfl., 2017; Altman, 1999).

Bidrag. Artikkelens overordnede bidrag er å gi prinsippet i definisjon 1 et presist, etterprøvbart innhold og å vise hvilke betingelser som må være oppfylt for at det skal kunne gjennomføres. Konkret gir artikkelen seks bidrag. *For det første* samler vi elementer som i litteraturen oftest behandles hver for seg, i én konsistent tilstandsrommodell der prinsippet er styringsregelen, formulert som en beslutningsregel fra observasjoner til en samlet tiltaksvektor (avsnittene 2–6). *For det andre* utleder vi en eksakt stabilitetsbetingelse for den tidsforsinkede styringssløyfen (proposisjon 1) og en tommelfingerregel for hvor kraftig en proporsjonal respons kan være ved en gitt forsinkelse (korollar 1). Betingelsen viser at det ikke er reaktiviteten eller forsinkelsen hver for seg som avgjør, men forholdet mellom dem og samfunnets egen dempning; den angir dermed hvor avdempet innsatsen må følge bruddene for at tilpasningen ikke selv skal skape svingninger. *For det tredje* viser vi at måleproblemer — registrert kriminalitet avhenger av både tillit og håndhevelse — kan skape selvforsterkende sløyfer som ikke er synlige når registrerte tall behandles som faktisk kriminalitet (avsnitt 2.4). Innsatsen må følge de *faktiske* bruddene, ikke registreringen av dem. *For det fjerde* utvider vi modellen med lovverket og straffesakskjeden (avsnitt 7): lovgivningen behandles som sløyfens langsomste ledd, og vi argumenterer for at et lovbestemt handlingsrom må bygges inn på forhånd innenfor konstitusjonelle grenser, mens domstolene analyseres som et gjennomregnet eksempel på en flaskehals ved hjelp av køteori og *Theory of Constraints*. Domstolskøen kobles til kriminalitets- og tillitsdynamikken, slik at virkningen av flaskehalsen på styringen kan analyseres og simuleres (avsnitt 8.2). Robustheten undersøkes mot tre utvidelser: et politi med begrenset kapasitet, en domstol som prioriterer alvorlige saker, og selvforsterkende rekruttering som kan gi vippepunkter (avsnittene 8.3 og 8.4). *For det femte* formulerer vi en syklisk beslutningsmetode som forener responsprinsippet med *Theory of Constraints* og reguleringsteorien (avsnitt 7.7), og lar underordning, kapasitet og prosessuell kvalitet bestemmes av regler som svarer på observasjoner med forsinkelse; faste vedtak, desentraliserte regler og en koordinert regel etter metoden sammenlignes i avsnitt 8.5. *For det sjette* skisserer vi hvordan modellen kan kalibreres og, etter validering, brukes som beslutningsstøtte i en norsk budsjett- og prioriteringssyklus, med metoden som arbeidsform (avsnitt 9).

Forholdet til tidligere arbeid. Byggesteinene er ikke nye hver for seg. Straffesakskjeden ble modellert som ett system allerede av Blumstein og Larson (1969). At et overbelastet håndhevelsesapparat kan gi selvforsterkende kriminalitet og vippepunkter, er kjent som *enforcement swamping* (Kleiman, 1993; Kleiman og Kilmer, 2009), og ligger nær politimetningen og rekrutteringsleddet her (avsnittene 4.3 og 8.4). Alvorlighetsvektede kriminalitetsmål finnes i bruk, blant annet Cambridge Crime Harm Index (Sherman mfl., 2016). Håndhevelse som trappes opp og ned etter etterlevelsen, er kjernen i *responsive regulation* (Ayres og Braithwaite, 1992), og regulatorer med proporsjonale, integrerende og deriverte ledd ble brukt på økonomisk stabiliseringspolitikk av Phillips (1954). Det artikkelen tilfører, er koblingen: tidsforsinkelse, måleproblem, tillit, kø og rettsstatlige skranker i samme modell, og en metode som følger av den.

Status og avgrensning. Artikkelen er et konseptnotat, skrevet for å inspirere til videre forskning og til en samtale med dem som arbeider i kjeden. Modellen er teoretisk. Ingen parametere er estimert på norske data, simuleringene illustrerer mekanismer og rangerer tiltak under gitte forutsetninger, og de samfunnsøkonomiske kostnadene og gevinstene er definert, men ikke tallfestet. Definisjon 1 gjelder lover og regler generelt; modellen dekker bare straffesakskjeden, ikke tilsyn og administrative sanksjoner. Avsnitt 10 samler de kjente begrensningene og en veiplan for det som gjenstår.

Disposisjon. Rekkefølgen følger i hovedsak de fire forutsetningene ovenfor. Avsnitt 2 definerer kriminalitetsbelastning, trygghet, tillit og måleligningen, avsnitt 3 de makroøkonomiske

driverne. Avsnitt 4 formulerer responsprinsippet som en beslutningsregel og analyserer sløyfen og stabiliteten. Avsnitt 5 modellerer strategisk tilpasning og forskyvning, avsnitt 6 tillitsdynamikken, målfunksjonen og rettsstatens skranker. Avsnitt 7 utvider modellen med lovverkets handlingsrom og straffesakskjedens flaskehalser og munner ut i beslutningsmetoden. Avsnitt 8 gir den numeriske illustrasjonen, avsnitt 9 viser bruken i praksis, avsnitt 10 drøfter begrensninger, og avsnitt 11 konkluderer. Tillegg A gir en symboloversikt og tillegg B bevisene.

2 Kriminalitetsbelastning, opplevd trygghet og tillit

2.1 Objektiv kriminalitetsbelastning

Et rasjonelt håndhevelsesregime kan ikke styre etter det uvektede antallet anmeldelser alene: ett drap og ett sykkeltyveri teller da likt. Vi definerer den *objektive kriminalitetsbelastningen* $K_{\text{obj}}(t)$ ved tidspunkt t over n lovbruddskategorier $i \in \{1, \ldots, n\}$ som

$$K_{\text{obj}}(t) \;=\; \sum_{i=1}^{n} N_i(t)\, A_i\, S_i^{\theta}, \tag{2}$$

der

- $N_i(t) \geq 0$ er *aktivitetsnivået*: omfanget av den faktiske (ikke bare registrerte) kriminelle virksomheten i kategori i, målt slik at antallet lovbrudd som begås per tidsenhet, er proporsjonalt med N_i;
- $A_i > 0$ er *alvorlighetsgraden*: en normativ, juridisk vekt utledet av strafferammen og krenkelsens art;
- $S_i > 0$ er *skadeomfanget*: den gjennomsnittlige økonomiske, fysiske og sosiale skadekostnaden per hendelse, og eksponenten $\theta \geq 1$ uttrykker samfunnets aversjon mot svært skadelige hendelsestyper.

Merk at θ virker på *gjennomsnittlig* skade per kategori. Skal man fange opp aversjon mot sjeldne, katastrofale enkelthendelser, må fordelingshalen modelleres eksplisitt, for eksempel ved å skille ut slike hendelser som egne kategorier.

To presiseringer hører med. For det første er N_i i dynamikken (avsnitt 4.3) en *bestand* av pågående virksomhet, som fylles på av tilstrømningen og tappes av håndhevelse og naturlig avgang. Lovbrudd som er begått, kan ikke fjernes; det håndhevelsen fjerner, er virksomheten som ville gitt nye lovbrudd. Statistikken måler derimot *strømmer* av anmeldte og oppklarte lovbrudd. Overgangen mellom de to er ikke modellert, og effektparameterne kan derfor ikke leses direkte ut av oppklaringsstatistikk (avsnitt 10). For det andre er K_{obj} en prioriteringsindeks i slekt med etablerte alvorlighetsvektede mål som Cambridge Crime Harm Index (Sherman mfl., 2016). Bare med $A_i = 1$ og $\theta = 1$ er den målt i kroner. Med normative vekter og $\theta > 1$ er den enhetsløs, og samfunnskostnadene i avsnitt 6 må da verdsettes for seg.

2.2 Opplevd trygghet og institusjonell tillit

Kriminalpolitikk styres ikke bare av objektive tall, men også av befolkningens opplevelse. Gjennom regelmessige undersøkelser — i Norge blant annet Politiets innbyggerundersøkelse (Politidirektoratet, 2024) og Nasjonal trygghetsundersøkelse (Løvgren mfl., 2023) — kan to tilstandsvariabler, normalisert til intervallet $[0, 1]$, estimeres:

- $U(t) \in [0, 1]$, *opplevd trygghet*: andelen av befolkningen som føler seg trygg i nærmiljøet og i samfunnet;
- $\Lambda(t) \in [0, 1]$, *institusjonell tillit*: befolkningens tillit til politiets og rettsstatens integritet, rettferdighet og kompetanse.

2.3 Effektivt kriminalitetspress

Vi definerer det *effektive samfunnsmessige kriminalitetspresset* som en konveks kombinasjon av normalisert objektiv skade og subjektiv utrygghet:

$$\tilde{K}(t) \;=\; \omega\,\frac{K_{\text{obj}}(t)}{K_{\text{ref}}} \;+\; (1-\omega)\,\frac{\Phi\big(1-U(t),\,1-\Lambda(t)\big)}{\Phi_{\text{ref}}}, \qquad \omega \in [0,1], \tag{3}$$

der $\Phi : [0,1]^2 \to [0,1]$ er en økende aggregator for utrygghet og tillitssvikt, for eksempel $\Phi(x,y) = \alpha x + (1-\alpha)y$. Normaliseringen med referanseverdiene K_{ref} og Φ_{ref} (for eksempel historiske gjennomsnitt) er nødvendig fordi K_{obj} og Φ har ulike enheter; uten den ville vekten ω ikke ha noen meningsfull tolkning. Med normaliseringen er $\tilde{K} = 1$ i referansesituasjonen. Parameteren ω angir hvor stor vekt myndighetene legger på harde kriminalitetsdata i forhold til befolkningens utrygghet og tillitssvikt.

Merknad 1 (Tillit som mål eller som skranke). Dersom Λ inngår i $\tilde{K}$ og styringsregelen svarer på et økt $\tilde{K}$ med *mer* håndhevelse, oppstår en mulig selvforsterkende sløyfe: uforholdsmessig håndhevelse svekker tilliten (avsnitt 6.1), som øker $\tilde{K}$, som utløser enda mer håndhevelse. Tillitssvikt bør derfor først og fremst møtes med tiltak rettet mot *kvaliteten* på håndhevelsen (prosessuell rettferdighet), ikke mot volumet. I den operative styringen anbefaler vi å bruke Λ som en skranke (avsnitt 6.3) heller enn som en del av målvariabelen, det vil si å velge Φ uavhengig av Λ ($\alpha = 1$).

2.4 Måleligningen: det vi ser, er ikke det som skjer

Myndighetene observerer ikke N_i direkte, men registrerte lovbrudd. For *offerbaserte* lovbrudd (vinning, vold) avhenger registreringen av befolkningens anmeldelsestilbøyelighet, som øker med tilliten. For *kontrollavhengige* lovbrudd (narkotika, trafikk, økonomisk kriminalitet) avhenger registreringen i stor grad av politiets egen innsats. Vi skriver derfor

$$N_i^{\text{obs}}(t) \;=\; r_i\big(\Lambda(t),\,T_i(t)\big)\,N_i(t) + \varepsilon_i(t), \qquad 0 < r_i \le 1, \tag{4}$$

der r_i er registreringsgraden og ε_i er målefeil. Offerundersøkelser viser at mørketallene er betydelige (Løvgren mfl., 2023). Ligning (4) har to viktige konsekvenser:

(a) Økt tillit kan gi *økt* registrert kriminalitet selv om den faktiske kriminaliteten faller (et «tillitsparadoks»).

(b) For kontrollavhengige lovbrudd gir økt innsats mer registrert kriminalitet, som en naiv styringsregel tolker som behov for enda mer innsats. Dette er samme mekanisme som gir løpske tilbakekoblingssløyfer i prediktiv politivirksomhet (Ensign mfl., 2018; Lum og Isaac, 2016).

Styringen bør derfor bygge på et *estimat* av N_i — for eksempel fra et Kalman-filter som kombinerer registerdata og offerundersøkelser (Kalman, 1960) — og ikke på rå registerstatistikk (avsnitt 9). Dette er den første forutsetningen for prinsippet i definisjon 1: skal innsatsen stå i forhold til hvor mye lovene *faktisk* brytes, må det være de faktiske bruddene, og ikke registreringen av dem, som styres etter.

Observerbarhet. Estimatet er bare så godt som identifikasjonen tillater. Høyresiden i (4) er et produkt av to uobserverte størrelser; på logaritmisk form, $\log N_i^{\text{obs}} = \log r_i + \log N_i$, gir ett registrert tall per periode én ligning med to ukjente. Ut fra registerdata alene kan en økning i registrert kriminalitet ikke skilles fra en økning i anmeldelsestilbøyeligheten; tilstanden $(\log N_i, \log r_i)$ er ikke observerbar i Kalmans forstand. Offerundersøkelsene, som anslår r_i direkte, foreligger årlig og med etterslep, mens styringen er månedlig. Mellom målepunktene hviler identifikasjonen på antakelser som bør gjøres eksplisitt i tilstandsrommodellen: (i) r_i

varierer *langsomt* (en tilfeldig gang med liten varians) mens N_i følger dynamikken i (19); da er paret observerbart fordi dynamikkene er forskjellige, men svakt når N_i selv er svært persistent, og bare for endringer i N_i som er raskere enn driften i r_i; (ii) r_i gis en *strukturell* form $r_i(\Lambda, T_i)$ med Λ målt i innbyggerundersøkelsene og T_i kjent, slik at få parametere gjenstår; (iii) flere indikatorer med *uavhengige* registreringsmekanismer (forsikringskrav, skadedata fra helsevesenet, mørketallsspørsmål i løpende undersøkelser) observerer samme N_i; (iv) offerundersøkelsene brukes som forsinkede ankerpunkter for nivået på r_i. Konsekvensen for styringen er todelt. Usikkerheten i $\hat{N}_i$ skal inn i beslutningsgrunnlaget (avsnitt 9), og den har en pris i stabilitet: et filter som glatter registerstøy for å skille signal fra registreringsendringer, innfører en fasetreghet som virker som en ekstra forsinkelse, slik at den effektive τ_{info} i (10) øker med filterets tidskonstant. Svak observerbarhet er dermed i seg selv et argument for lavere reaktivitet etter korollar 1, og for å la raske omprioriteringer bygge på kategorier der registreringsgraden er best kjent.

3 Makroøkonomiske indikatorer som strukturelle drivere

Tilfanget av lovbrudd påvirkes av økonomiske rammebetingelser (Freeman, 1999; Chalfin og McCrary, 2017). Vi definerer en makroøkonomisk tilstandsvektor

$$\mathbf{E}(t) = \begin{bmatrix} u(t) & G(t) & \varphi(t) & Y_{\text{gap}}(t) \end{bmatrix}^{\top}, \tag{5}$$

der u er arbeidsledighetsraten (særlig blant unge i risikogrupper), G er inntektsulikhet målt ved Gini-koeffisienten, φ er inflasjon og levekostnadspress (fall i disponibel realinntekt), og Y_{gap} er produksjonsgapet (avviket fra potensielt BNP).

Det *strukturelle kriminalitetspådraget* Ω — tilstrømningen av lovbrudd som er uavhengig av politiets innsats — skrives i sin enkleste, lineære form som

$$\Omega(\mathbf{E}(t)) = \beta_0 + \beta_1 u(t-\ell_1) + \beta_2 G(t-\ell_2) + \beta_3 \varphi(t-\ell_3) + \beta_4 Y_{\text{gap}}(t-\ell_4), \tag{6}$$

med kategorispesifikke versjoner Ω_i i avsnitt 5. Etterslepene $\ell_k \geq 0$ fanger opp at makroøkonomiske endringer virker med forsinkelse. Fortegnene bindes ikke på forhånd. Empirien gir gode holdepunkter for $\beta_1 > 0$ for vinningskriminalitet (Raphael og Winter-Ebmer, 2001) og $\beta_2 > 0$ for voldskriminalitet (Kelly, 2000; Fajnzylber mfl., 2002). Virkningen av konjunkturer er derimot tvetydig: høykonjunktur svekker motivet for vinningslovbrudd, men øker samtidig mulighetene (flere verdier, mer aktivitet og alkohol), slik at β_4 kan ha ulikt fortegn i ulike kategorier (Cook og Zarkin, 1985). Fortegn og størrelse må derfor estimeres per kategori.

4 Tilbakekobling, tidsforsinkelse og pendelsvingninger

4.1 Responsprinsippet som beslutningsregel

Sammenhengen (1), $T \leftarrow f(K, \mathcal{O})$, sier at tiltakene er et bevisst svar på kriminaliteten slik den observeres. To presiseringer gjør den til en modell. For det første er politiets innsats bare én komponent i en *tiltaksvektor* som også omfatter reaksjonsformer, kapasitet, prosessuell kvalitet og lovbestemt handlingsrom,

$$\boldsymbol{v}(t) = (T_1, \ldots, T_n;\ \xi_k;\ Z_k;\ q;\ T_i^{\text{lov}})(t), \tag{7}$$

der ξ_k er andelen saker som føres videre til ledd k i straffesakskjeden, Z_k kapasiteten i leddet og q den prosessuelle kvaliteten (avsnittene 6.1 og 7.2). For det andre kan ingen av komponentene bygge på kriminaliteten K selv, som ikke observeres, men bare på *observasjonene*

$$\mathcal{O}(t) = \{N_i^{\text{obs}}(s),\ \text{undersøkelser},\ J_k(s),\ Q_k(s),\ \tau_k^{\text{kø}}(s),\ v_k(s),\ \mathbf{E}(s)\ :\ s \leq t - \tau_{\text{info}}\}, \tag{8}$$

det vil si registrerte lovbrudd, offer- og innbyggerundersøkelser, kjedestatistikk og makroindikatorer, som foreligger med hver sin forsinkelse, og på tilstandsestimatet som dannes fra dem gjennom måleligningen (4). Responsprinsippet i definisjon 1 er da en *beslutningsregel*

$$\boldsymbol{v}(t) = \Psi\big(\hat{\boldsymbol{x}}(t),\, \mathcal{O}(t)\big), \qquad \hat{\boldsymbol{x}}(t) = \mathbb{E}\big[\boldsymbol{x}(t) \,\big|\, \mathcal{O}(t)\big], \quad \boldsymbol{x} = (N_i, U, \Lambda, Q_k, \dots), \tag{9}$$

som avbilder det observerte bildet på et samlet sett av tiltak. Dette er den formelle versjonen av kortformen (1): T blir tiltaksvektoren $\boldsymbol{v}$, K inngår gjennom estimatet $\hat{\boldsymbol{x}}$, og f blir regelen Ψ; den enkle sammenhengen $T = f(K)$ er grensetilfellet med fullt observert K uten forsinkelse og med politiets innsats som eneste tiltak. Tre egenskaper ved (9) bærer resten av artikkelen. (i) Regelen ser aldri K; avviket $K - \hat{K}$ er en forstyrrelse i sløyfen, og observerbarheten (avsnitt 2.4) avgjør hvor stor den er og hvor mye glatting, og dermed forsinkelse, estimatet krever. (ii) Komponentene i $\boldsymbol{v}$ har ulik beslutningstakt og ledetid, fra uker for innsatsen til flere år for det lovbestemte handlingsrommet; i praksis er Ψ derfor satt sammen av regler i flere lag, hver med sin eier og tidsskala, slik hierarkisk regulering legger raske sløyfer innenfor rammer satt av langsomme (Scattolini, 2009) (avsnitt 7.7). (iii) Styringsregelen (11) nedenfor og den modellprediktive regelen (24) er spesialtilfeller med $\boldsymbol{v} = (T)$: den første avbilder det forsinkede, estimerte avviket lineært på innsatsen, den andre velger innsatsen som første trinn i en optimal plan gitt $\hat{\boldsymbol{x}}$. Tiltak som settes inn på forhåndsbestemte tidspunkter uavhengig av $\mathcal{O}$, er *åpen sløyfe*; avsnitt 8.5 viser hva som skjer når også de gjøres til responser.

4.2 Styringsregelen

For å regulere kriminalitetspresset $\tilde{K}$ mot et samfunnsmessig akseptabelt målnivå K^* setter staten inn håndhevelsesinnsatsen $T(t) = \sum_i T_i(t)$. Dette er prinsippet i definisjon 1 i operativ form: innsatsen følger avviket mellom det målte, alvorlighetsvektede kriminalitetspresset og det akseptable nivået. Mellom et lovbrudd og det operative svaret ligger en samlet systemisk tidsforsinkelse

$$\tau = \tau_{\text{info}} + \tau_{\text{pol}} + \tau_{\text{oper}}, \tag{10}$$

sammensatt av tiden til statistikken foreligger (τ_{info}), til politiske vedtak er fattet (τ_{pol}) og til tiltakene er iverksatt (τ_{oper}); figur 1 viser hvor i sløyfen de oppstår. Vi modellerer myndighetenes atferd som en tidsforsinket PID-regulator (Minorsky, 1922; Åström og Murray, 2021) basert på det observerte styringsavviket $e(t-\tau) = \hat{\tilde{K}}(t-\tau) - K^*$, der $\hat{\tilde{K}}$ er estimatet av kriminalitetspresset dannet fra observasjonene (i simuleringene antas estimatet feilfritt, slik at forsinkelsen er den eneste måleimperfeksjonen):

$$T(t) = \operatorname{sat}_{[0,\,T^{\max}]}\Big[\, T_0 + \underbrace{k_p\, e(t-\tau)}_{\text{reaktivt ledd}} + \underbrace{k_i \int_{t-W}^{t} e(t'-\tau)\,\mathrm{d}t'}_{\text{strukturelt ledd}} + \underbrace{k_d\, \dot{e}(t-\tau)}_{\text{trendledd}} \Big], \tag{11}$$

der $\operatorname{sat}_{[0,T^{\max}]}$ avgrenser innsatsen til det fysisk og rettslig mulige (innsatsen kan ikke være negativ, og kapasiteten er begrenset), og W er «budsjetthukommelsen» i det integrerende leddet. Det deriverte leddet reagerer på *trenden* i avviket og er foregripende i reguleringsteknisk forstand. Det må ikke forveksles med kriminalitets*forebygging*, som i denne modellen virker gjennom Ω.

4.3 Kriminalitetsdynamikken

Kobler vi styringsregelen til kriminalitetsutviklingen, får vi en tidsforsinket differensialligning (*delay differential equation*, DDE). På aggregert nivå ($n = 1$) skriver vi

$$\frac{\mathrm{d}K_{\text{obj}}(t)}{\mathrm{d}t} \;=\; \Omega(\mathbf{E}(t)) \;-\; \gamma\, \Lambda(t)^{\eta}\, T(t)\, K_{\text{obj}}(t) \;-\; \delta\, K_{\text{obj}}(t), \tag{12}$$

der $\gamma > 0$ er effektiviteten per enhet innsats, $\eta \geq 0$ er tillitselastisiteten (tillit multipliserer effektiviteten fordi den gir flere anmeldelser, vitneutsagn og tips), og $\delta \geq 0$ er en «naturlig» avgangsrate (aldring ut av kriminalitet, endrede muligheter). Uten δ kunne kriminaliteten bare reduseres gjennom håndhevelse, noe som er empirisk urimelig og gjør modellen ustabil når $T = 0$. Med $\delta = 0$ får man formuleringen uten naturlig avgang som spesialtilfelle.

Politiets kapasitet. Fjerningsleddet i (12) er et masseaksjonsledd: ved fast innsats T fjerner politiet dobbelt så mye når kriminaliteten dobles. Det forutsetter at politiet aldri er mettet, en forutsetning som står i et spenningsforhold til utviklingen i avsnitt 1, der henleggelser av kapasitetsgrunner øker nettopp når sakstilfanget øker. Vi lar derfor fjerningen mettes i kriminalitetsnivået etter samme form som Hollings type II-respons for predatorer med begrenset håndteringstid (Holling, 1959):

$$F(K, T, \Lambda) \;=\; \frac{\gamma' \, \Lambda^{\eta} \, T \, K}{1 + K/K_{\mathrm{sat}}} \;=\; Z_{\mathrm{E}}(T, \Lambda) \, \frac{K}{K + K_{\mathrm{sat}}}, \qquad Z_{\mathrm{E}} \coloneqq \gamma' \Lambda^{\eta} T K_{\mathrm{sat}}, \tag{12$'$}$$

der K_{sat} er *halvmetningsnivået*: ved $K = K_{\mathrm{sat}}$ arbeider politiet med halvparten av den kapasiteten Z_{E} som innsatsen gir. Den andre skrivemåten viser at (12′) har samme form som kapasitetsfunksjonen for domstolsleddet i avsnitt 7.3 (ligning (28)): etterforskningsleddet har en kapasitet proporsjonal med innsatsen, og fjerningen er den delen av kapasiteten som belastningen utløser. Med $\gamma' = \gamma(1 + K_{\mathrm{ref}}/K_{\mathrm{sat}})$ er fjerningen i referanselikevekten uavhengig av K_{sat}, og grensetilfellet $K_{\mathrm{sat}} \to \infty$ gir eksakt (12). Hovedanalysen i avsnitt 4.4 og avsnitt 8 bruker dette grensetilfellet; merknad 2 og avsnitt 8.3 viser hvordan endelig K_{sat} endrer stabiliteten og resultatene, og at metningen gir den fallende oppklaringsintensiteten som kjennetegner et presset politi: ved fast innsats er fjerningen per lovbrudd F/K proporsjonal med $1/(1 + K/K_{\mathrm{sat}})$, som avtar når belastningen øker.

4.4 Stabilitetsanalyse

For å forstå når pendelsvingninger oppstår, ser vi på et rent proporsjonalt styringsregime ($k_i = k_d = 0$) med $\omega = 1$ og $K_{\mathrm{ref}} = 1$, og holder tillit og pådrag konstante på $\bar{\Lambda}$ og $\bar{\Omega}$ (de endrer seg langsomt sammenlignet med K_{obj}). La $c \coloneqq \gamma \bar{\Lambda}^{\eta}$ være den tillitsjusterte effektiviteten, og la $(\bar{K}, \bar{T})$ være likevekten der $\bar{\Omega} = (c\bar{T} + \delta)\bar{K}$ og $\bar{T} = T_0 + k_p(\bar{K} - K^*)$, med $0 < \bar{T} < T^{\max}$. Definer

$$a \coloneqq c\,\bar{T} + \delta \text{ (samfunnets egen dempning)}, \qquad b \coloneqq c\,\bar{K}\,k_p \text{ (effektiv politisk reaktivitet)}. \tag{13}$$

Proposisjon 1 (Stabilitet under forsinkelse)**.** *Linearisert rundt $(\bar{K}, \bar{T})$ følger avviket $x = K_{\mathrm{obj}} - \bar{K}$ ligningen $\dot{x}(t) = -a\,x(t) - b\,x(t - \tau)$. Da gjelder:*

(a) Hvis $b \leq a$, er likevekten asymptotisk stabil for alle *forsinkelser $\tau \geq 0$.*

(b) Hvis $b > a$, er likevekten asymptotisk stabil hvis og bare hvis $\tau < \tau_c$, der

$$\tau_c \;=\; \frac{\arccos(-a/b)}{\sqrt{b^2 - a^2}}. \tag{14}$$

Ved $\tau = \tau_c$ krysser et par komplekskonjugerte egenverdier den imaginære aksen med frekvens $\varpi = \sqrt{b^2 - a^2}$ og periode $2\pi/\varpi$. I den ikke-lineære modellen svarer dette generisk til en Hopf-bifurkasjon: for $\tau > \tau_c$ vokser avvikene til en vedvarende svingning (grensesyklus), der amplituden begrenses av modellens ikke-lineariteter (det bilineære leddet TK_{obj} i (12), tillitsdynamikken og i siste instans metningen i (11)).

Beviset, som bygger på Hayes (1950), står i tillegg B. Proposisjonen presiserer en intuisjon som ofte uttrykkes løst: det er ikke produktet $k_p\tau$ alene som avgjør stabiliteten, men forholdet

mellom den effektive reaktiviteten b (som også avhenger av tillit og kriminalitetsnivå) og samfunnets egen dempning a. For prinsippet i definisjon 1 betyr dette at innsatsen kan følge bruddene uten å skape svingninger bare dersom reaksjonen er avdempet nok i forhold til forsinkelsen i målingen av dem. Følgende tilstrekkelige betingelse er enkel nok til bruk i praksis:

Korollar 1 (Tommelfingerregel)**.** *Siden* $\arccos(-a/b) \geq \pi/2$ *og* $\sqrt{b^2 - a^2} \leq b$*, er* $\tau_c \geq \pi/(2b)$*. Sløyfen er derfor stabil dersom*

$$k_p \, \tau \;<\; \frac{\pi}{2\,\gamma\,\bar{\Lambda}^{\eta}\,\bar{K}} \;=\; \frac{\pi}{2\,c\,\bar{K}}. \tag{15}$$

Tre politikkimplikasjoner følger direkte:

(i) *Moderat reaktivitet er robust.* Med $k_p \leq a/(c\bar{K})$ er sløyfen stabil uansett hvor lang forsinkelsen er.

(ii) *Kortere forsinkelse gir rom for kraftigere respons.* Når $b \gg a$, er $\tau_c \approx \pi/(2b)$, så en halvering av τ — for eksempel med raskere statistikk eller hyppigere omprioritering — omtrent dobler den største stabile reaktiviteten.

(iii) *Tillit gjør systemet mer følsomt.* Høyere $\bar{\Lambda}$ øker effektiviteten, men øker — alt annet likt — også c og dermed både b og svingningsfrekvensen, slik at τ_c blir kortere for en gitt k_p. Effektive institusjoner bør derfor reagere *mer* avdempet på kortsiktige utslag, ikke mindre.

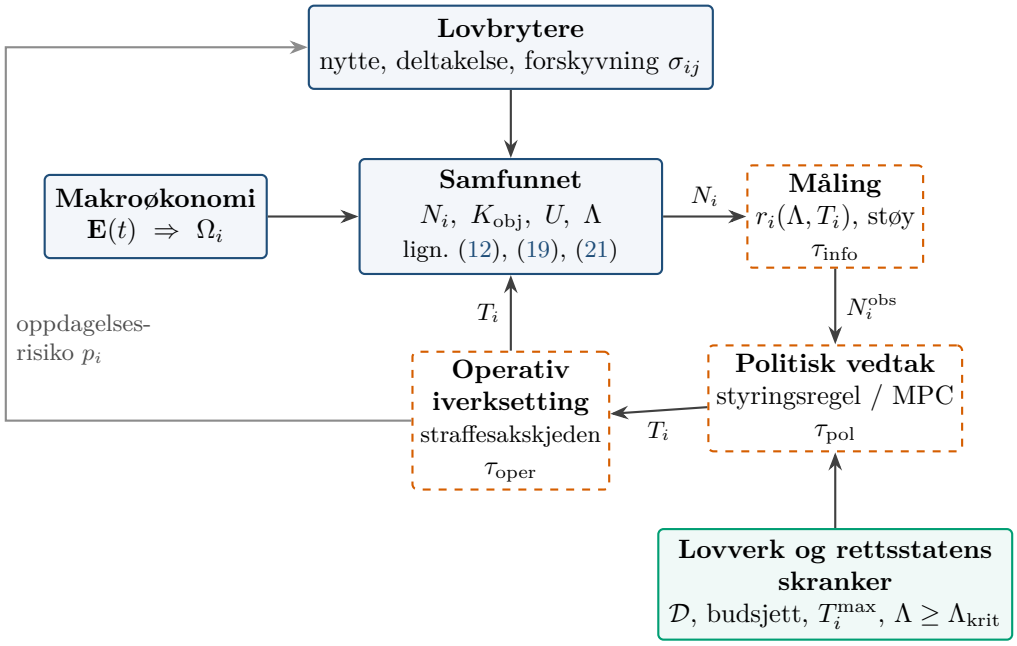
Merknad 2 (Politimetning og stabilitet)*.* Med fjerningsleddet (12′) i stedet for masseaksjonsleddet gjelder proposisjon 1 med

$$a_{\mathrm{sat}} = \bar{r}\, s(\bar{K}) + \delta, \qquad b_{\mathrm{sat}} = \bar{r}\,\bar{K}\,k_p, \qquad s(K) \coloneqq \frac{1}{1 + K/K_{\mathrm{sat}}} \in (0,1], \tag{13′}$$

der $\bar{r} \coloneqq F(\bar{K}, \bar{T}, \bar{\Lambda})/\bar{K}$ er fjerningen per lovbrudd i likevekten (tillegg B). Med γ' kalibrert som i (12′) er $\bar{r} = c\bar{T}$ uavhengig av K_{sat}, så den effektive reaktiviteten b er *uendret*, mens dempningen a reduseres med faktoren $s(\bar{K})$: når belastningen øker, øker fjerningen mindre enn proporsjonalt, og samfunnets egen dempning av et avvik blir svakere. Siden τ_c i (14) er voksende i a ved fast b, *reduserer* metningen den kritiske forsinkelsen og grensen $k_p \leq a/(c\bar{K})$ for stabilitet uansett forsinkelse. Med parameterne i tabell 2 faller τ_c fra 7,2 til 6,7 måneder når $K_{\mathrm{sat}} = 3$ og til 6,4 måneder når $K_{\mathrm{sat}} = 1{,}5$, mens grensen for stabilitet uansett forsinkelse faller fra $k_p = 1{,}2$ til 0,95 og 0,8. Et presset politi gjør altså styringssløyfen mer følsom, ikke mindre, og tommelfingerregelen (15) gjelder uendret fordi den ikke bruker a.

Det integrerende leddet fjerner varige avvik, men tilfører fasetap og senker stabilitetsgrensen ytterligere. Det deriverte leddet kan i prinsippet kompensere for noe av forsinkelsen, men forsterker målestøy, som er betydelig i kriminalstatistikk. En prinsipielt bedre løsning er å *predikere* virkningen av allerede vedtatte tiltak, slik en Smith-prediktor (Smith, 1957) eller en modellprediktiv regulator (avsnitt 6.3) gjør.

Forbeholdet om det integrerende leddet har praktisk betydning. Offentlige budsjetter er i hovedsak inkrementelle: neste års ramme er fjorårets pluss en justering. Det er integrerende styring, ikke proporsjonal styring rundt en fast grunninnsats T_0. For en slik prosess gjelder verken tommelfingerregelen (15) eller resultatet om at moderat reaktivitet er stabil uansett forsinkelse uten videre; proposisjon 1 gir da en øvre grense for hva sløyfen tåler, ikke en garanti. Stabilitetsgrensen for en inkrementell regel er ikke utledet her og hører til det som gjenstår (avsnitt 10).



Figur 1. Styringssløyfen. Stiplede bokser markerer leddene som bidrar til den samlede forsinkelsen $\tau = \tau_{\text{info}} + \tau_{\text{pol}} + \tau_{\text{oper}}$. Makroøkonomien virker som en ytre forstyrrelse, lovbryternes tilpasning som en strategisk motkraft og lovverket og rettsstatens skranker som grenser for beslutningsrommet. Den operative iverksettingen omfatter hele straffesakskjeden, som analyseres nærmere i avsnitt 7.

5 Strategisk tilpasning: lovbrytere og politi

5.1 Nytte, oppdagelsesrisiko og deltakelse

Vi modellerer samspillet mellom rettshåndhevere (spiller 1) og potensielle lovbrytere (spiller 2) som et dynamisk spill der myndighetene fordeler den samlede håndhevelsesinnsatsen $T(t) = \sum_{i=1}^{n} T_i(t)$ på kategoriene. I tråd med Becker (1968) er den forventede nytten for en potensiell lovbryter av å begå et lovbrudd i kategori i

$$\mathbb{E}[\mathcal{U}_i^C(t)] \;=\; (1 - p_i(t))\, B_i(\mathbf{E}(t)) \;-\; p_i(t)\, F_i(A_i), \tag{16}$$

der B_i er gevinsten (som avhenger av de økonomiske rammebetingelsene), F_i er den forventede sanksjonen (som øker med alvorlighetsgraden A_i), og p_i er oppdagelsesrisikoen. Formuleringen forutsetter at gevinsten går tapt ved oppdagelse. Oppdagelsesrisikoen avhenger både av innsatsen og av befolkningens tillit, fordi tillit gir flere anmeldelser, tips og vitner:

$$p_i(t) \;=\; 1 - \exp\Big(-\gamma_i\, T_i(t)\, \Lambda(t)^{\eta} \Big). \tag{17}$$

Den eksponentielle formen gir avtagende grenseavkastning av innsats. Størrelsen $h_i(t) := \gamma_i T_i(t) \Lambda(t)^{\eta}$ tolkes som oppdagelsesintensiteten. Empirisk er det godt belegg for at *sannsynligheten* for å bli tatt avskrekker mer enn *strengheten* i straffen (Nagin, 2013; Chalfin og McCrary, 2017).

For å knytte individuelle valg til aggregerte størrelser lar vi deltakelsesandelen følge en logistisk modell for diskrete valg:

$$\Pi_i(t) = \frac{1}{1 + \exp\big(-\mathbb{E}[\mathcal{U}_i^C(t)]/s_i \big)}, \qquad \Omega_i(t) = M_i(\mathbf{E}(t))\, \Pi_i(t), \tag{18}$$

der M_i er det makroøkonomisk bestemte potensialet av motiverte aktører og muligheter (jf. ligning (6)), og $s_i > 0$ måler heterogeniteten i befolkningen. Innsats virker dermed gjennom

to kanaler: *avskrekking*, som senker tilstrømningen Ω_i via p_i, og *inkapasitering og forstyrrelse*, som fjerner pågående aktivitet med intensitet h_i i ligning (19) nedenfor. I den aggregerte modellen (12) er avskrekkingskanalen absorbert i Ω.

Merknad 3 (Tsebelis' paradoks). I et simultant inspeksjonsspill der både lovbrytere og politi optimerer, fører en strengere straff i likevekt til at politiet reduserer innsatsen, mens kriminaliteten forblir uendret (Tsebelis, 1989). Når myndighetene derimot *forplikter* seg til et innsatsnivå gjennom budsjett og planer (et Stackelberg-spill), gjelder ikke paradokset på samme måte. Modellen her forutsetter forpliktelse på budsjettnivå, men paradokset kan være relevant for kortsiktige operative valg.

5.2 Forskyvning mellom kategorier

Når politiet trapper opp innsatsen mot én type kriminalitet j, kan en del av aktiviteten flytte til andre kategorier eller områder i (Guerette og Bowers, 2009). Vi modellerer dette med en *forskyvningsmatrise* $\Sigma = [\sigma_{ij}]$:

$$\frac{\mathrm{d}N_i(t)}{\mathrm{d}t} \;=\; \Omega_i(t) \;-\; \big(h_i(t)+\delta_i\big)N_i(t) \;+\; \sum_{j\neq i}\sigma_{ij}\,h_j(t)\,N_j(t), \qquad \sum_{i\neq j}\sigma_{ij} \le 1 \;\; \forall j. \tag{19}$$

Andelen σ_{ij} av aktiviteten som forstyrres i kategori j, dukker opp igjen i kategori i. Skranken på kolonnesummene sikrer at forskyvningen aldri overstiger det volumet som faktisk fortrenges. Negative verdier av σ_{ij} representerer *spredning av gevinster* (*diffusion of benefits*): innsats mot j reduserer også i. Systematiske oversikter finner at forskyvningen i gjennomsnitt er klart mindre enn full, og at spredning av gevinster er minst like vanlig (Guerette og Bowers, 2009; Braga mfl., 2019). Med $n = 1$ og $\sigma \equiv 0$ reduseres (19) til (12). I diskret tid (for eksempel månedlig, $\Delta = 1$) brukes $N_i(t+\Delta) = N_i(t) + \Delta \cdot [\text{høyresiden i (19)}]$, eventuelt med eksponentiell diskretisering av fjerningsleddet for å sikre $N_i \ge 0$.

6 Tillit, samfunnsøkonomi og rettsstatens skranker

6.1 Trygghet og tillit

Opplevd trygghet tilpasser seg gradvis det faktiske kriminalitetsnivået:

$$\frac{\mathrm{d}U(t)}{\mathrm{d}t} \;=\; \kappa_U\Big(U^{\mathrm{eq}}\big(K_{\mathrm{obj}}(t)\big) - U(t)\Big), \qquad U^{\mathrm{eq}}(K) = U_0 - \psi\,\big(K/K_{\mathrm{ref}} - 1\big), \tag{20}$$

projisert til $[0,1]$, der κ_U er tilpasningshastigheten og ψ er følsomheten. Medieoppslag og enkelthendelser kan legges til som eksogene sjokk.

For lite håndhevelse svekker tilliten fordi staten ikke klarer å beskytte borgerne (U faller). Uforholdsmessig mye håndhevelse svekker tilliten fordi staten selv oppleves som overgriper (Tyler, 1990; Sunshine og Tyler, 2003). Vi modellerer dette som

$$\begin{aligned}\frac{\mathrm{d}\Lambda(t)}{\mathrm{d}t} \;=\; & \kappa_\Lambda\Big(\Lambda_0 + \mu_1\big(U(t)-U_0\big) + \mu_3\big(q(t)-q_0\big) - \Lambda(t)\Big) \\ & - \mu_2\,\Lambda(t)\,\max\Big(0,\; T(t) - T_{\mathrm{prop}}\big(K_{\mathrm{obj}}(t)\big)\Big)^2,\end{aligned} \tag{21}$$

der $T_{\mathrm{prop}}(K)$ er den høyeste innsatsen som oppleves som forholdsmessig ved kriminalitetsnivået K (for eksempel $T_{\mathrm{prop}}(K) = \nu_0 K$), og $q(t) \in [0,1]$ er den *prosessuelle kvaliteten* på håndhevelsen: respektfull behandling, nøytralitet, begrunnelse og mulighet til å bli hørt. Det første leddet lar tilliten vende tilbake mot et nivå bestemt av trygghet og prosessuell kvalitet. Uten en slik tilbakevending ville tilliten drive ubegrenset ved ethvert varig avvik $U \neq U_0$. Det siste leddet straffer overhåndhevelse kvadratisk, og faktoren Λ hindrer at tilliten blir negativ. Leddet med q fanger opp kjernen i Tylers teori: hvordan politiet opptrer, betyr minst like mye for legitimiteten som hvor mye politiet gjør. Det kausale belegget for at prosessuell rettferdighet gir bedre etterlevelse og samarbeid, er omdiskutert (Nagin og Telep, 2017); tillitselastisiteten η og parameterne i (21) bør derfor behandles som usikre.

6.2 Samfunnsøkonomisk målfunksjon

Myndighetenes overordnede styringsproblem er å minimere det diskonterte samfunnstapet

$$\min_{\{T_i(\cdot),\, q(\cdot)\}} \mathcal{J} = \int_0^\infty e^{-\rho t}\Big[C_{\text{skade}}\big(K_{\text{obj}}(t)\big) + C_{\text{utrygghet}}\big(1 - U(t)\big) + C_{\text{budsjett}}\big(T(t), q(t)\big) + C_{\text{rettsstat}}\big(T(t),\, 1 - \Lambda(t)\big)\Big]\, \mathrm{d}t \tag{22}$$

under dynamikken i ligningene (19)–(21), måleligningen (4) og skrankene i avsnitt 6.3; $\rho > 0$ er diskonteringsraten. Kostnadsfunksjonene antas økende og konvekse. Med $A_i = 1$ og $\theta = 1$ er K_{obj} målt i skadekostnad, og C_{skade} kan være identitetsfunksjonen; ellers må indeksen oversettes til kroner. $C_{\text{rettsstat}}$ representerer kostnader ved inngrep i borgernes rettigheter og ved svekket legitimitet ut over virkningen på kriminaliteten. Rettsøkonomisk tilsvarer (22) kriteriet om å minimere summen av skadekostnader og håndhevelseskostnader (Becker, 1968; Polinsky og Shavell, 2000), utvidet med trygghet og legitimitet.

Fire forbehold er nødvendige. For det første er kostnadsfunksjonene ikke tallfestet i denne artikkelen. Simuleringene i avsnitt 8 styrer mot målnivåer og sammenligner ikke regimene på J, og skyggeprisene som omtales senere, er ikke beregnet. For det andre må en verdsetting unngå dobbelttelling: utrygghet kan ikke ligge både i den sosiale skadekostnaden S_i, i $C_{\text{utrygghet}}$ og i $\tilde{K}$, og tillit kan ikke uten videre være både kostnadsledd og hard skranke. For det tredje dekker C_{budsjett} her politiets innsats. En fullstendig analyse må ta med kostnadene i påtale, domstoler, forsvar og bistand, kriminalomsorg og konfliktråd, sanksjonenes egne samfunnskostnader og virkninger på tilbakefall og arbeid (Polinsky og Shavell, 2000; Bhuller mfl., 2020), skattefinansieringskostnaden og fordelingsvirkningene, i tråd med rammeverket for samfunnsøkonomiske analyser i staten. For det fjerde er (22) og definisjon 1 ikke samme kriterium. Å minimere J betyr å sette innsatsen inn der den siste kronen reduserer mest skade, og det avhenger av hvor godt håndhevelsen virker (γ_i) og hva den koster (w_i), ikke bare av omfang og alvorlighet. Der håndhevelse virker svakt, kan det riktige svaret på økte brudd være forebygging snarere enn mer innsats. Definisjon 1 angir utgangspunktet og retningen; (22) er kriteriet tiltakene til slutt må vurderes etter.

6.3 Rettsstaten som skrankeoppfyllelses- og optimeringsproblem

Fra et informatikk- og operasjonsanalyseperspektiv kan styringsproblemet formuleres som et *Constraint Satisfaction Problem* (CSP) $\langle \mathcal{X}, \mathcal{D}, \mathcal{C} \rangle$ (Rossi mfl., 2006), der $\mathcal{X} = \{T_1(t), \ldots, T_n(t)\}$ er beslutningsvariablene, $\mathcal{D}$ er domenet av lovlige virkemidler, og $\mathcal{C} = \mathcal{C}_{\text{hard}} \cup \mathcal{C}_{\text{soft}}$ er skrankene:

$$\begin{aligned}
\mathcal{C}_{\text{hard}}: \quad & \textstyle\sum_{i=1}^{n} w_i\, T_i(t) \le \mathcal{B}(t) && \text{(budsjett og kapasitet i politi og domstoler)} \\
& T_i^{\min} \le T_i(t) \le T_i^{\max} && \text{(rettsstatlige skranker: gulv og tak)} \\
& \Lambda(t) \ge \Lambda_{\text{krit}} && \text{(minsteterskel for institusjonell legitimitet)} \\
\mathcal{C}_{\text{soft}}: \quad & K_{\text{obj}}(t) \le K^* && \text{(politisk målnivå for kriminalitet)} \\
& U(t) \ge U^* && \text{(målnivå for opplevd trygghet)}
\end{aligned} \tag{23}$$

Her er w_i enhetskostnaden for innsats i kategori i og $\mathcal{B}(t)$ budsjettrammen. Grunnloven og menneskerettighetene setter både tak og gulv. Taket $T_i^{\max}$ begrenser hvor inngripende håndhevelsen kan være. Gulvet $T_i^{\min} \ge 0$ følger av statens plikt til å sikre rettighetene (Grunnloven § 92 og de positive forpliktelsene etter EMK): alvorlige krenkelser skal etterforskes og forfølges effektivt, så i enkelte kategorier kan innsatsen ikke prioriteres ned uansett hvordan resten av bildet ser ut. Begge er stiliserte størrelser: rettighetene begrenser først og fremst *metoder*, ikke ressursmengde. (I styringsregelen (11) betegner $T^{\max}$ det samlede kapasitetstaket; her er $T_i^{\max}$ den rettslige grensen per kategori.) De to siste harde skrankene er av ulik type: $T_i \le T_i^{\max}$ er en skranke på beslutningsvariablene, mens $\Lambda \ge \Lambda_{\text{krit}}$ er en *tilstandsskranke*. Den kan bare håndheves gjennom modellens prediksjoner av hvordan tilliten vil utvikle seg.

Overbestemte situasjoner. Dersom pådraget Ω øker kraftig mens budsjettet $\mathcal{B}$ og rettsstatsskrankene $T_i^{\max}$ ligger fast, kan det hende at ingen tillatt tiltaksplan oppfyller alle skrankene: problemet er *overbestemt* (*over-constrained*). Man må da gå over til et skrankebundet optimeringsproblem (COP), der de myke skrankene erstattes av straffeledd med slakkvariabler, $\zeta_K \left[K_{\text{obj}} - K^*\right]_+^2 + \zeta_U \left[U^* - U\right]_+^2$. Beslutningstakerne står da overfor et reelt valg: enten akseptere brudd på de myke skrankene (høyere kriminalitet eller lavere trygghet) eller endre en hard skranke. Her er det en vesentlig normativ asymmetri. Budsjettet $\mathcal{B}$ er en legitim politisk størrelse som kan endres. De konstitusjonelle skrankene $T_i^{\max}$ og menneskerettighetene er derimot ikke styringsvariabler i en rettsstat. En forsvarlig bruk av modellen behandler dem som absolutte og rapporterer heller hvilken *pris* (skyggepris) de har, målt i kriminalitet og utrygghet. Skrankesettet avgrenser dermed det rommet prinsippet i definisjon 1 gjelder innenfor: innsatsen tilpasses bruddene, mens de konstitusjonelle grensene og legitimitetsterskelen ligger fast uansett hvordan kriminalitetsbildet utvikler seg. Mellom disse ytterpunktene ligger de grensene som er fastsatt i ordinær lov; i avsnitt 7.1 deles $T_i^{\max}$ derfor i en konstitusjonell og en lovbestemt del, og kapasiteten i domstolene behandles som en egen skranke.

Modellprediktiv styring. I dynamisk form løses problemet som *modellprediktiv styring* (MPC) (Rawlings mfl., 2017). Ved hvert beslutningstidspunkt t estimeres dagens tilstand $\hat{\boldsymbol{x}}(t) = (\hat{N}_1, \ldots, \hat{N}_n, \hat{U}, \hat{\Lambda})$, medregnet virkningen av tiltak som er vedtatt, men ennå ikke har virket. Deretter løses

$$\begin{aligned} \min_{T_{0|t},\ldots,T_{H-1|t}} \; & \sum_{k=0}^{H-1} g(\boldsymbol{x}_{k|t}, T_{k|t}) + V_f(\boldsymbol{x}_{H|t}) \\ \text{s.a.} \;\; & \boldsymbol{x}_{k+1|t} = f\big(\boldsymbol{x}_{k|t}, T_{k|t}, \hat{\mathbf{E}}_{t+k}\big), \quad \boldsymbol{x}_{0|t} = \hat{\boldsymbol{x}}(t), \quad (\boldsymbol{x}_{k|t}, T_{k|t}) \text{ oppfyller (23)}, \end{aligned} \tag{24}$$

der $\boldsymbol{x}_{k|t}$ er den predikerte tilstanden k perioder fram, g er det diskrete motstykket til integranden i (22), H er planleggingshorisonten og V_f er en sluttkostnad. Bare det første steget $T_{0|t}$ iverksettes. Ved neste beslutningstidspunkt gjentas prosedyren med oppdatert informasjon. Er usikkerheten vesentlig, kan problemet formuleres som en skrankebundet Markov-beslutningsprosess (CMDP) (Altman, 1999). Fordi MPC eksplisitt predikerer virkningen av allerede vedtatte tiltak, motvirker den mekanismen som gir pendelsvingninger i avsnitt 4.4.

Beregningskompleksitet. Det er fristende å hevde at ressursfordeling med mange kategorier og forskyvningseffekter er «NP-hard», og at dette i seg selv begrunner desentralisering. Med kontinuerlige tiltak, konvekse kostnader og lineære budsjettskranker er hvert MPC-steg imidlertid et glatt, lavdimensjonalt problem som løses effektivt, selv om dynamikken er bilineær; NP-hardhet oppstår først med udelelige ressurser, kompliserte planleggingsskranker eller lovbrytere av ukjente typer, som i bayesianske Stackelberg-spill (Conitzer og Sandholm, 2006; Korzhyk mfl., 2010; Tambe, 2011), og også da finnes gode tilnærmingsmetoder. Det sterkeste argumentet for lokal tilpasning er derfor ikke beregningsmessig, men *informasjonsmessig*: kunnskap om lokale forhold er spredt og vanskelig å sentralisere (Hayek, 1945).

7 Lovverk og domstoler: handlingsrom og flaskehalser i straffesakskjeden

Til nå har lovverket og domstolene inngått i modellen bare som en fast ramme: lovverket gjennom domenet $\mathcal{D}$ og grensene $T_i^{\max}$, domstolene gjennom en udifferensiert kapasitetsskranke i (23). Forenklingen skjuler to forhold som avgjør om prinsippet i definisjon 1 lar seg realisere, og som svarer til den tredje og fjerde forutsetningen i avsnitt 1. For det første bestemmer lovverket hvilke virkemidler som er tilgjengelige, og hvor raskt de kan tas i bruk. Et styringssystem som kan omprioritere innsatsen månedlig, men bare innenfor et regelverk som endres over flere år, er dynamisk bare innenfor de grensene regelverket setter. For det

andre er politiets innsats det første leddet i en kjede — etterforskning, påtale, domstol og straffegjennomføring — der den samlede gjennomstrømningen bestemmes av det svakeste leddet. Når domstolene er dette leddet, omsettes en opptrapping av politiinnsatsen i restanser, ikke i reaksjoner. Dette avsnittet utvider modellen på begge punkter. Hovedbudskapet kan sammenfattes i to teser: *(i) lovverket må gi et lovbestemt handlingsrom som gjør dynamisk rettshåndhevelse mulig innenfor rettsstatens faste grenser; (ii) hvert ledd i kjeden må dimensjoneres og utformes slik at det ikke blir en flaskehals som nøytraliserer styringen.*

Kjeden analyseres med domstolen som gjennomregnet eksempel. Valget er pedagogisk, ikke empirisk: tallene i avsnitt 1 peker i dag mot politi og påtale og mot straffegjennomføringen snarere enn mot tingrettene. Mekanismene — kø, frafall, forsinket virkning og et signal oppstrøms som flaskehalsen selv blåser opp — gjelder for ethvert ledd som binder. Tiltakene har derimot ulik form fra ledd til ledd: ved en flaskehals i etterforskning og påtale er underordning prioritering mellom saker, og ved en flaskehals i straffegjennomføringen alternative gjennomføringsformer. En kjede med kø og frafall i alle fire ledd, og en diagnose på faktiske tall, hører til det som gjenstår (avsnitt 10).

7.1 Lovgivningen som den langsomste sløyfen

Styringen av rettshåndhevelsen foregår på minst tre tidsskalaer. Operativ omprioritering skjer over uker og måneder, budsjettstyring over år, og lovgivning over flere år: utredning, høring, proposisjon, stortingsbehandling og ikrafttredelse gir en samlet lovgivningsforsinkelse $\tau_{\text{lov}} \gg \tau$. Lovverket bestemmer domenet $\mathcal{D}$ av lovlige virkemidler — hvilke metoder, reaksjonsformer og prosessformer som finnes — og det setter de rettslige grensene for hvor intensivt hvert virkemiddel kan brukes. Den øvre grensen $T_i^{\max}$ i (23) bør derfor deles i to:

$$T_i^{\min} \leq T_i(t) \leq T_i^{\text{lov}}(t) \leq T_i^{\text{grl}}, \tag{25}$$

der T_i^{grl} er den *konstitusjonelle* grensen, gitt av Grunnloven og de menneskerettighetene Norge er bundet av (EMK), og $T_i^{\text{lov}}(t)$ er den *lovbestemte* grensen, som lovgiver legitimt kan flytte innenfor T_i^{grl} gjennom ordinær lovgivningsprosess. Oppdelingen presiserer den normative asymmetrien i avsnitt 6.3: T_i^{grl} er ikke en styringsvariabel, mens T_i^{lov} er det, men bare for lovgiver og bare på tidsskalaen τ_{lov}.

Konsekvensen for styringsproblemet følger direkte. Når $\mathcal{D}$ og T_i^{lov} ligger fast mens kriminalitetsbildet endrer seg — digital kriminalitet er det tydeligste eksempelet, med nye handlingsmønstre, nye bevistyper og aktører utenfor norsk jurisdiksjon — blir problemet i avsnitt 6.3 *systematisk* overbestemt. Ingen omfordeling av innsatsen innenfor et uendret domene oppfyller skrankene, og modellen vil rapportere en vedvarende høy skyggepris på de lovbestemte grensene. En slik skyggepris er et signal til lovgiver, ikke til politiet.

Det er likevel avgjørende hvordan lovgivningen skal gi rom for dynamikk i håndhevelsen. Innsikten i proposisjon 1 gjelder som analogi også for den lovgivende sløyfen: en sløyfe med lang forsinkelse tåler bare lav reaktivitet. Hendelsesdrevet hastelovgivning etter enkeltsaker er lovgivningens motstykke til en for høy k_p ved lang τ. Den svarer på et avvik som ble observert lenge etter at det oppsto, virker først når situasjonen kan være en annen, og gir samme risiko for overreaksjon som i avsnitt 4.4 — med den forskjellen at straffelovgivning i tillegg har varige virkninger for forutberegnelighet og likebehandling. Dynamikken bør derfor ikke legges i lovgivningens *reaksjonshastighet*, men bygges inn *på forhånd* som et regelstyrt handlingsrom:

(i) *Rammelover med avgrenset delegasjon.* Loven fastsetter formål, vilkår og ytre grenser, mens prioritering, organisering og nærmere gjennomføring reguleres i forskrift som kan endres på budsjettsyklusens tidsskala.

(ii) *Forsøkshjemler.* Tidsavgrensede forsøk med nye prosess- eller reaksjonsformer i et utvalg distrikter eller domstoler, med forhåndsdefinert evaluering, gir samtidig den kvasieks-

perimentelle variasjonen som kalibreringen i avsnitt 9 er avhengig av (Ranchordás, 2014).

(iii) *Solnedgangsklausuler.* Inngripende hjemler gis begrenset varighet og faller bort med mindre de fornyes etter evaluering. Det gjør det mulig å møte en reell trussel uten at reaksjonen blir permanent (Ranchordás, 2014).

(iv) *Innebygd etterhåndsevaluering.* Lovvedtak følges av indikatorer og en plikt til evaluering etter en fastsatt periode, slik at lovgivningssløyfen får en måling å styre etter.

Handlingsrommet er ikke ubegrenset. Legalitetsprinsippet (Grunnloven § 113) krever at myndighetenes inngrep overfor den enkelte har grunnlag i lov, og setter grenser for hvor mye som kan delegeres til forskrift. Grunnloven § 96 (ingen kan dømmes uten etter lov eller straffes uten etter dom) og EMK artikkel 7 krever at straffbarhet og strafferamme er fastsatt i lov før handlingen, og Grunnloven § 97 forbyr lover med tilbakevirkende kraft. Grunnloven § 95 og EMK artikkel 6 gir rett til en rettferdig rettergang innen rimelig tid, og Grunnloven § 98 krever likhet for loven. Handlingsrommet bør derfor først og fremst ligge i *prioritering, organisering, prosessformer og reaksjonsformer*, ikke i hva som er straffbart eller hvor strengt det straffes. Det er den samme avgrensningen som ligger i presiseringen til definisjon 1: håndhevelsen er dynamisk, lovene er faste. Det er også der den største gevinsten ligger, som neste delavsnitt viser.

7.2 Straffesakskjeden som kønettverk

Politiets innsats T virker ikke som reaksjon i seg selv. Den setter i gang en kjede av ledd $k \in \{\mathrm{E}, \mathrm{P}, \mathrm{D}, \mathrm{S}\}$ — etterforskning, påtale, domstol (tingrett og eventuelt lagmannsrett) og straffegjennomføring — som hvert har en innstrømning av saker $J_k(t)$, en kapasitet $Z_k(t)$ (saker som kan avgjøres per måned) og en kø $Q_k(t)$ (figur 2). Når køen ikke er tom, gjelder

$$\frac{\mathrm{d}Q_k(t)}{\mathrm{d}t} = J_k(t) - Z_k(t), \qquad \upsilon_k(t) \coloneqq \frac{J_k(t)}{Z_k(t)}, \qquad \tau_k^{\text{kø}}(t) = \frac{Q_k(t)}{Z_k(t)}, \tag{26}$$

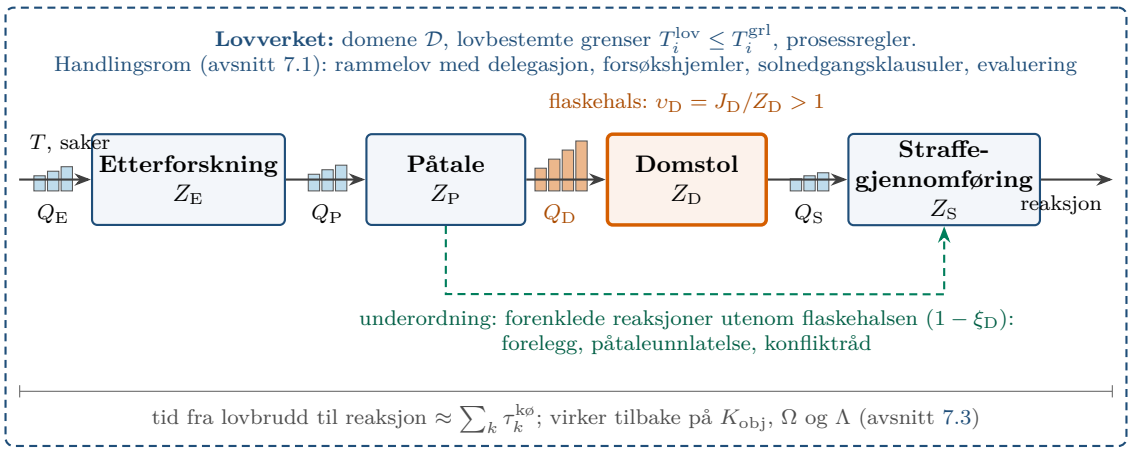
der υ_k er utnyttelsesgraden og $\tau_k^{\text{kø}}$ den forventede køtiden etter Littles lov (Little, 1961). Kjedens gjennomstrømning kan ikke overstige $\min_k Z_k$. Innstrømningen til domstolene er en andel av den aktiviteten politiet avdekker og fjerner, det vil si av fjerningsleddet i (12):

$$J_{\mathrm{D}}(t) = \xi_{\mathrm{D}}\, \gamma\, \Lambda(t)^{\eta}\, T(t)\, K_{\text{obj}}(t), \tag{27}$$

der $\xi_{\mathrm{D}} \in (0, 1]$ er andelen av de oppklarte sakene som bringes inn for retten, mens resten avgjøres med forenklede reaksjoner (forelegg, påtaleunnlatelse, overføring til konfliktråd). Strømmen måles i samme normaliserte enheter som K_{obj}. Tre innsikter følger.

Mer politiinnsats øker køen, ikke antall avgjorte saker. Når domstolen er flaskehalsen ($\upsilon_{\mathrm{D}} > 1$), er antall avgjorte saker per måned lik Z_{D} uansett hvor stor T er. En opptrapping av politiinnsatsen øker J_{D}, og restansen vokser lineært med $J_{\mathrm{D}} - Z_{\mathrm{D}}$: køtiden øker med $\upsilon_{\mathrm{D}} - 1$ måneder for hver måned overbelastningen varer. Det forlenger tiden fra lovbrudd til reaksjon. Den aggregerte modellen i avsnitt 4 lar hele virkningen av innsatsen inntre ved pågripelse; den delen av virkningen som først realiseres ved dom — inkapasitering gjennom soning og den individuelle og allmenne avskrekkingen — forsinkes i virkeligheten med $\tau_{\mathrm{D}}^{\text{kø}}$, og for saker som henlegges eller foreldes, uteblir den. Hvordan dette virker inn på styringssløyfen, er ikke opplagt, og modelleres i avsnitt 7.3.

Ventetiden vokser sterkt før kapasiteten er brukt opp. Selv under full utnyttelse er køtiden sterkt ikke-lineær. Med tilfeldig variasjon i innstrømning og behandlingstid (M/M/1-køen,



Figur 2. Straffesakskjeden som kønettverk innenfor lovverkets ramme. Hvert ledd k har kapasitet Z_k og kø Q_k; gjennomstrømningen begrenses av det leddet med lavest kapasitet, her domstolen (uthevet). Den stiplede grønne pilen viser underordning: saker med lav alvorlighet ledes utenom flaskehalsen med forenklede reaksjoner. Lovverket avgrenser både virkemidlene og prosessformene, og er derfor selv en mulig regelskranke.

Kleinrock, 1975) er forventet tid i systemet $1/(1-\upsilon)$ ganger gjennomsnittlig behandlingstid: fem ganger ved $\upsilon = 0{,}8$ og tjue ganger ved $\upsilon = 0{,}95$. En kapasitetsbuffer i flaskehalsen er derfor ikke sløsing, men en forutsetning for forutsigbar saksbehandlingstid, og for at kjeden skal kunne absorbere de svingningene i J_{D} som en ustabil sløyfe skaper (avsnitt 8.2). En domstol med mange dommere (M/M/c-køen) har mindre kø ved samme utnyttelse, og en domstol som prioriterer alvorlige saker gir disse kort ventetid selv når restansen er stor; begge forholdene tas opp i avsnitt 8.3.

Tid til reaksjon berører avskrekking, tillit og rettigheter. Teoretisk svekker lang tid fra lovbrudd til reaksjon den avskrekkende virkningen, fordi sanksjonen diskonteres. Det empiriske belegget for denne «celerity»-effekten er imidlertid klart svakere enn belegget for virkningen av oppdagelsessannsynligheten (Nagin, 2013; Chalfin og McCrary, 2017). Vi tar den derfor med som en moderat og begrenset kanal med eksplisitt følsomhetsanalyse (avsnittene 7.3 og 8.2). Det som er godt dokumentert, er at lang saksbehandlingstid rammer fornærmede og siktede, svekker bevisene og tilliten til at rettsstaten virker (q og Λ i (21)), og at den har en rettslig grense i kravet om avgjørelse innen rimelig tid (Grunnloven § 95, EMK artikkel 6).

7.3 Kjeden i den dynamiske modellen

For at kjeden skal inngå i den dynamiske modellen, trengs tre ting: hvordan domstolen avvikler køen, hvilken del av fjerningsleddet i (12) som forutsetter en rettskraftig avgjørelse, og hvordan saksbehandlingstiden virker tilbake på tilstrømning og tillit. Vi konsentrerer oss om domstolsleddet som gjennomregnet eksempel på et ledd som binder, og lar Q_{D} være beholdningen av saker som er kommet inn, men ikke avgjort:

$$\frac{\mathrm{d}Q_{\text{D}}(t)}{\mathrm{d}t} = J_{\text{D}}(t) - R_{\text{D}}(t) - \varsigma_{\text{D}}\, Q_{\text{D}}(t), \quad R_{\text{D}} = \frac{Z_{\text{D}}\, Q_{\text{D}}}{Q_{\text{D}} + m_{\text{D}} Z_{\text{D}}}, \quad \tau_{\text{D}}^{\text{kø}} = \frac{Q_{\text{D}}}{R_{\text{D}}} = m_{\text{D}} + \frac{Q_{\text{D}}}{Z_{\text{D}}}, \tag{28}$$

der R_{D} er antall avgjorte saker per måned, m_{D} er den rene behandlingstiden per sak i en ubelastet domstol, og ς_{D} er raten for saker som faller ut av køen uten reaksjon. I virkeligheten skjer slikt frafall i hovedsak som henleggelse i påtaleleddet og ved foreldelse; i modellen er alt frafall etter oppklaring lagt til denne ene køen. Kapasitetsfunksjonen er glatt: en liten kø ($Q_{\text{D}} \ll m_{\text{D}} Z_{\text{D}}$) avgjøres i takt med innstrømningen; en stor kø gir full kapasitetsutnyttelse, $R_{\text{D}} \approx Z_{\text{D}}$, og (28) går over i (26). Saksbehandlingstiden følger av

Littles lov med R_{D} som gjennomstrømning. Uten frafall har (28) en stasjonær løsning bare når $\upsilon_{\mathrm{D}} < 1$, og da er $\tau_{\mathrm{D}}^{\text{kø}} = m_{\mathrm{D}}/(1-\upsilon_{\mathrm{D}})$, som er M/M/1-sammenhengen ovenfor. Med $\upsilon_{\mathrm{D}} > 1$ vokser køen inntil frafallet $\varsigma_{\mathrm{D}} Q_{\mathrm{D}}$ fyller gapet, det vil si inntil saker henlegges eller foreldes i det omfanget overbelastningen krever. Linearisert rundt en stasjonær kø virker domstolen som et førsteordens lavpassfilter på strømmen av avgjorte saker, med tidskonstant $1/\big((1-R_{\mathrm{D}}/Z_{\mathrm{D}})^2/m_{\mathrm{D}} + \varsigma_{\mathrm{D}}\big)$, som er lang ved høy utnyttelse (28 måneder med parameterne i tabell 2): endringer i politiets innsats slår sent og svakt ut i avgjørelser. Ligning (28) beskriver én felles kø uten prioritering; i avsnitt 8.3 utvides den til to sakstyper med forrang for de alvorlige og til c parallelle spor, med (28) som eksakt grensetilfelle.

Inkapasitering og forsinket virkning. Fjerningsleddet $\gamma\Lambda^{\eta}TK_{\mathrm{obj}}$ i (12) lar hele virkningen av innsatsen inntre når politiet griper inn. En del av virkningen inntrer i virkeligheten først ved rettskraftig avgjørelse: inkapasitering gjennom soning, den individualpreventive virkningen av dommen og den allmennpreventive virkningen av at reaksjonen faktisk kommer. La $\vartheta \in [0,1]$ være denne andelen for saker som går til domstolen. Kriminalitetsdynamikken blir da

$$\frac{\mathrm{d}K_{\mathrm{obj}}(t)}{\mathrm{d}t} \;=\; \Omega^{\mathrm{eff}}(t) \;-\; \big(1-\xi_{\mathrm{D}}\vartheta\big)\,\gamma\,\Lambda(t)^{\eta}\,T(t)\,K_{\mathrm{obj}}(t) \;-\; \vartheta\,R_{\mathrm{D}}(t) \;-\; \delta\,K_{\mathrm{obj}}(t), \tag{29}$$

der den første fjerningsstrømmen virker umiddelbart (saker som avgjøres med forenklede reaksjoner, og den delen av virkningen for domstolssaker som inntrer ved inngrepet: avbrutt aktivitet, beslag, varetekt), mens den andre inntrer i takt med domstolens avgjørelser. Uten flaskehals ($Z_{\mathrm{D}} \to \infty$, $m_{\mathrm{D}} \to 0$) er $R_{\mathrm{D}} = J_{\mathrm{D}} = \xi_{\mathrm{D}}\gamma\Lambda^{\eta}TK_{\mathrm{obj}}$, og (29) faller eksakt sammen med (12). Med flaskehals er $R_{\mathrm{D}} \le Z_{\mathrm{D}}$ uansett hvor stor T er: andelen ϑ av virkningen av en opptrapping realiseres ikke, og saker som faller ut av køen, gir ingen virkning overhodet.

Saksbehandlingstid, avskrekking og tillit. Saksbehandlingstiden virker tilbake gjennom to kanaler. For det første diskonteres sanksjonen: lang tid fra lovbrudd til reaksjon reduserer nåverdien av F_i i (16) og dermed avskrekkingen, slik at tilstrømningen øker. For det andre svekker lang saksbehandlingstid tilliten til at rettsstaten virker. Begge kanalene uttrykkes ved en felles metningsfunksjon av overskridelsen av en referansetid $\tau_{\mathrm{D}}^{\mathrm{ref}}$, som settes lik saksbehandlingstiden i utgangslikevekten, den tiden samfunnet og lovbryterne har tilpasset seg:

$$g(\tau_{\mathrm{D}}^{\text{kø}}) = \frac{[\tau_{\mathrm{D}}^{\text{kø}} - \tau_{\mathrm{D}}^{\mathrm{ref}}]_+}{[\tau_{\mathrm{D}}^{\text{kø}} - \tau_{\mathrm{D}}^{\mathrm{ref}}]_+ + \tau_{\mathrm{h}}}, \qquad \Omega^{\mathrm{eff}} = \Omega(\mathbf{E})\,(1+\chi_{\mathrm{D}}\,g), \qquad \Lambda_0 \;\to\; \Lambda_0 - \mu_4\,g \;\text{ i (21).} \tag{30}$$

Funksjonen g er null ved referansetiden, øker omtrent lineært med stigning $1/\tau_{\mathrm{h}}$ for små overskridelser og nærmer seg 1 for svært lange køer; τ_{h} er den overskridelsen som gir halvparten av full virkning. χ_{D} er dermed den største relative økningen i tilstrømningen som kan tilskrives en verdiløs (uendelig forsinket) sanksjon, og μ_4 det største tapet i tillitens likevektsnivå som kan tilskrives saksbehandlingstiden. Metningen sikrer at ingen av kanalene vokser ubegrenset. Siden belegget for celerity-effekten er svakt, settes χ_{D} moderat, og avsnitt 8.2 viser resultatene også med $\chi_{\mathrm{D}} = 0$.

Hva som ikke tas med, og hva køen gjør med stabiliteten. Informasjonsforsinkelsen τ i (11) lar vi ikke avhenge av køen: styringen observerer registrert kriminalitet, ikke domstolsutfall, og den delen av sløyfens forsinkelse som skyldes domstolen, fanges allerede endogent av (29). Konsekvensen er en annen enn man kunne vente. For en mettet domstol er $R_{\mathrm{D}} \approx Z_{\mathrm{D}}$ en konstant som ikke påvirkes av T. Holder vi Λ og Ω^{eff} faste, gjelder lineariseringen i proposisjon 1 med

$$a' = \big(1-\xi_{\mathrm{D}}\vartheta\big)\,c\,\bar{T} + \delta, \qquad b' = \big(1-\xi_{\mathrm{D}}\vartheta\big)\,c\,\bar{K}\,k_p, \tag{31}$$

det vil si at dempningen og den effektive reaktiviteten skaleres med $1-\xi_{\mathrm{D}}\vartheta$, mens δ beholdes. Siden $b'/a' < b/a$, øker den kritiske forsinkelsen (14): en mettet domstol gjør sløyfen *mer* stabil i lineær forstand, men bare fordi den kobler ut en del av styringen. Prisen er et høyere likevektsnivå for kriminaliteten og en voksende restanse, og stabiliteten er lånt: når kapasiteten utvides og R_{D} igjen følger J_{D}, gjenvinner sløyfen hele forsterkningen b, og eventuelle svingninger vender tilbake (avsnitt 8.2). I tillegg kommer en langsom, positiv tilbakekobling gjennom χ_{D} (mer kriminalitet, flere saker, lengre kø, svakere avskrekking), som frafallet ς_{D} og metningen i g holder begrenset.

7.4 *Theory of Constraints* anvendt på straffesakskjeden

Innsikten om at en kjedes gjennomstrømning bestemmes av det svakeste leddet, er systematisert i Goldratts *Theory of Constraints* (ToC) (Goldratt og Cox, 1984; Goldratt, 1990). ToC foreskriver fem fokuseringstrinn: identifiser skranken, utnytt den fullt ut, underordne alt annet skranken, utvid skranken, og begynn på nytt når skranken har flyttet seg. Tabell 1 oversetter trinnene til straffesakskjeden. Tre poeng fortjener utdyping.

For det første: å *utnytte* flaskehalsen betyr at ingen rettsmøtetime går tapt. Utsatte hovedforhandlinger, saker som ikke er ferdig forberedt når de berammes, og mangelfull påtalekvalitet er kapasitetstap i det leddet der kapasiteten er knappest, og dermed tap for hele kjeden.

For det andre: *underordning* innebærer at flaskehalsens kapasitet styrer takten på *sakene som sendes inn i den* («trommel–buffer–tau» i Goldratts terminologi). I straffesakskjeden bør tauet virke på påtalebeslutningene og valget av reaksjonsform, ikke på politiets oppdagelse og forstyrrelse av kriminell aktivitet (avsnitt 8.2). Saker bør prioriteres inn i domstolen etter $A_i S_i^{\theta}$ i (2), den samme alvorlighetsvektingen som styrer innsatsen etter definisjon 1, mens saker med lav alvorlighet avgjøres med forenklede reaksjoner utenfor flaskehalsen: forelegg, påtaleunnlatelse og konfliktråd. I notasjonen ovenfor er dette en reduksjon i ξ_{D} (avsnitt 8.3 tallfester hva prioriteringen betyr for de alvorlige sakene). Tilståelsesdom avsies av tingretten og er en forenklet behandling *i* domstolen: den senker behandlingstiden m_{D}, ikke ξ_{D}.

Underordning har en rettslig side som modellen ikke fanger. Når andelen saker til domstol varierer med køen, kan to like handlinger møtes med ulik reaksjon avhengig av når og hvor saken kommer inn. Det er forenlig med Grunnloven § 98 bare dersom forskjellen følger av saklige, generelle kriterier som er fastsatt på forhånd, for eksempel i riksadvokatens direktiver, og ikke av køtallet i den enkelte sak. I praksis er det derfor sakstyper, ikke prosentandeler, som kan flyttes, og ξ_{D} i simuleringene må leses som virkningen av slike kriterier. Påtalemyndigheten er dessuten uavhengig i den enkelte sak (straffeprosessloven § 55) og har som utgangspunkt påtaleplikt (§ 62 a). Samordningen i avsnitt 7.7 kan derfor bare skje gjennom generelle retningslinjer og felles informasjon, ikke gjennom instruks. Modellen antar også at forenklede reaksjoner virker like godt som dom og ikke koster noe i leddene som overtar sakene. Begge deler er forenklinger som trekker i favør av underordning (avsnitt 10).

For det tredje understreker Goldratt at mange skranker er *regelskranker* (*policy constraints*), ikke fysiske: det er prosedyrer, prosessregler og arbeidsdeling som begrenser gjennomstrømningen, ikke antall dommere eller rettssaler. I straffesakskjeden kan prosessreglene være den egentlige flaskehalsen. Da er trinn 4 — å utvide skranken — et lovgivningsspørsmål, og kjedeanalysen kobles tilbake til avsnitt 7.1: forenklede prosessformer, utvidet adgang til skriftlig og digital bevisføring og klarere regler for beramming er endringer i $\mathcal{D}$, ikke i budsjettet.

Noen grenser ligger fast. Uskyldspresumsjonen, retten til kontradiksjon og retten til forsvarer er ikke effektiviseringsparametere, men selve innholdet i en rettferdig rettergang (Grunnloven §§ 95 og 96, EMK artikkel 6). Muntlighet og bevisumiddelbarhet er bærende prosessprinsipper, men de er lovfestede og har unntak, og hører dermed til det handlingsrommet lovgiver kan justere, slik forslaget om utvidet adgang til skriftlig og digital bevisføring

Tabell 1. Goldratts fem fokuseringstrinn anvendt på straffesakskjeden.

Trinn	ToC-prinsipp	Anvendt på straffesakskjeden
1	Identifiser skranken	Mål innstrømning, restanse og gjennomløpstid per ledd; beregn υ_k og $\tau_k^{\text{kø}}$ fra (26). Skyggeprisen på kapasitetsskranken i (32) er positiv bare i flaskehalsen.
2	Utnytt skranken	Ingen tapt kapasitet i flaskehalsen: færre utsatte rettsmøter, ferdig forberedte saker ved beramming, høy påtalekvalitet, digital bevisføring der rettssikkerheten tillater det.
3	Underordne alt annet	La flaskehalsens kapasitet styre hvor mange saker som sendes til domstolen (trommel–buffer–tau), ikke politiets innsats. Prioriter saker etter $A_i S_i^{\theta}$; led saker med lav alvorlighet til forenklede reaksjoner ($\xi_{\text{D}} \downarrow$).
4	Utvid skranken	Flere dommere og rettssaler, men også prosessuelle lovendringer og forenklede prosessformer: regelskranker krever lovgivning (avsnitt 7.1).
5	Gjenta	Skranken flytter seg, for eksempel til straffegjennomføringen. Mål på nytt, og la ikke institusjonell treghet bli den nye skranken.

forutsetter. Grensen går der en justering griper inn i retten til en rettferdig rettergang. Å utnytte flaskehalsen berører dessuten berammelse og saksstyring, som er dommernes ansvar; domstolenes uavhengighet (Grunnloven § 95 annet ledd) innebærer at slike tiltak må eies av domstolene selv. ToC anvendt på domstolene handler om å fjerne tap av kapasitet og om å lede saker dit de hører hjemme, ikke om å senke beviskravene eller redusere den tiltaltes rettigheter.

7.5 Kapasitetsskranker i optimeringen

Kjedemodellen gir to tillegg til skrankesettet (23):

$$\begin{aligned} & T_i^{\min} \le T_i(t) \le T_i^{\text{lov}}(t) \le T_i^{\text{grl}} && \text{(lovbestemt og konstitusjonell grense, jf. (25))} \\ & J_k(T(t)) \le \upsilon_k^{\max} Z_k(t) \quad \forall k && \text{(kapasitet med buffer i hvert ledd av kjeden)} \end{aligned} \tag{32}$$

der $\upsilon_k^{\max} < 1$ er den høyeste utnyttelsesgraden som gir akseptabel køtid. Kapasitetene Z_k er beslutningsvariabler på budsjettsyklusens tidsskala med egen ledetid (rekruttering, opplæring, lokaler), mens domenet $\mathcal{D}$ og de lovbestemte grensene er beslutningsvariabler på lovgivningens tidsskala. Modellen får dermed tre lag av beslutninger med hver sin forsinkelse, og MPC-problemet (24) løses for det raskeste laget gitt det som er bestemt i de to tregere. Kapasitetsskranken bør rettes mot ξ_{D} og Z_{D}; å håndheve den ved å begrense politiinnsatsen T gir et dårlig utfall (avsnitt 8.2).

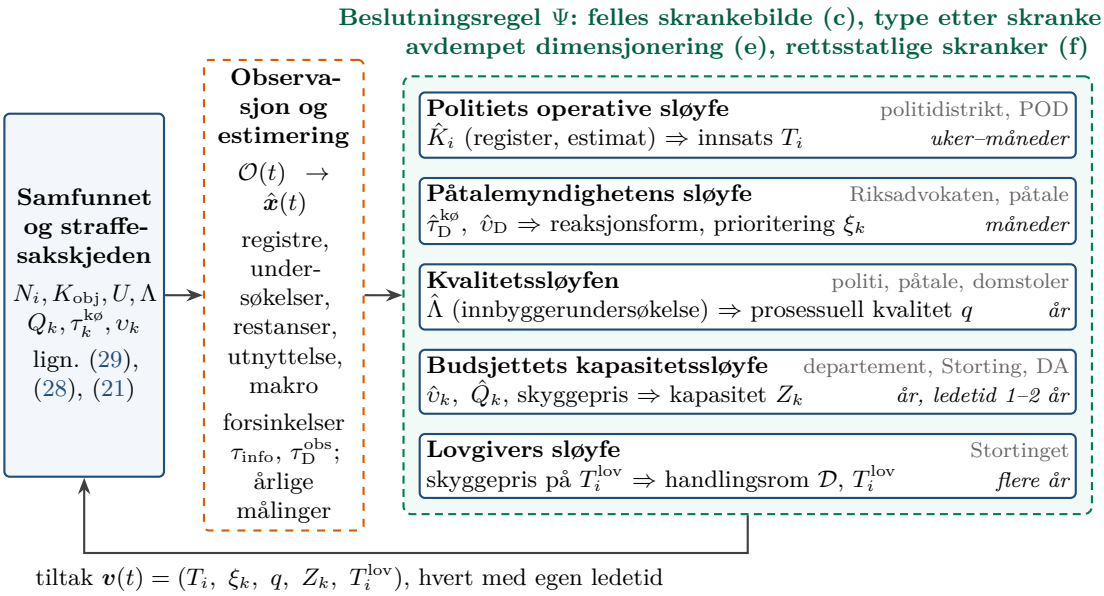
Koblingen til ToC er direkte. Skyggeprisen på kapasitetsskranken for ledd k, $-\partial \mathcal{J}^*/\partial Z_k$, er null for ledd med ledig kapasitet og positiv bare i flaskehalsen. Optimeringen *identifiserer* dermed skranken (trinn 1 i tabell 1) og tallfester *verdien* av å utvide den (trinn 4) i samme enhet som det samlede samfunnstapet. Tilsvarende måler skyggeprisen på T_i^{lov} hva det lovbestemte handlingsrommet er verdt, og gir lovgiver et kvantitativt grunnlag for å vurdere om det bør utvides — innenfor T_i^{grl}.

7.6 Regneeksempel: domstolen som flaskehals

Eksempel: Hva skal til for at domstolen holder tritt med en opptrapping?

Parameterne er som i tabell 2; strømmer måles i andeler av referansenivået per måned, og $\xi_{\text{D}} = 1$. Tallene er likevektsbetraktninger som dimensjonerer scenarioene i avsnitt 8.2.

1. Innstrømning før og etter sjokket. Før sjokket er fjerningsstrømmen $\gamma \Lambda T K_{\text{obj}} = 0{,}125 \cdot 0{,}8 \cdot 1 \cdot 1 = 0{,}10$ per måned. I den nye likevekten etter sjokket uten flaskehals (regime A: $K_{\text{obj}} \approx 1{,}086$, $T \approx 1{,}257$, $\Lambda \approx 0{,}787$) er den $0{,}125 \cdot 0{,}787 \cdot 1{,}257 \cdot 1{,}086 \approx 0{,}134$, en økning på 34 prosent.



Figur 3. Metoden som sløyfer i flere lag. Alle tiltak er responser på det observerte bildet $\mathcal{O}(t)$ og tilstandsestimatet $\hat{\boldsymbol{x}}(t)$, men med ulik eier, beslutningstakt og ledetid. Den stiplede rammen er beslutningsregelen Ψ i (9): sløyfene deler ett bilde av hvilken skranke som binder, og dimensjoneringen regner med de andre sløyfenes tiltak. Lovgivers sløyfe gjøres ikke raskere, men gir handlingsrommet de øvrige virker innenfor.

2. Utnyttelse og kø. Anta at domstolene før sjokket arbeidet med 95 prosent utnyttelse, $Z_{\text{D}} = 0{,}10/0{,}95 \approx 0{,}105$. Etter sjokket er $\upsilon_{\text{D}} = 0{,}134/0{,}105 \approx 1{,}28$, altså 28 prosent overbelastning. Antall avgjorte saker øker ikke; restansen vokser med $0{,}134 - 0{,}105 = 0{,}029$ per måned, og køtiden øker etter (26) med $\upsilon_{\text{D}} - 1 \approx 0{,}28$ måneder per måned, det vil si om lag 3,4 måneder for hvert år overbelastningen varer.

3. Underordning. Ledes 20 prosent av sakene til forenklede reaksjoner utenfor domstolen ($\xi_{\text{D}} = 0{,}8$), faller innstrømningen til 0,107 og υ_{D} til 1,02. Overbelastningen er nesten borte, men restansen avvikles ikke, og ventetiden er svært følsom når υ_{D} ligger nær 1.

4. Utvidelse. Skal utnyttelsen tilbake til 95 prosent med full innstrømning, må kapasiteten økes med 34 prosent (til 0,141). Kombinert med underordning er en økning på 7 prosent (til 0,113) tilstrekkelig. Skyggeprisen på Z_{D} i (32) angir hva utvidelsen er verdt målt i samfunnstap, og om den bør skje gjennom flere dommere eller gjennom prosessuelle lovendringer (avsnitt 7.1).

Diagnose: en opptrapping av politiinnsatsen med 26 prosent som ikke følges av kapasitet nedstrøms, omsettes i restanser. Likevektsbetraktningene sier hva som skal til for å gjenopprette balansen, men ikke hva som skjer underveis, når køen virker tilbake på kriminalitet, innsats og tillit. Det viser simuleringen i avsnitt 8.2.

7.7 Metoden: fra observasjon til tiltak

Avsnittene 2–7.5 gir hver sin del av svaret på hvordan tiltak bør velges: måleligningen sier hva som kan observeres, stabilitetsanalysen hvor kraftig det kan reageres, tillitsdynamikken hvilke tiltak som verner legitimiteten, og *Theory of Constraints* hvor i kjeden et tiltak virker. Beslutningsregelen (9) binder delene sammen, men sier ikke hvordan Ψ skal konstrueres. Vi foreslår at den realiseres som en syklisk metode i sju trinn:

(a) *Observer og estimer.* Dann $\hat{\boldsymbol{x}}(t)$ fra $\mathcal{O}(t)$ gjennom måleligningen (4), med usikkerhet, medregnet innstrømning, restanse, saksbehandlingstid og utnyttelse per ledd (26), undersøkelsene og makroindikatorene (5).

(b) *Sammenlign med målene, alvorlighetsvektet.* Avvikene beregnes per kategori med vektene

Tabell 2. Parametere i den numeriske illustrasjonen (tidsenhet: måned).

Parameter	Verdi	Tolkning
$\gamma,\ \eta$	0,125, 1	effektivitet; gir fjerningsrate $\gamma\Lambda T = 0{,}10$ i likevekt
δ	0,02	naturlig avgang
Ω_0	$0{,}12 \to 0{,}156$	strukturelt pådrag før/etter sjokket i måned 12
$T_0,\ T^{\max},\ K^*$	1, 3, 1	grunninnsats, kapasitetstak, målnivå
k_p	3	politisk reaktivitet (proporsjonal regel)
$\kappa_U,\ U_0,\ \psi$	0,25, 0,80, 0,30	trygghetsdynamikk
$\kappa_\Lambda,\ \Lambda_0,\ \mu_1,\ \mu_2$	0,10, 0,80, 0,50, 0,30	tillitsdynamikk ($\mu_3 = 0$)
$T_{\text{prop}}(K)$	$1{,}6\,K$	grense for forholdsmessig innsats
MPC: H, λ_T, λ_Λ, Λ_{krit}	24, 0,05, 20, 0,75	horisont, vekt på innsatsavvik, vekt på tillitsbrudd, myk tillitsskranke
Kjede: $Z_{\text{D}},\ m_{\text{D}},\ \varsigma_{\text{D}}$	0,105, 0,233, 0,01	domstolskapasitet ($\upsilon_{\text{D}} = 0{,}95$ før sjokket); ren behandlingstid, kalibrert slik at $\tau_{\text{D}}^{\text{ref}} = 3$ mnd.; frafall fra køen per måned
$\vartheta,\ \chi_{\text{D}},\ \mu_4,\ \tau_{\text{h}}$	0,5, 0,15, 0,12, 9	andel av virkningen som krever dom; største celerity-økning i Ω; største tillitstap av saksbehandlingstid; halvmetningstid (mnd.), jf. (29) og (30)
Robusthet: K_{sat}	∞; 3; 1,5	politiets halvmetningsnivå, jf. (12′); ∞ i hovedanalysen
$\pi_{\text{A}},\ \vartheta_{\text{A}},\ \vartheta_{\text{B}},\ c_{\text{D}}$	0,4, 0,8, 0,3, 1 (10)	alvorlige sakers andel av K_{obj}; andel av virkningen som krever dom per klasse (vektet 0,5); antall parallelle spor i domstolen, jf. (33)
$\epsilon,\ K_{\text{h}},\ \varrho$	0,30, 2, 4	rekrutteringsledd: styrke, halvmetningsnivå og bratthet, jf. (34)
Regler: $\tau_{\text{D}}^{\text{obs}},\ \xi_{\min},\ \tau_{\text{D}}^{\text{gr}},\ k_\xi,\ \tau_\xi$	3; 0,8; 6; 0,2/år; 6	observasjonsforsinkelse for kjedestatistikk; gulv for andel saker til domstol; terskel for saksbehandlingstid; forsterkning og tidskonstant for påtalepraksis, jf. (35)
$\upsilon^*,\ k_Z,\ \tau_Z,\ \tau_{\text{ned}},\ H_Q,\ H_{\text{p}}$	0,95; 1; 6–24; 36; 24; 36	målutnyttelse; budsjettforsterkning; realisering etter vedtak; nedbyggingstid; horisont for restanseavvikling; prognosehorisont (mnd.), jf. (36)–(37); bestilling når $\hat{\upsilon}_{\text{D}} > 1$, nedbygging når $\hat{\upsilon}_{\text{D}} < 0{,}85$
$\mu_3,\ \Lambda^*,\ k_q,\ q_{\max}$	0,5; 0,78; 1/år; 0,2	kvalitetssløyfen: tillitsvirkning av q, tillitsmål, forsterkning, øvre grense (avsnitt 8.5)
$a,\ b$	0,12, 0,30	dempning og effektiv reaktivitet i referanselikevekten, jf. (13)
τ_c	7,2 mnd.	kritisk forsinkelse fra (14)

$A_i S_i^\theta$ fra (2), altså etter hvor mye og hvor alvorlig lovene brytes, og for trygghet, tillit og saksbehandlingstid mot U^*, Λ_{krit} og τ_k^{ref}.

(c) *Identifiser den bindende skranken i hele kjeden* (ToC trinn 1) ved utnyttelsesgrader, restansevekst og skyggeprisene på kapasitets-, budsjett-, tillits- og lovskrankene i (23) og (32); som regel binder én om gangen.

(d) *Velg type tiltak etter skranken* (ToC trinn 2–4): binder politiets kapasitet: innsats og organisering (T_i); binder domstolen: først utnytte, så underordne ($\xi_{\text{D}} \downarrow$, prioritering etter $A_i S_i^\theta$) og bare om det ikke rekker, utvide ($Z_{\text{D}} \uparrow$); nærmer tilliten seg Λ_{krit}: prosessuell kvalitet q, ikke volum (merknad 1); er skyggeprisen på T_i^{lov} vedvarende positiv: et signal til lovgiver (avsnitt 7.1).

(e) *Dimensjoner mengde og tidspunkt med reguleringsteorien.* Reaksjonen holdes innenfor stabilitetsgrensen for den aktuelle sløyfen (proposisjon 1 og korollar 1), lavere jo svakere

observerbarheten er, men tidlig og tilstrekkelig der vippepunkter er plausible (avsnitt 8.4). Virkningen av tiltak som er vedtatt, men ennå ikke har virket, regnes inn før nye vedtas, og kapasitet dimensjoneres etter strømmen som gjenstår etter underordning og etter restansen som skal avvikles, ikke etter strømmen flaskehalsen selv har blåst opp (avsnitt 8.5).

(f) *Kontroller de rettsstatlige skrankene:* T_i^{grl} og T_i^{lov} (25), saklig begrunnelse i omfang og alvorlighet (Grunnloven § 98), forholdsmessighet (T_{prop}) og avgjørelse innen rimelig tid.

(g) *Iverksett, evaluer og gjenta* (ToC trinn 5): logg prognosefeil, re-estimer, og begynn på nytt, fordi skranken flytter seg når den forrige er løst.

Figur 3 viser hvordan metoden fordeler seg på fem sløyfer med hver sin eier og tidsskala. Hver sløyfe er en tilbakekobling med forsinkelse, og innsikten i proposisjon 1 gjelder for hver av dem, eksakt der sløyfen har samme form (merknad 4) og ellers som analogi: en langsom sløyfe tåler bare lav reaktivitet og skal ikke gjøres rask, men gis et handlingsrom på forhånd (avsnitt 7.1). Det som skiller metoden fra at hver aktør svarer på sitt eget signal, er trinn (c) og (e): ett felles bilde av hvilken skranke som binder, og en dimensjonering som regner med de andre aktørenes tiltak og med tiltak som allerede er vedtatt. Avsnitt 8.5 undersøker hva forskjellen betyr, og algoritme 1 i avsnitt 9 gjør metoden operativ.

8 Numerisk illustrasjon

8.1 Fem styringsregimer

For å illustrere dynamikken simulerer vi den aggregerte modellen ($n = 1$) med ligningene (12), (20) og (21) og månedlig tidsenhet. Parameterne er stiliserte og valgt slik at referansesituasjonen er en likevekt med $K_{\text{obj}} = 1$, $T = 1$ og $\Lambda = 0{,}8$; de er ikke estimert på norske data (tabell 2). I måned 12 inntreffer et varig makroøkonomisk sjokk som øker pådraget Ω med 30 prosent. Vi sammenligner fem styringsregimer:

(A) proporsjonal regel (11) med $k_p = 3$ og forsinkelse $\tau = 3$ måneder;

(B) samme regel med $\tau = 6$ måneder;

(C) samme regel med $\tau = 11$ måneder (for eksempel statistikk 3, budsjett 6 og iverksetting 2);

(D) forenklet MPC med $\tau = 11$ måneder: myndighetene kjenner ikke sjokket før det er observert, men predikerer virkningen av egne vedtatte tiltak og velger hver måned ett konstant innsatsnivå for de neste 24 månedene som minimerer et kvadratisk tap med en myk skranke $\Lambda \geq 0{,}75$;

(E) avdempet proporsjonal regel med $k_p = 1{,}2$ og $\tau = 11$ måneder, det vil si samme forsinkelse som C, men med reaktiviteten på grensen for stabilitet uansett forsinkelse.

Proposisjon 1 gir $\tau_c = \arccos(-0{,}4)/\sqrt{0{,}30^2 - 0{,}12^2} \approx 7{,}2$ måneder i referanselikevekten; i likevekten etter sjokket ($\bar{K} \approx 1{,}09$, $\bar{T} \approx 1{,}26$, $\bar{\Lambda} \approx 0{,}79$) er grensen om lag 7,1 måneder. Regime A og B ligger dermed under stabilitetsgrensen og regime C over. Figur 4 bekrefter dette. Med kort forsinkelse (A) stabiliserer kriminaliteten seg raskt. Proporsjonalregelen etterlater et varig avvik ($K_{\text{obj}} \approx 1{,}09$), som et integrerende ledd ville fjernet. Med $\tau = 6$ (B) oppstår dempede svingninger. Med $\tau = 11$ (C) vokser svingningene til en grensesyklus med periode på om lag 28 måneder: innsatsen pendler mellom 0,7 og 2,2, kriminaliteten mellom 0,91 og 1,41, og i hver fase med overhåndhevelse faller tilliten, ned mot 0,47. Lavere tillit reduserer igjen effektiviteten og forsterker neste oppsving. Svingningene er ikke tilsiktet: de oppstår selv om hver enkelt beslutning er bevisst og rimelig ut fra de tallene som foreligger når den tas, fordi tallene beskriver en situasjon som allerede er endret av tiltak som ennå ikke har virket. Det er sløyfen, ikke den enkelte beslutningen, som skaper mønsteret.

MPC-regimet (D) har nøyaktig samme informasjonsforsinkelse. Den første toppen ($K_{\text{obj}} \approx 1{,}22$) er derfor nesten like høy som i C: ingen styringsregel kan reagere på et sjokk før det er observert. Forskjellen ligger i det som skjer etterpå. Fordi MPC tar hensyn til at allerede vedtatte tiltak ennå ikke har virket, unngår den overreaksjonen. Innsatsen topper seg på 1,4, kriminaliteten stabiliserer seg nær målet ($K_{\text{obj}} \approx 1{,}03$), og tilliten holder seg over 0,78.

Sammenligningen mellom C og D blander to ting: prognosen og reaksjonsstyrken. Regime E skiller dem. Med samme forsinkelse som C, men avdempet reaksjon, uteblir svingningene: første topp er 1,23, innsatsen kommer aldri over 1,3, og tilliten holder seg over 0,77, på linje med MPC. Prisen er et større varig avvik ($K_{\text{obj}} \approx 1{,}16$, mot 1,09 i A og 1,03 med MPC), fordi en svakere proporsjonal respons lukker mindre av gapet. Det viktigste grepet mot pendelsvingninger er altså å dempe reaksjonen. Det prognosebasert styring tilfører ut over dette, er at det varige avviket kan fjernes uten at stabiliteten ofres. MPC-regimet har dessuten fordeler det ikke vil ha i praksis: det kjenner modellen eksakt og observerer trygghet og tillit månedlig (avsnitt 8.2 viser hva feil modell koster, og avsnitt 8.5 bruker årlige tillitsmålinger).

Hvor representative er parameterne?. To valg driver svingningene i regime C, og begge ligger trolig i ytterkant. Reaktiviteten $k_p = 3$ betyr at ti prosent mer kriminalitet utløser 30 prosent mer innsats. Vi kjenner ingen budsjettprosess som er dokumentert å reagere så kraftig, og for $k_p \leq 1{,}2$ er sløyfen stabil uansett forsinkelse. Effekten av innsats er også sterk: parameterne gir en likevektselastisitet for kriminaliteten med hensyn på innsatsen på $-c\bar{T}/(c\bar{T}+\delta) = -0{,}10/0{,}12 \approx -0{,}83$, mens de fleste anslag i litteraturen ligger mellom $-0{,}1$ og $-0{,}5$ (Chalfin og McCrary, 2017). En svakere effekt gir lavere b og større stabilitetsmargin. Regime C viser derfor hva som *kan* skje når reaksjonen er kraftig og forsinkelsen lang, ikke hva som skjer i norsk rettshåndhevelse i dag. Om den faktiske sløyfen ligger nær stabilitetsgrensen, må avgjøres ved å estimere reaktivitet og forsinkelse på norske tidsserier (avsnitt 9.3 og avsnitt 10). Ligger den trygt innenfor, er det kjede- og kapasitetsanalysen i resten av avsnittet som har størst praktisk interesse.

Figur 5 viser stabilitetsgrensen (14) som funksjon av reaktiviteten k_p, sammen med den konservative tommelfingerregelen (15). For $k_p \leq a/(c\bar{K}) = 1{,}2$ er sløyfen stabil uansett forsinkelse. Over denne verdien faller den tillatte forsinkelsen raskt. Ved $\tau = 11$ måneder er den største stabile reaktiviteten $k_p \approx 2{,}3$, mens tommelfingerregelen gir den forsiktigere grensen $k_p < \pi/(2 \cdot 0{,}1 \cdot 11) \approx 1{,}4$.

8.2 Simulering med straffesakskjeden

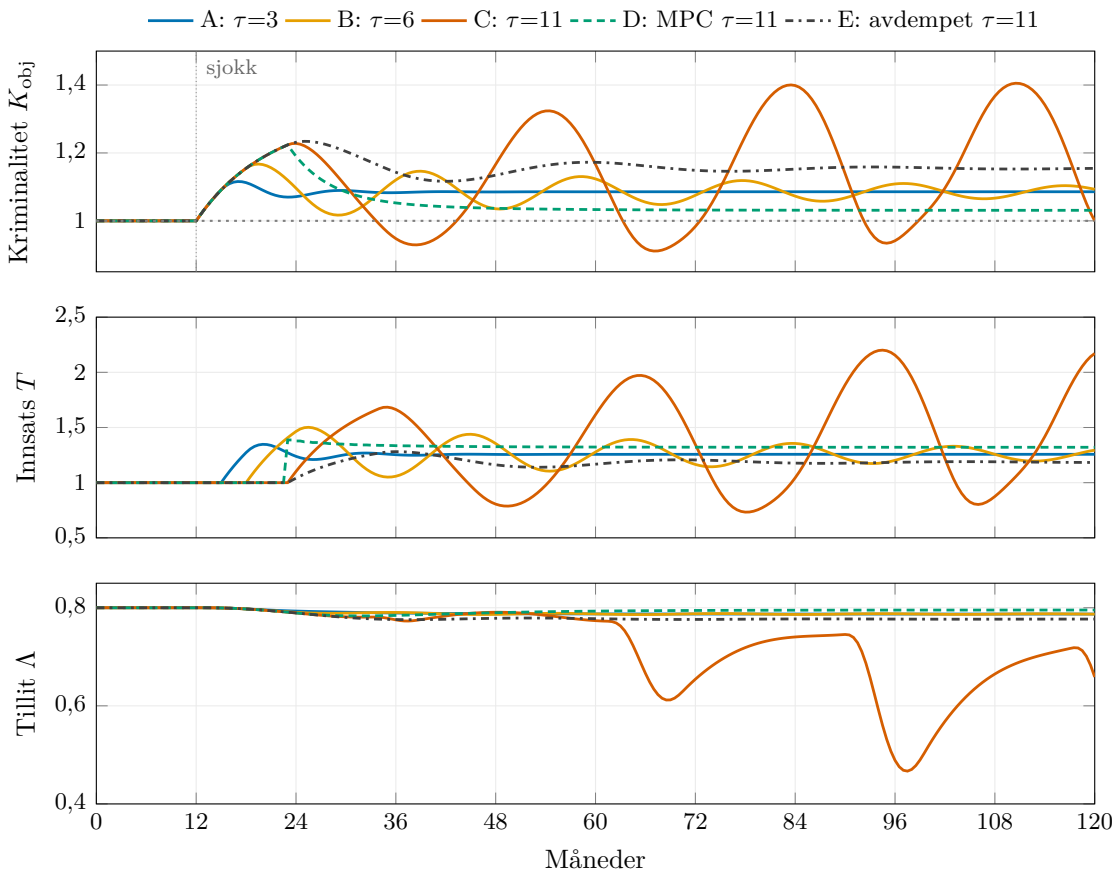
Vi kobler nå domstolsleddet til dynamikken etter avsnitt 7.3: (12) erstattes av (29), køen følger (28), og tilstrømning og tillit påvirkes etter (30). Parameterne står nederst i tabell 2. Domstolen er kalibrert som i avsnitt 7.6: før sjokket arbeider den med $\upsilon_{\text{D}} = 0{,}95$ og en saksbehandlingstid på 3 måneder, hvorav den rene behandlingstiden utgjør 0,23 måneder; 2,9 prosent av sakene faller ut av køen. For å isolere virkningen av kjeden bruker vi det stabile regimet A ($k_p = 3$, $\tau = 3$) og samme sjokk som før, og sammenligner fem scenarioer:

(1) *Ingen flaskehals* ($Z_{\text{D}} \to \infty$). Dette er regime A fra figur 4; den koblede modellen gjengir det til maskinpresisjon (største avvik i K_{obj} under 10^{-14}, og tilsvarende for regime B–D), og likeså med endelig kapasitet når $\vartheta = \chi_{\text{D}} = \mu_4 = 0$. Med stor, men endelig kapasitet ($Z_{\text{D}} = 1$) er største avvik 0,005.

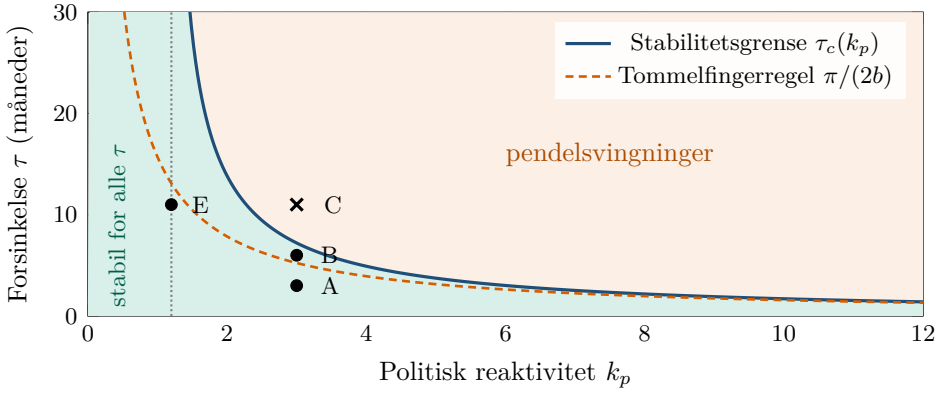
(2) *Domstolen som flaskehals, ingen tiltak*: innsatsen trappes opp etter (11) som før.

(3) *Underordning* (trinn 3 i tabell 1): fra måned 24, når restansen har vært synlig i ett år, ledes 20 prosent av sakene til forenklede reaksjoner (ξ_{D}: $1 \to 0{,}8$).

(4) *Utvidelse* (trinn 4): kapasiteten økes med 34 prosent (til 0,141), besluttet i måned 24 og realisert lineært fram til måned 48 (rekruttering, opplæring, lokaler og lovendringer tar



Figur 4. Simulert respons på et varig makrosjokk (Ω +30 % i måned 12) under fem styringsregimer. A–C: proporsjonal regel med $k_p = 3$ og økende forsinkelse. D: forenklet modellprediktiv styring med samme forsinkelse som C. E: avdempet proporsjonal regel ($k_p = 1{,}2$) med samme forsinkelse som C. Prikket horisontal linje: målnivå $K^* = 1$.



Figur 5. Stabilitetskart for den lineariserte sløyfen med parameterne i tabell 2. Til venstre for den stiplede vertikale linjen er systemet stabilt for alle forsinkelser. Punktene A–C og E viser de proporsjonale regimene i figur 4.

tid).

(5) *Underordning og utvidelse* kombinert.

Figur 6 viser forløpene. *Uten tiltak* (2) blir utfallet varig dårligere på alle mål. Innstrømningen til domstolen overstiger kapasiteten fra sjokket av, og saksbehandlingstiden er nær 8 måneder etter ett år, 22 måneder etter tre år og 54 måneder ved slutten av perioden.

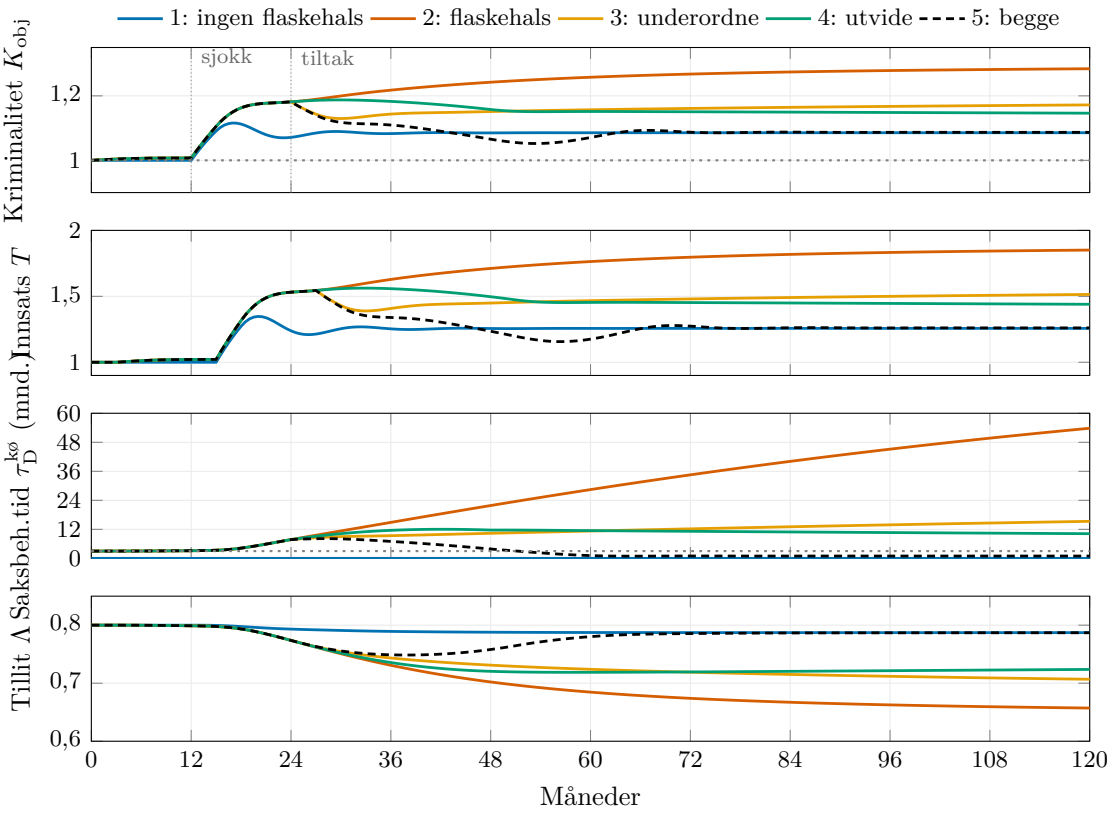
Styringsregelen svarer på den høye kriminaliteten med å trappe opp innsatsen til 1,85, men den delen av virkningen som forutsetter dom, uteblir: domstolen avgjør 0,105 saker per måned uansett, innstrømningen er 0,195, og 0,056 per måned, mer enn en fjerdedel av sakene, henlegges eller foreldes. Kriminaliteten stabiliserer seg på 1,28 (mot 1,09 uten flaskehals) og tilliten på 0,66 (mot 0,79): innsatsen omsettes i kø og frafall i stedet for reaksjoner, og den lengre saksbehandlingstiden svekker avskrekking og tillit.

Underordning (3) er det eneste virkemidlet som virker raskt, og det krever ingen ny kapasitet: veksten i køen bremses kraftig (15 måneder ved slutten, mot 54), kriminaliteten ender på 1,17 og tilliten på 0,71, og innsatsen (1,51) er lavere enn i scenario 2 fordi kriminaliteten er lavere. Domstolen er likevel fortsatt overbelastet ($\upsilon_{\mathrm{D}} = 1{,}21$), så restansen fortsetter å vokse langsomt. *Utvidelse alene* (4) gir et liknende sluttresultat for kriminalitet og tillit (1,15 og 0,72), men virker senere: saksbehandlingstiden topper seg på 12 måneder i måned 42 og er fortsatt 10 måneder ved slutten, fordi utvidelsen er dimensjonert for strømmen og ikke for den beholdningen som ble bygget opp i de 36 månedene før kapasiteten var på plass. Først *kombinasjonen* (5) gjenoppretter referansen: saksbehandlingstiden topper seg på 8 måneder i måned 28 og er tilbake under referansetiden fra måned 52; kriminalitet, innsats og tillit ender på 1,09, 1,26 og 0,79, som i scenario 1. Rekkefølgen i *Theory of Constraints*, underordne først og utvid deretter, er altså ikke bare organisatorisk fornuftig, men nødvendig i modellen: underordning er det som virker raskt, utvidelse det som gjenoppretter balansen varig. Tiltakene er her lagt inn på faste tidspunkter og i full størrelse, dimensjonert med kjennskap til den nye likevekten; om rekkefølgen og dimensjoneringen oppstår av seg selv når tiltakene i stedet er responser på observasjoner, undersøkes i avsnitt 8.5.

Å underordne seg ved å holde igjen. Hva om politiet i stedet begrenser innsatsen slik at $J_{\mathrm{D}} \leq 0{,}95\, Z_{\mathrm{D}}$, det vil si håndhever kapasitetsskranken i (32) gjennom T? Med $\xi_{\mathrm{D}} = 1$ gir dette det dårligste utfallet av alle: innsatsen presses ned til 0,54, kriminaliteten stiger til 2,66 og tilliten faller til 0,56, fordi fjerningen av pågående aktivitet begrenses til det domstolen kan avgjøre, mens tilstrømningen er upåvirket. Kombinert med underordning ($\xi_{\mathrm{D}} = 0{,}8$) ender kriminaliteten på 1,59, mens tilliten (0,71) og saksbehandlingstiden (3 måneder) er bedre enn i scenario 2: moderat innsats med underordning gir kortere kø og høyere tillit enn opptrapping uten tiltak, men høyere kriminalitet. Kapasitetsskranken bør derfor håndheves gjennom ξ_{D} og Z_{D}, ikke gjennom T.

Modellprediktiv styring med og uten kjedemodell. MPC er bare så god som modellen den bygger på. En MPC som bruker (12) uten kjede som intern modell, men styrer det koblede systemet i scenario 2, overvurderer virkningen av egen innsats og ender med $K_{\mathrm{obj}} = 1{,}52$, dårligere enn proporsjonalregelen (1,28). Med (28)–(30) som intern modell oppnår MPC 1,22 ved flaskehals og 1,07 med underordning (mot 1,17 for proporsjonalregelen), men kan ikke gjøre noe med køen så lenge T er eneste virkemiddel (57 og 18 måneders saksbehandlingstid); en hard skranke $J_{\mathrm{D}} \leq \upsilon_{\mathrm{D}}^{\max} Z_{\mathrm{D}}$ i (24) tilsvarer trommel–buffer–tau ovenfor. Nytten av kjedemodellen i MPC er at prediksjonen av køen gjør kapasitetsproblemet synlig for dem som rår over ξ_{D} og Z_{D}.

Følsomhet. Celerity-effekten er den mest usikre antakelsen. Tabell 3 viser sluttverdiene for χ_{D} lik null, moderat og sterk. Uten celerity-effekt er hovedbildet det samme: flaskehalsen gir $K_{\mathrm{obj}} = 1{,}19$, $\Lambda = 0{,}68$ og 36 måneders saksbehandlingstid, underordning 1,11, 0,74 og 8 måneder. Rangeringen av scenarioene er uavhengig av χ_{D}, og kombinasjonen (5) er upåvirket, fordi køen der ikke overskrider referansetiden lenge. Andelen ϑ betyr mer for nivåene: med $\vartheta = 0$ ender scenario 2 på $K_{\mathrm{obj}} = 1{,}16$, med $\vartheta = 0{,}75$ på 1,48. Parameteren μ_4 flytter først og fremst tilliten (0,77 uten kanalen, 0,55 med dobbel styrke), og ς_{D} først og



Figur 6. Simulert respons på samme makrosjokk som i figur 4 under regime A ($k_p = 3$, $\tau = 3$) når domstolskøen virker tilbake på dynamikken (avsnitt 7.3). 1: ingen flaskehals (identisk med regime A i figur 4). 2: domstol med 95 prosent utnyttelse før sjokket, ingen tiltak. 3: 20 prosent av sakene ledes til forenklede reaksjoner fra måned 24. 4: kapasiteten økes med 34 prosent over månedene 24–48. 5: begge tiltak. Prikkede horisontale linjer: målnivå $K^* = 1$ og referansetiden $\tau_D^{ref} = 3$ måneder. Saksbehandlingstiden i scenario 1 er ikke definert og vises som null.

fremst køen (88 måneder uten frafall, 36 med dobbelt frafall), mens kriminaliteten ligger på 1,27–1,29 i alle tilfellene.

Tabell 3. Følsomhet for celerity-parameteren χ_D: kriminalitet K_{obj}, tillit Λ og saksbehandlingstid $\tau_D^{kø}$ (måneder) i måned 120 for scenarioene i figur 6.

	$\chi_D = 0$ (ingen)			$\chi_D = 0{,}15$ (moderat)			$\chi_D = 0{,}30$ (sterk)		
Scenario	K_{obj}	Λ	$\tau_D^{kø}$	K_{obj}	Λ	$\tau_D^{kø}$	K_{obj}	Λ	$\tau_D^{kø}$
2: flaskehals, ingen tiltak	1,19	0,68	36	1,28	0,66	54	1,38	0,64	74
3: underordning	1,11	0,74	8	1,17	0,71	15	1,25	0,68	27
5: underordning og utvidelse	1,09	0,79	1	1,09	0,79	1	1,09	0,79	1

Køen og stabiliteten. For regime C ($\tau = 11$) endrer flaskehalsen svingningsbildet. Uten kjede vokser svingningene til en grensesyklus med topper på 1,40 og bunner på 0,91. Med en mettet domstol dempes de: toppene ligger på 1,32, 1,32 og 1,31, bunnene stiger fra 1,16 til 1,27, og perioden øker fra om lag 28 til 35 måneder. Dette er ligning (31) i praksis: i sluttilstanden til scenario 2 er $a' = 0{,}096$ og $b' = 0{,}158$, som gir $\tau_c' = 17{,}7$ måneder, mens samme tilstand uten metning ville gitt $\tau_c = 8{,}1$. Dempningen er lånt: gjennomføres underordning og utvidelse i regime C, kobles styringen inn igjen, og svingningene vender tilbake med nesten full styrke

Tabell 4. Robusthet for politiets kapasitet: lineær stabilitet i referanselikevekten, regime C uten kjede, og kjedescenarioene (sluttverdier i måned 120). k_p^∞ er den største reaktiviteten som er stabil uansett forsinkelse.

			Regime C		2: flaskehals			3	4	5
K_{sat}	τ_c (mnd.)	k_p^∞	$K_{\text{obj}}^{\max}$	$\Lambda^{\min}$	K_{obj}	Λ	$\tau_{\text{D}}^{\text{kø}}$	K_{obj}	K_{obj}	K_{obj}
∞ (grunnmodell)	7,2	1,20	1,41	0,47	1,28	0,66	54	1,17	1,15	1,09
3	6,7	0,95	1,47	0,37	1,32	0,65	53	1,19	1,16	1,09
1,5	6,4	0,80	1,51	0,33	1,34	0,65	52	1,20	1,16	1,10

(topper 1,31, 1,21 og 1,30; bunner 0,97, 0,95 og 0,91). En stabilitetsdiagnose (avsnitt 9.3) basert på data fra en periode med mettede domstoler kan derfor undervurdere sløyfens reaktivitet, og en kapasitetsutvidelse bør følges av en ny vurdering av k_p og τ.

8.3 Robusthet: politiets kapasitet og prioriterte domstolskøer

Grunnmodellen i avsnittene 8.1 og 8.2 bygger på to forenklinger som kan trekke resultatene i hver sin retning: et politi uten kapasitetsgrense, og en domstol som lar restansen ramme de alvorligste sakene like hardt som de øvrige. Vi løsner dem hver for seg med figurene 4–6 som referanse. Begge utvidelsene har grunnmodellen som eksakt grensetilfelle: med $K_{\text{sat}} \to \infty$ gjengis regime A–C og scenario 2 uten avvik, og med to sakstyper uten forrang og lik ϑ gjengis scenario 2 med største avvik under 10^{-14} i K_{obj}.

Politiets kapasitet. Vi erstatter fjerningsleddet med (12′) og lar halvmetningsnivået være $K_{\text{sat}} = 3$ (politiet er merkbart presset først ved en tredobling av belastningen) og $K_{\text{sat}} = 1{,}5$ (presset allerede ved 50 prosent økning), med γ' kalibrert slik at referanselikevekten er uendret (tabell 4). Retningen følger merknad 2: dempningen faller, reaktiviteten står, og den kritiske forsinkelsen krymper fra 7,2 til 6,7 og 6,4 måneder i referanselikevekten og fra 7,1 til 6,5 og 6,3 måneder i likevekten etter sjokket. Regime B ($\tau = 6$) er fortsatt stabilt, men marginen faller fra 1,1 til 0,3 måneder, og svingningene dør langt langsommere ut: etter 100 måneder pendler kriminaliteten fortsatt mellom 1,04 og 1,16 med $K_{\text{sat}} = 1{,}5$, mot 1,07 og 1,11 uten metning. I regime C blir grensesyklusen kraftigere på alle mål (figur 7, venstre): toppene stiger fra 1,41 til 1,47 og 1,51, innsatsen pendler opp til 2,4 og 2,5 i stedet for 2,2, og tillitens bunnpunkt faller fra 0,47 til 0,37 og 0,33. Mekanismen er den fallende oppklaringsintensiteten: ved toppen i regime C er fjerningen per lovbrudd for gitt innsats 17 prosent lavere med $K_{\text{sat}} = 1{,}5$ enn i referansen, slik at hvert oppsving krever mer innsats, og mer innsats gir større tillitstap. Dette er modellens kvalitative motstykke til at oppklaringsandelen faller når sakstilfanget øker (avsnitt 1). I kjedescenarioene forskyves nivåene moderat i samme retning (scenario 2: K_{obj} fra 1,28 til 1,34 og innsatsen fra 1,85 til 2,01 med $K_{\text{sat}} = 1{,}5$), mens saksbehandlingstid og tillit er nesten upåvirket og rangeringen av tiltakene står (begge tiltak 1,10, utvidelse 1,16, underordning 1,20, ingen tiltak 1,34). Modellprediktiv styring tåler metningen selv når dens interne modell ikke kjenner den: regime D anvendt på det mettede systemet gir en første topp på 1,25 og 1,27 (mot 1,22) og ender på 1,04 og 1,05, med tilliten på om lag 0,78 gjennom hele forløpet. Politimetning forskyver altså resultatene i avsnitt 8.1 i den retningen som gjør forsiktighet viktigere, og lar resultatene i avsnitt 8.2 stå.

Prioriterte domstolskøer og flere spor. Vi deler sakene i to typer, alvorlige (A) og øvrige (B), med andel π_{A} av den alvorlighetsvektede belastningen i type A, og lar hver type ha egen kø, egen saksbehandlingstid og egen andel ϑ_k av virkningen som krever dom. Domstolen gir

Tabell 5. Domstol med to sakstyper: sluttverdier i måned 120 under regime A med felles kø (FIFO) og med forrang for alvorlige saker. K_{A} og K_{B} er den alvorlighetsvektede belastningen i hver type relativt til referansen (0,4 og 0,6); τ_{A} og τ_{B} er saksbehandlingstid i måneder. Uten flaskehals er $K_{\mathrm{obj}} = 1{,}09$ og $\Lambda = 0{,}79$.

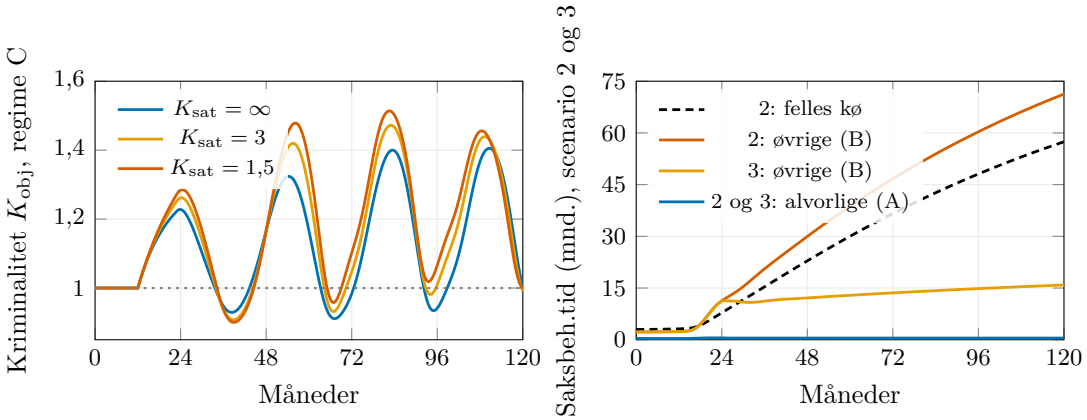
	Felles kø (FIFO)					Forrang for alvorlige saker					
Scenario	K_{obj}	K_{A}	K_{B}	Λ	$\tau^{\mathrm{kø}}$	K_{obj}	K_{A}	K_{B}	Λ	τ_{A}	τ_{B}
2: flaskehals, ingen tiltak	1,30	1,52	1,15	0,65	57	1,18	1,00	1,30	0,71	0,5	71
3: underordning	1,19	1,28	1,13	0,70	19	1,13	1,06	1,18	0,74	0,5	16
4: utvidelse	1,15	1,17	1,13	0,72	10	1,11	1,07	1,13	0,76	0,4	6,5
5: underordning og utvidelse	1,09	1,09	1,08	0,79	1,0	1,09	1,09	1,09	0,79	0,4	0,6
Trommel–buffer–tau via T	2,66	2,68	2,65	0,56	3,0	2,62	2,61	2,62	0,57	0,4	2,3

type A forrang: A-sakene disponerer hele kapasiteten, B-sakene det som blir igjen,

$$R_{\mathrm{A}} = \frac{Z_{\mathrm{D}} Q_{\mathrm{A}}}{Q_{\mathrm{A}} + m_{\mathrm{D}} Z_{\mathrm{D}}}, \qquad R_{\mathrm{B}} = \frac{Z_{\mathrm{B}} Q_{\mathrm{B}}}{Q_{\mathrm{B}} + m_{\mathrm{D}} Z_{\mathrm{B}}}, \quad Z_{\mathrm{B}} \coloneqq Z_{\mathrm{D}} - R_{\mathrm{A}}, \qquad \tau_k^{\mathrm{kø}} = \frac{Q_k}{R_k}, \tag{33}$$

og tilstrømning, inkapasitering og tillit følger (29) og (30) typevis, med g_k fra hver types egen ventetid og referansetid og π-vektet g i tillitsligningen. Prioriteten tolkes som ikke-preemptiv (en påbegynt hovedforhandling avbrytes ikke); i væskemodellen (33) skiller de to formene seg bare med én behandlingstid m_{D}, uten betydning på månedlig tidsskala. Flere parallelle spor (dommere, rettssaler) legges inn ved å erstatte kapasitetsfunksjonen i (28) med sammenhengen mellom beholdning og utnyttelse for M/M/c-køen med c_{D} spor (Erlang C, Kleinrock, 1975), invertert numerisk; $c_{\mathrm{D}} = 1$ gir (28) eksakt, og m_{D} rekalibreres slik at referansetiden fortsatt er 3 måneder (1,54 måneder for $c_{\mathrm{D}} = 10$). Større variasjon i behandlingstidene kan tas inn ved å skalere ventetidsdelen med $(c_a^2 + c_s^2)/2$ etter Kingmans tungtrafikk-tilnærming (Kingman, 1961); vi holder faktoren lik 1. Med $\pi_{\mathrm{A}} = 0{,}4$, $\vartheta_{\mathrm{A}} = 0{,}8$ og $\vartheta_{\mathrm{B}} = 0{,}3$ (tabell 2) er det π-vektede gjennomsnittet $\vartheta = 0{,}5$ som før. Underordning gjelder nå bare de øvrige sakene: ξ_{B} senkes fra 1 til 0,67, de samme 20 prosentene av det samlede sakstilfanget som i scenario 3. Referansen er totypemodellen uten forrang (felles kø), som gir $K_{\mathrm{obj}} = 1{,}30$ i scenario 2 mot 1,28 i figur 6, fordi typer med ulik ϑ ikke beholder samme vekt gjennom forløpet.

Tabell 5 og figur 7 (høyre) viser resultatet. Med felles kø rammer restansen de alvorlige sakene hardest, fordi en større del av virkningen mot dem forutsetter dom: K_{A} ender 52 prosent over referansen og K_{B} 15 prosent over. Med forrang snus bildet. De alvorlige sakene avgjøres på under en måned gjennom hele forløpet, hele virkningen mot dem realiseres, og K_{A} ender på referansenivået: de alvorlige sakene nyter godt av den opptrappingen de øvrige sakene utløser. Hele kostnaden ved flaskehalsen flyttes til de øvrige sakene, som får 71 måneders saksbehandlingstid og 30 prosent høyere belastning, og dit konsentreres frafallet. Énkømodellen overvurderer dermed det samlede tapet: merbelastningen i forhold til scenario 1 mer enn halveres, fra 0,21 til 0,10, og tillitstapet faller fra 0,13 til 0,08. Den er riktig i at kapasitetstapet koster, men tar feil om hvem som bærer kostnaden. Rangeringen av tiltakene endres ikke i noen variant: begge tiltak er best, deretter utvidelse, underordning og ingen tiltak, mens trommel–buffer–tau gjennomført ved å holde igjen politiinnsatsen fortsatt er klart dårligst ($K_{\mathrm{obj}} \approx 2{,}6$). Underordningen betyr mindre for den samlede belastningen med forrang (fra 1,18 til 1,13, mot fra 1,30 til 1,19 uten), fordi forrangen allerede har vernet det som veier tyngst, men den er fortsatt det eneste virkemidlet som raskt bremser køen for de øvrige sakene. Prioritering etter $A_i S_i^{\theta}$ (tabell 1) er altså i seg selv et virkemiddel som verner det alvorlighetsvektede målet, men det flytter kapasitetsproblemet i stedet for å fjerne det, og ventetiden for de øvrige sakene har sin egen rettslige grense i kravet om avgjørelse innen rimelig tid. Antall spor betyr lite: med $c_{\mathrm{D}} = 10$ endres saksbehandlingstiden i de overbelastede



Figur 7. Robusthet. Venstre: regime C ($k_p = 3$, $\tau = 11$) uten kjede, med politiets kapasitet modellert etter (12′); metning gjør grensesyklusen kraftigere. Høyre: saksbehandlingstid i den koblede modellen (regime A) med to sakstyper; med forrang avgjøres de alvorlige sakene på under en måned gjennom hele forløpet, mens hele restansen rammer de øvrige sakene.

scenarioene med under én måned, fordi $R_{\text{D}} = Z_{\text{D}}$ ved overbelastning uansett køens finstruktur; i scenario 5 er sluttiden 1,7 måneder mot 1,0 fordi den rene behandlingstiden er lengre når referansetiden holdes fast. M/M/1-formen i (28) har derfor betydning for kalibreringen og for forløpet nær full utnyttelse, ikke for hovedresultatene, som drives av overbelastningen. Forrangen endrer derimot den lånte stabiliteten (avsnitt 8.2) kvantitativt: R_{A} følger fortsatt J_{A}, så bare B-delen av sløyfen kobles ut, og i regime C svinger kriminaliteten med amplitude 0,21 ved mettet domstol, mot under 0,1 med felles kø og 0,49 uten flaskehals. En prioriterende domstol maskerer en ustabil sløyfe i mindre grad enn énkømodellen tilsier.

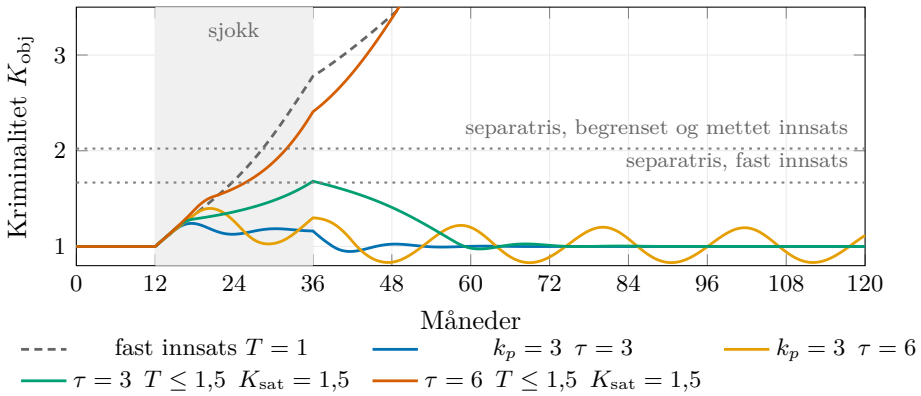
8.4 Selvforsterkende dynamikk og vippepunkter

Modellen har til nå ingen mekanisme der kriminalitet avler kriminalitet: tilstrømningen Ω er uavhengig av nivået. Rekruttering til kriminelle miljøer, sosial smitte og ulovlige markeder der omsetningen gir grunnlag for ny omsetning, gir en tilstrømning som vokser med nivået, og slike mekanismer kan gi flere likevekter og vippepunkter som en lineær analyse rundt én likevekt ikke fanger. Vi legger derfor til et rekrutteringsledd med metning,

$$\Omega^{\text{eff}}(t) = \Omega(\mathbf{E}(t)) + \epsilon\,\mathcal{H}(K_{\text{obj}}), \qquad \mathcal{H}(K) := \frac{K^{\varrho}}{K_{\text{h}}^{\varrho} + K^{\varrho}}, \tag{34}$$

der ϵ er den største rekrutteringen per måned, K_{h} nivået der halvparten er nådd, og $\varrho > 1$ gir en terskel: rekrutteringen er ubetydelig ved lave nivåer og tiltar raskt når miljøene når en kritisk størrelse. Med $\varrho = 1$ er $\mathcal{H}$ konkav og likevekten entydig; det er terskelen som skaper bistabilitet. Ved fast innsats løser likevektene $\Omega + \epsilon\mathcal{H}(K) = (c\bar{T} + \delta)K$, det vil si skjæringene mellom en S-kurve og en rett linje. En nødvendig betingelse for tre skjæringer er at rekrutteringens største stigning overstiger dempningen, $\epsilon \max_K \mathcal{H}'(K) > c\bar{T} + \delta$, og vippepunktet (den mellomste, ustabile likevekten) smelter sammen med en av de stabile der linjen tangerer kurven, en sadelnode-bifurkasjon (Scheffer mfl., 2009). Lokalt rundt den lave likevekten gjelder proposisjon 1 fortsatt, men med dempningen $a - \epsilon\mathcal{H}'(\bar{K})$: rekrutteringen spiser av samfunnets egen dempning, og τ_c faller. Nær vippepunktet går den effektive dempningen mot null, som gir den kritiske nedbremsingen (*critical slowing down*) som er kjent som tidlig varsel om kritiske overganger: langsommere gjenoppretting etter forstyrrelser og økende autokorrelasjon i tidsserien (Scheffer mfl., 2009).

Med $\epsilon = 0{,}30$, $K_{\text{h}} = 2$ og $\varrho = 4$ (tabell 2) og Ω rekalibrert slik at $K_{\text{obj}} = 1$ forblir likevekt ved $T = 1$, har den åpne sløyfen tre likevekter, 1, 1,67 og 2,88; den lave er stabil



Figur 8. Vippepunkt med rekrutteringsleddet (34). Et forbigående sjokk (+60 prosent i det strukturelle pådraget, skravert) presser kriminaliteten mot separatrisen. Uten respons, og med begrenset og mettet innsats når responsen kommer for sent ($\tau = 6$), passeres separatrisen og overgangen til den høye likevekten (utenfor aksen) blir varig. Med tidlig respons vender kriminaliteten tilbake.

med dempning 0,054 i stedet for 0,12, slik at τ_c faller fra 7,2 til 5,9 måneder, og separatrisen ligger 67 prosent over referansen. Lukkes sløyfen med proporsjonalregelen, forsvinner den høye likevekten: fordi fjerningen i (12) er proporsjonal med K_{obj}, er det nok at innsatsen kan økes varig med 10 prosent for at linjen skal ligge over S-kurven overalt. Det gjelder ikke når innsatsen både er begrenset og politiet mettet. Med et innsatstak $T \leq 1{,}5$ (en budsjettskranke) og $K_{\text{sat}} = 1{,}5$ er fjerningen begrenset oppad av Z_{E}, og de tre likevektene vender tilbake: 1, 2,02 og 5,2. Figur 8 viser et forbigående sjokk, en økning i det strukturelle pådraget på 60 prosent i månedene 12–36. Uten respons passerer kriminaliteten separatrisen i løpet av sjokket og går over til den høye likevekten, forsterket av at tryggheten og dermed tilliten faller; den vender ikke tilbake når sjokket er over. Med proporsjonalregelen og $\tau = 3$ stopper utslaget på 1,24 før kriminaliteten vender tilbake; med $\tau = 6$ stopper det på 1,40, hvoretter kriminaliteten pendler i en grensesyklus mellom 0,83 og 1,20 rundt referansen, siden τ_c med rekruttering er 5,9 måneder. I ingen av tilfellene passeres separatrisen. Med begrenset og mettet innsats skiller forsinkelsen mellom to kvalitativt ulike utfall: ved $\tau = 3$ når kriminaliteten 1,68, under separatrisen på 2,02, og faller tilbake til referansen; ved $\tau = 6$ har den passert 2,4 når sjokket opphører, og overgangen blir varig. Tilbakeveien er dessuten asymmetrisk: fra den høye likevekten kreves en innsats som fjerner denne likevekten, langt mer enn det som ville ha hindret overgangen.

Tre konsekvenser følger for konseptet. For det første er *tidlig og tilstrekkelig* respons avgjørende der selvforsterkende mekanismer er plausible; den avdempede responsen i korollar 1 må da kombineres med et handlingsrom og en kapasitet som gjør responsen tilstrekkelig før separatrisen nås, og budsjettskranken i (23) får en skyggepris som stiger bratt nær vippepunktet. For det andre må den lokale stabilitetsanalysen suppleres med en global: den relevante marginen er ikke bare $\tau_c - \tau$, men avstanden til separatrisen, som krymper når politiet er mettet. For det tredje er dette situasjonen der modellprediktiv styring med riktig intern modell har størst verdi, fordi den kan forutsi at et forbigående sjokk vil krysse separatrisen før det skjer. Om og hvor slike terskler finnes i norske lovbruddskategorier, er et empirisk spørsmål som artikkelen ikke tar stilling til.

8.5 Endogene tiltak: fra observasjon til beslutning

I avsnitt 8.2 ble underordning og utvidelse lagt inn på faste tidspunkter og i full størrelse, dimensjonert med kjennskap til den nye likevekten; bare politiets innsats var en respons på observasjoner. Vi lar nå også ξ_{D}, Z_{D} og q bestemmes av beslutningsregler som svarer på

forsinkede observasjoner, og sammenligner tre måter å organisere responsen på. Politiet følger regime A, sjokket er som før, og horisonten forlenges til 180 måneder fordi kapasitetssløyfen er langsom. Kjedestatistikken ($\hat{\tau}_{\mathrm{D}}^{\mathrm{kø}}$, $\hat{J}_{\mathrm{D}}$, $\hat{v}_{\mathrm{D}}$, $\hat{Q}_{\mathrm{D}}$) observeres med forsinkelse $\tau_{\mathrm{D}}^{\mathrm{obs}} = 3$ måneder; tilliten måles årlig og publiseres tre måneder etter årsskiftet (tabell 2).

Desentraliserte regler. Hver aktør svarer på sitt eget signal uten å kjenne de andres planer. *Påtalemyndigheten* leder flere saker til forenklede reaksjoner når observert saksbehandlingstid overskrider terskelen $\tau_{\mathrm{D}}^{\mathrm{gr}} = 6$ måneder, og fører dem tilbake når tiden er under referansen; praksis endres med tidskonstant τ_ξ, og andelen til domstol kan ikke gå under $\xi_{\min} = 0{,}8$, fordi de alvorlige sakene skal for retten:

$$\Delta\xi^{\mathrm{mål}} = -k_\xi\Big[\tfrac{\hat{\tau}_{\mathrm{D}}^{\mathrm{kø}}}{\tau_{\mathrm{D}}^{\mathrm{gr}}} - 1\Big]_+ + k_\xi\Big[1 - \tfrac{\hat{\tau}_{\mathrm{D}}^{\mathrm{kø}}}{\tau_{\mathrm{D}}^{\mathrm{ref}}}\Big]_+ \text{ per måned, i } [\xi_{\min}, 1], \qquad \dot{\xi}_{\mathrm{D}} = \frac{\xi^{\mathrm{mål}} - \xi_{\mathrm{D}}}{\tau_\xi}. \tag{35}$$

Budsjettet vedtas årlig. Overstiger gjennomsnittlig observert innstrømning siste tolv måneder, $\bar{J}_{\mathrm{D}}$, den kapasiteten som er realisert eller bevilget, Z^{bev}, bestilles kapasitet for utnyttelsen $v^* = 0{,}95$; er utnyttelsen under 0,85, bygges kapasiteten langsomt ned ved naturlig avgang:

$$\Delta Z^{\mathrm{best}} = k_Z\Big(\frac{\bar{J}_{\mathrm{D}}}{v^*} - Z^{\mathrm{bev}}\Big) \text{ hvis } \bar{J}_{\mathrm{D}} > Z^{\mathrm{bev}}, \qquad \dot{Z}_{\mathrm{D}} = -\frac{Z_{\mathrm{D}} - \bar{J}_{\mathrm{D}}/v^*}{\tau_{\mathrm{ned}}} \text{ hvis } \bar{J}_{\mathrm{D}} < 0{,}85\, Z_{\mathrm{D}}. \tag{36}$$

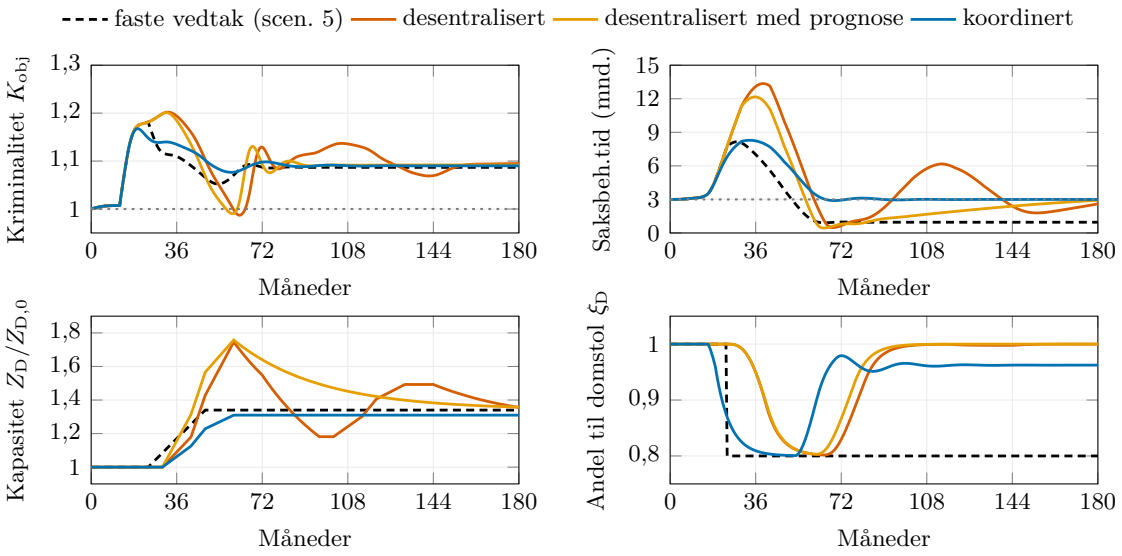
Bestilt kapasitet realiseres lineært 6–24 måneder etter vedtaket, og budsjettet regner med det som allerede er bevilget: regelen reagerer bare på det den ser, men den er ikke naiv.

Koordinert regel etter metoden. Trinn (c)–(e) i avsnitt 7.7 gir én felles kapasitetsbalanse: innstrømningen pluss avviklingen av restansen over en horisont H_Q skal holdes innenfor kapasiteten med buffer, $\xi_{\mathrm{D}}\hat{F} + [\hat{Q}_{\mathrm{D}} - Q^{\mathrm{ref}}]_+/H_Q \le v^* Z_{\mathrm{D}}$, der $\hat{F} = \hat{J}_{\mathrm{D}}/\xi_{\mathrm{D}}$ er observert fjerningsstrøm og Q^{ref} restansen ved referansetiden. Når domstolen er identifisert som bindende ($\hat{v}_{\mathrm{D}} > 1$), underordnes det først, og budsjettet bestiller bare det underordningen ikke tar unna:

$$\begin{aligned} \xi^{\mathrm{mål}} &= \Pi_{[\xi_{\min},1]}\left[\frac{v^* Z_{\mathrm{D}} - [\hat{Q}_{\mathrm{D}} - Q^{\mathrm{ref}}]_+/H_Q}{\hat{F}}\right] && \text{(månedlig, påtalemyndigheten)},\\ \Delta Z^{\mathrm{best}} &= k_Z\Big(\frac{\xi_{\min}\hat{F}^{\mathrm{p}} + [\hat{Q}_{\mathrm{D}} - Q^{\mathrm{ref}}]_+/H_Q}{v^*} - Z^{\mathrm{bev}}\Big)_+ && \text{(årlig, budsjettet)}. \end{aligned} \tag{37}$$

Den første regelen er trommel–buffer–tau på påtalebeslutningene (tabell 1). I den andre er $\hat{F}^{\mathrm{p}}$ en *prognose* for fjerningsstrømmen i den gjenopprettede likevekten, beregnet ved å kjøre den interne modellen uten flaskehals $H_{\mathrm{p}} = 36$ måneder fram fra det observerte tilstandsestimatet med samme styringsregel for innsatsen (trinn e). Kapasitet bygges ned bare når den er overflødig selv om alle saker går til domstol ($\hat{F}^{\mathrm{p}} < 0{,}85\, Z_{\mathrm{D}}$). Som mellomform bruker vi en *desentralisert regel med prognose*, der budsjettet dimensjonerer etter $\hat{F}^{\mathrm{p}}$ og restansen, men regner med $\xi_{\mathrm{D}} = 1$ fordi det ikke kjenner påtalemyndighetens plan. *Kvalitetssløyfen* lar q øke med $k_q(\Lambda^* - \hat{\Lambda})$ per måned, begrenset til $[0, q_{\max}]$, når siste publiserte årsmåling $\hat{\Lambda}$ ligger under målet $\Lambda^* = 0{,}78$, med $\mu_3 = 0{,}5$ i (21). Med faste tidsplaner for ξ_{D} og Z_{D} gjengir den utvidede koden scenario 2 og 5 i figur 6 uten avvik.

Resultater. Figur 9 og tabell 6 besvarer tre spørsmål. *Oppstår rekkefølgen underordne først, utvid deretter av seg selv?* Ikke med desentraliserte regler. Budsjettet ser overbelastningen i utnyttelsesgraden straks og bestiller 27 prosent mer kapasitet i måned 24, mens påtalemyndighetens signal, saksbehandlingstiden, passerer terskelen så sent at andelen først faller fra måned 31 og når gulvet i måned 54. Kapasiteten kommer før underordningen, og begge tiltak



Figur 9. Endogene tiltak i den koblede modellen (regime A, samme sjokk som før, 180 måneder). Faste vedtak er scenario 5 i figur 6. Med desentraliserte regler bestiller budsjettet etter den innstrømningen flaskehalsen selv har blåst opp, og kapasiteten bygges ut til 1,75 ganger utgangsnivået før den bygges ned og opp igjen; med den koordinerte regelen underordnes det først, og kapasiteten dimensjoneres etter prognosen for den gjenopprettede likevekten og restansen.

settes inn i full størrelse mot samme problem. I den koordinerte regelen følger rekkefølgen av to forhold: påtalemyndighetens sløyfe er den raskeste og reagerer så snart overbelastningen er observert (andelen faller fra måned 17 og når gulvet i måned 36), og budsjettet dimensjonerer bare det som gjenstår etter underordningen (19 prosent i måned 24 og 12 prosent i måned 36). Rekkefølgen skyldes tidsskalaene og at dimensjoneringen regner med den andre sløyfens tiltak, ikke ToC-trinnene i seg selv.

Får kapasitetssløyfen egne svingninger og overutbygging? Ja, og årsaken er først og fremst signalet, ikke forsinkelsen. Mens domstolen er flaskehals, er innstrømningen oppblåst av flaskehalsen selv: kriminaliteten ligger høyt, styringsregelen trapper opp innsatsen, og flere saker avdekkes ($\hat{J}_D = 0{,}17$ i måned 33, mot 0,13 i den gjenopprettede likevekten). Et budsjett som dimensjonerer etter denne strømmen, bestiller 27 og 48 prosent i to påfølgende år, og kapasiteten når 1,75 ganger utgangsnivået i måned 60, mot de 1,34 likevekten krever, det vil si om lag 30 prosent over behovet. Når kapasiteten er på plass, faller kriminalitet, innsats og innstrømning, utnyttelsen synker til 0,44, kapasiteten bygges ned, sakene føres tilbake til domstolen, og i måned 96 er domstolen igjen overbelastet. Resultatet er en dempet svingning i kapasiteten med en periode på seks til sju år ($Z_D/Z_{D,0}$ fra 1,75 ned til 1,18 og opp til 1,49), med følgebølger i reaksjonsform og kriminalitet (topper i K_{obj} på 1,20, 1,13 og 1,14), det mønsteret som i eksperimenter oppstår når beslutningstakere overser forsyningslinjen og tilbakekoblingen fra egne tiltak (Sterman, 1989). Dempet forsterkning ($k_Z = 0{,}5$) senker toppen til 1,50, men svingningen består og utfallet blir dårligere. Prognose uten koordinering hjelper heller ikke: et budsjett som regner med at alle saker går til domstol mens påtalemyndigheten leder hver femte sak utenom, bestiller 47 og 29 prosent og ender på samme overkapasitet. Den koordinerte regelen bestiller 19 og 12 prosent, kapasiteten stabiliserer seg på 1,31, og andelen til domstol går tilbake til 0,96 når kapasiteten er på plass.

Gir koordinering bedre og mer stabile utfall? Målt mot de faste vedtakene, som er dimensjonert med kjennskap til likevekten, gir den koordinerte regelen nesten samme utfall for kriminalitet, tillit og saksbehandlingstid med 14 prosent færre kapasitetsår, fordi reaksjonsformene bærer en del av tilpasningen. De desentraliserte reglene gir høyere kriminalitet, lavere

Tabell 6. Endogene tiltak, regime A, månedene 0–180: gjennomsnittlig kriminalitet fra måned 12, laveste tillit, lengste saksbehandlingstid, høyeste og endelige kapasitet relativt til utgangsnivået, andel saker til domstol ved slutten, kapasitetsår over utgangsnivået og kumulativt frafall fra køen målt i måneder av referanseinnstrømningen.

Regel	$\bar{K}_{\text{obj}}$	$\Lambda^{\min}$	$\tau_{\text{D}}^{\max}$	$Z^{\max}$	Z_{180}	ξ_{180}	kap.år	frafall
2: flaskehals, ingen tiltak	1,259	0,651	69	1,00	1,00	1,00	0	73
5: faste vedtak (måned 24)	1,092	0,749	8,2	1,34	1,34	0,80	4,1	4,4
Desentralisert	1,106	0,719	13,4	1,75	1,36	1,00	4,9	10,3
Desentralisert, $k_Z = 0{,}5$	1,114	0,714	14,8	1,50	1,25	0,97	3,8	12,4
Desentralisert med prognose	1,099	0,724	12,2	1,76	1,36	1,00	5,7	7,8
Koordinert, $H_Q = 24$	1,098	0,743	8,3	1,31	1,31	0,96	3,5	8,3
Koordinert, $H_Q = 48$	1,102	0,740	8,5	1,19	1,19	0,88	2,2	8,6
Koordinert, $H_Q \to \infty$	1,128	0,735	8,8	1,07	1,07	0,80	0,9	13,5
Koordinert med kvalitetssløyfe	1,095	0,743	8,3	1,31	1,31	0,96	3,5	8,3

tillit i bunnen, lengre kø, mer frafall og 38 prosent flere kapasitetsår enn den koordinerte regelen. Prisen for koordineringen er at en del av sakene varig avgjøres med forenklede reaksjoner. Hvor stor den delen skal være, er et normativt valg som i modellen styres av horisonten H_Q: en kort horisont ($H_Q = 12$) gir mer permanent kapasitet, en lang ($H_Q = 48$) mindre kapasitet og lavere andel til domstol, og uten restanseavvikling ($H_Q \to \infty$) økes kapasiteten med de 7 prosentene regneeksempelet i avsnitt 7.6 angir, men restansen avvikles ikke: saksbehandlingstiden blir stående på 6 måneder og tilliten på 0,75. Kvalitetssløyfen endrer lite: fordi tilliten måles årlig og publiseres med forsinkelse, er bunnpunktet upåvirket, mens nivået i gjenopprettingsfasen løftes (0,80 mot 0,79 i måned 120). Kvalitet er et virkemiddel for nivået på tilliten, ikke for å fange opp et fall som allerede har skjedd.

Merknad 4 (Kapasitetssløyfens stabilitet). Kapasitetssløyfen alene har samme form som politisløyfen. Med gitt innstrømning J og en regel som hvert budsjettår av lengde Δ korrigerer andelen k_Z av gapet mellom nødvendig kapasitet J/υ^* og *realisert* kapasitet, realisert τ_Z etter vedtaket, følger avviket $z = Z_{\text{D}} - J/\upsilon^*$ ligningen $\dot{z} = -b\,z(t - \tau_Z)$ med $b = k_Z/\Delta$ og $a = 0$ (tillegg B); proposisjon 1 gir stabilitet hvis og bare hvis $k_Z\tau_Z/\Delta < \pi/2$, for $\Delta = 12$ og $\tau_Z = 21$ måneder $k_Z < 0{,}9$. To forhold gjør sløyfen mindre utsatt enn grensen tilsier. Regner budsjettet med kapasitet som er bevilget, men ennå ikke realisert (forsyningslinjen i systemdynamikkens terminologi, Sterman, 2000), faller forsinkelsen ut av sløyfen, og regelen treffer behovet uten svingninger for alle $k_Z \leq 1$; realiseres kapasiteten gradvis, dempes svingningene også når bevilget kapasitet ikke regnes med, men langsomt (høyst 21 prosent overskytende kapasitet for $k_Z = 1$ og realisering over 6–36 måneder, med merkbare svingninger i flere tiår). Overutbyggingen i tabell 6 skyldes derfor ikke sløyfens forsterkning, men signalet: innstrømningen budsjettet styrer etter, er endogen til den flaskehalsen det skal fjerne. Derfor krever trinn (e) i avsnitt 7.7 en prognose for den gjenopprettede likevekten, ikke en fremskrivning av dagens strøm.

9 Slik brukes modellen i praksis

Prinsippet i definisjon 1 — at innsatsen løpende skal stå i forhold til hvor mye, og hvor alvorlig, lovene faktisk brytes — kan ikke gjennomføres uten et apparat som måler bruddene, beregner en avdempet respons og viser hva responsen betyr for resten av kjeden. Modellen er en skisse til dette apparatet. Avsnittet beskriver hvordan den er tenkt brukt når den er kalibrert og validert, ikke et ferdig verktøy (avsnitt 10). Den er ment som *beslutningsstøtte*, ikke som autopilot: den skal gjøre konsekvensene av valg synlige, tallfeste avveininger og avdekke dynamiske feller, mens avgjørelsene fortsatt tas av ansvarlige politikere og ledere. Naturlige brukere er Politidirektoratet og politidistriktene (prioritering og ressursfordeling),

Tabell 7. Bruksområder, spørsmål modellen besvarer, og hva den leverer.

Bruksområde	Spørsmål	Leveranse
1. Diagnose	Er dagens styringspraksis stabil, eller skaper den pendelsvingninger? Hvor i straffesakskjeden ligger flaskehalsen?	Estimert implisitt reaktivitet $\hat{k}_p$, forsinkelse $\hat{\tau}$, stabilitetsmargin $\tau_c - \tau$ (proposisjon 1); utnyttelsesgrad υ_k og køtid $\tau_k^{\text{kø}}$ per ledd (avsnitt 7.2).
2. Scenarioanalyse	Hva skjer med kriminalitet, trygghet, tillit og restanser ved et konjunkturfall, et budsjettkutt, en opptrapping av politiinnsatsen eller en endring i domstolskapasiteten?	Prognosebaner med usikkerhetsbånd for N_i, U, Λ, Q_{D} og $\tau_{\text{D}}^{\text{kø}}$ under alternative forløp av $\mathbf{E}$, $\mathcal{B}$, ξ_{D} og Z_{D} (avsnitt 8.2).
3. Beslutningsstøtte	Hvordan bør innsatsen fordeles mellom kategorier og over tid innenfor budsjett og rettsstatlige grenser, og hvilken type tiltak svarer til den skranken som binder?	Anbefalt tiltaksplan fra metoden (algoritme 1) med MPC for innsatsen, skyggepriser, varsel om forskyvning og tillitsrisiko.
4. Evaluering	Virket tiltaket, eller skyldtes endringen makroforhold eller registreringspraksis?	Dekomponering av observert endring i makrobidrag, tiltaksbidrag og målebidrag (ligningene (4) og (6)).

Justis- og beredskapsdepartementet (budsjett og styringsdialog) og Stortinget (vurdering av satsinger og kutt). Dette avsnittet beskriver hva modellen kan brukes til, hvordan den settes i drift, og hvordan metoden i avsnitt 7.7 blir en arbeidsform for aktørene i kjeden.

9.1 Fire bruksområder

Modellen kan brukes på fire måter, med økende krav til data og kalibrering. Tabell 7 oppsummerer hvilke spørsmål hver av dem besvarer, og hva beslutningstakeren får.

9.2 Trinn for trinn

Trinn 1. Avgrens styringsnivået. Velg nivå (nasjonalt eller politidistrikt), tidsenhet (måned er et godt utgangspunkt) og et lite antall lovbruddskategorier, typisk $n = 4$–8, for eksempel vold, vinning, seksuallovbrudd, narkotika, økonomisk kriminalitet og trafikk. Skill mellom offerbaserte og kontrollavhengige kategorier, fordi de har ulik måleligning (4).

Trinn 2. Fastsett de normative vektene. Alvorlighetsvektene A_i kan utledes av straffe-rammene, for eksempel som normalisert maksimumsstraff. Skadekostnadene S_i kan bygge på samfunnsøkonomiske kostnadsanslag per lovbrudd. Start med $\theta = 1$ og vis følsomheten for $\theta = 1{,}5$. Vektene A_i, S_i, θ og ω er politiske valg og skal dokumenteres åpent.

Trinn 3. Bygg datagrunnlaget. Tabell 8 viser aktuelle kilder. Datasett med ulik frekvens (måned, kvartal, år) kobles i en felles tilstandsrommodell. Ta med innstrømning, restanse og saksbehandlingstid for hvert ledd i straffesakskjeden, slik at υ_k og $\tau_k^{\text{kø}}$ i (26) kan beregnes.

Trinn 4. Estimer tilstanden. Rå registerdata er ikke et mål på N_i. Bruk et Kalman-filter eller en bayesiansk tilstandsrommodell (Kalman, 1960) som kombinerer registrerte lovbrudd, offerundersøkelser (som anslår registreringsgraden r_i) og innbyggerundersøkelser. Resultatet er løpende estimater $\hat{N}_i(t)$, $\hat{U}(t)$, $\hat{\Lambda}(t)$ med usikkerhet. Gjør identifikasjonsantakelsene fra avsnitt 2.4 eksplisitte: modeller $\log r_i$ som en langsom

tilfeldig gang eller strukturelt som $r_i(\hat{\Lambda}, T_i)$, bruk offerundersøkelsene som forsinkede ankerpunkter, og ta med uavhengige indikatorer der de finnes. Rapporter usikkerheten i $\hat{N}_i$ sammen med estimatet, og la filterets tidskonstant inngå i τ_{info} når stabilitetsmarginen beregnes i trinn 7: jo svakere observerbarheten er, desto mer må registerstøyen glattes, og desto lavere reaktivitet tåler sløyfen.

Trinn 5. Kalibrer parameterne. Den største metodiske utfordringen er at politiinnsatsen selv svarer på kriminaliteten. En naiv regresjon av kriminalitet på innsats gir derfor skjeve anslag på γ_i (Levitt, 1997); også instrumentbaserte anslag er sårbare, jf. McCrary (2002). Effektparametere bør hentes fra kvasieksperimenter og randomiserte forsøk, for eksempel med *hot spots*-patruljering (Sherman og Weisburd, 1995; Braga mfl., 2019). Forskyvningsmatrisen Σ kan anslås fra studier med kontrollområder (Guerette og Bowers, 2009). Makrokoeffisientene β_k estimeres i paneldata over distrikter med etterslep, og tillitsparameterne (κ_Λ, μ_1–μ_3) fra gjentatte undersøkelser. Bayesiansk estimering med forhåndsfordelinger fra litteraturen gir et realistisk bilde av usikkerheten.

Trinn 6. Valider. Estimer modellen på en tidlig periode og test prognosene på en senere periode som ikke er brukt i estimeringen. Sjekk om modellen gjenskaper historiske sykluser i innsats og kriminalitet. Gjennomfør følsomhetsanalyser for de normative vektene og de mest usikre parameterne.

Trinn 7. Ta modellen i bruk for diagnose, scenarioer, beslutningsstøtte og evaluering (tabell 7). En enkel stabilitetsdiagnose (avsnitt 9.3) og en flaskehalsdiagnose (avsnitt 9.5) kan gjøres allerede etter trinn 3. Den fulle beslutningssyklusen (algoritme 1) krever at trinnene 4–6 er gjennomført.

9.3 Regneeksempel: stabilitetsdiagnose

Eksempel: Er budsjettsyklusen en kilde til svingninger?

Tallene er tenkte og viser framgangsmåten; de er ikke estimater for Norge, og reaktiviteten i trinn 2 er trolig langt høyere enn i noen faktisk budsjettprosess (avsnitt 8.1). Kriminalitet og innsats måles i andeler av referansenivået, slik at $\bar{K} = \bar{T} = 1$ (jf. tabell 2).

1. Mål forsinkelsen. Anta at statistikken foreligger etter 3 måneder, at omprioritering skjer i den årlige budsjettprosessen (i gjennomsnitt 6 måneder etter at tallene foreligger), og at iverksetting tar 2 måneder: $\tau = 3 + 6 + 2 = 11$ måneder.

2. Estimer den implisitte reaktiviteten. Anta at en regresjon av innsatsendringer på tidligere avvik fra målnivået viser at et avvik på 10 prosent over målet har gitt om lag 30 prosent mer innsats, altså $\hat{k}_p \approx 0{,}30/0{,}10 = 3$.

3. Beregn a og b. Med fjerningsrate $c\bar{T} = \gamma\bar{\Lambda}^\eta\bar{T} = 0{,}10$ per måned og naturlig avgang $\delta = 0{,}02$ gir (13) $a = 0{,}10 + 0{,}02 = 0{,}12$ og $b = c\bar{K}\hat{k}_p = 0{,}10 \cdot 3 = 0{,}30$.

4. Sammenlign med stabilitetsgrensen. Siden $b > a$, gir (14)

$$\tau_c = \frac{\arccos(-0{,}4)}{\sqrt{0{,}30^2 - 0{,}12^2}} = \frac{1{,}98}{0{,}275} \approx 7{,}2 \text{ måneder } < \tau = 11.$$

Diagnose: med disse forutsetningene er sløyfen ustabil. Ved stabilitetsgrensen er svingeperioden $2\pi/\sqrt{b^2 - a^2} \approx 23$ måneder, altså om lag to år, det vil si et mønster der satsinger og kutt ville avløse hverandre over budsjettårene.

5. Vurder tiltak. (a) *Dempe reaksjonen:* Ved $\tau = 11$ er største stabile reaktivitet $k_p \approx 2{,}3$; den konservative regelen (15) gir $k_p < 1{,}4$. I praksis: la et avvik på 10 prosent utløse høyst 14–23 prosent endring i innsatsen, ikke 30. (b) *Korte ned forsinkelsen:* Månedlig statistikk ($\tau_{\text{info}} = 1$) og fullmakt til omprioritering i løpet av året ($\tau_{\text{pol}} = 2$) gir $\tau = 1 + 2 + 2 = 5 < 7{,}2$, slik at dagens $k_p = 3$ blir stabil. (c) *Styre modellbasert:* MPC predikerer virkningen av vedtatte tiltak og fjerner hovedårsaken til overreaksjon (figur 4, regime D).

Tabell 8. Datagrunnlag og estimeringsmetode for modellens størrelser (norske forhold).

Størrelse	Kilde	Frekvens	Bruk / metode
N_i^{obs}	Politiets straffesaksregister; SSBs kriminalstatistikk	måned	observasjoner i (4)
r_i	Nasjonal trygghetsundersøkelse (offerundersøkelse) (Løvgren mfl., 2023)	årlig	anslag på mørketall per kategori
$U,\ \Lambda$	Politiets innbyggerundersøkelse (Politidirektoratet, 2024)	årlig, distrikt	målinger i tilstandsrommodellen
$u,\ \varphi$	SSB/NAV (ledighet), SSB (KPI)	måned	**E** i (6)
G	SSBs inntektsstatistikk	årlig	**E** i (6)
Y_{gap}	Norges Banks anslag på produksjonsgapet	kvartal	**E** i (6)
$T_i,\ w_i,\ \mathcal{B}$	Ressursregnskap, tidsregistrering, statsbudsjettet	måned/år	innsats, enhetskostnader, budsjettskranke
$\tau_{\text{info}}, \tau_{\text{pol}}, \tau_{\text{oper}}$	Tidsstempler i saksbehandling; budsjettkalender	løpende	forsinkelsen τ i (10)
$J_k,\ Z_k,\ Q_k$	Politiets og påtalemyndighetens saksbehandlingsstatistikk; Domstoladministrasjonens statistikk over restanser og saksbehandlingstid per instans; kriminalomsorgens soningskø	måned, kvartal	kapasitetsskranker i (32); flaskehalsdiagnose (tabell 1)
$T_i^{\text{lov}},\ \mathcal{D}$	Gjeldende lov og forskrift; forsøkshjemler med utløpsdato	ved lovendring	lovbestemte grenser og domene i (25)
$\gamma_i,\ \Sigma$	Kvasieksperimenter, randomiserte forsøk, meta-analyser	—	forhåndsfordelinger

9.4 Beslutningssyklusen: metoden som algoritme

Algoritme 1 gjør metoden i avsnitt 7.7 operativ som en fast beslutningssyklus. Syklusen kjøres i hvert lag av figur 3 i lagets takt: månedlig eller kvartalsvis for innsats og reaksjonsform, årlig for kapasitet og kvalitet, og ved lovrevisjoner for handlingsrommet. Trinnene er de samme i alle lag; det som skiller dem, er hvilke komponenter av (7) som kan endres, og med hvilken ledetid. MPC (24) er trinn (e) for innsatsen; for reaksjonsform og kapasitet kan reglene i avsnitt 8.5 brukes, så lenge de bygger på samme skrankebilde og prognose.

9.5 Diagnose av straffesakskjeden og lov- og prosessendringer

Flaskehalsanalysen i avsnittene 7.2 og 7.4 kan gjennomføres uavhengig av resten av modellen, fordi den bare krever data som allerede registreres: innstrømning, restanse og saksbehandlingstid for hvert ledd. Utnyttelsesgraden υ_k og køtiden $\tau_k^{\text{kø}}$ fra (26) beregnes per ledd og per sakstype, og leddet med høyest utnyttelse og raskest voksende restanse identifiseres som skranken. Når kjedeskrankene er lagt inn i optimeringen (32), gir skyggeprisene det samme svaret i verdienheter. Tiltakene som følger, hører hjemme hos ulike aktører, og det er et poeng i seg selv at modellen gjør arbeidsdelingen synlig:

- *Politiet og påtalemyndigheten* styrer takten inn i flaskehalsen (underordning): prioritering etter alvorlighet, ferdig forberedte saker og bruk av forenklede reaksjoner der loven åpner for det.
- *Domstolene og Domstoladministrasjonen* utnytter og utvider kapasiteten: beramming, færre utsettelser, digital saksavvikling og dimensjonering av dommerressurser.
- *Justis- og beredskapsdepartementet og Stortinget* forvalter regelskrankene: prosessuelle lov-

Algoritme 1. Beslutningssyklus etter metoden (avsnitt 7.7) for ett lag i figur 3

Inndata: kalibrert modell f med estimerte parametere; skranker (23) og (32); horisont H; vekter i målfunksjonen (22); makroscenarier; lagets tiltakskomponenter i $\boldsymbol{v}$ og deres ledetider

Utdata: anbefalte tiltak, prognoser og beslutningsgrunnlag for hver periode

for hver *beslutningsperiode t i lagets takt* **gjør**

- **(a) Observer og estimer:** hent $\mathcal{O}(t)$ (registrerte lovbrudd, undersøkelser, restanse og saksbehandlingstid per ledd, utnyttelse, ressursbruk, makroindikatorer); oppdater $\hat{\boldsymbol{x}}(t)$ med Kalman-filter, korrigert for $r_i(\Lambda, T_i)$, med usikkerhet;
- **(b) Sammenlign med målene:** alvorlighetsvektede avvik per kategori, trygghet, tillit og saksbehandlingstid mot K^*, U^*, Λ_{krit}, τ_k^{ref};
- **(c) Identifiser skranken:** beregn υ_k, restansevekst og skyggepriser på kapasitets-, budsjett-, tillits- og lovskranker; pek ut den som binder;
- **(d) Velg type tiltak** etter skranken: innsats og organisering, utnytte, underordne (ξ), utvide (Z), kvalitet (q), eller signal til lovgiver (T^{lov});
- **(e) Dimensjoner:** prediker tilstanden fram til nye tiltak kan virke, gitt allerede vedtatte tiltak i alle lag;
- **for hver** *makroscenario* $\hat{\mathbf{E}}^{(s)}$ *(basis, nedside, oppside)* **gjør**
  - løs (24) for innsatsen og kapasitetsbalansen (37) for reaksjonsform og kapasitet, med prognosen for den gjenopprettede likevekten, ikke dagens strøm;
  - **hvis** *problemet er uten tillatt løsning* **så**
    - løs COP med slakk på myke skranker; rapporter hvilke mål som må vike, og hvor mye;
- kontroller at reaksjonen ligger innenfor stabilitetsgrensen for laget (proposisjon 1 og korollar 1), og at den er tidlig nok der vippepunkter er plausible;
- **(f) Kontroller rettsstatlige skranker:** T_i^{grl}, T_i^{lov}, likhet for loven, forholdsmessighet, rimelig tid;
- presenter beslutningsgrunnlaget: anbefalte tiltak, prognoser med usikkerhet, stabilitetsmargin, skyggepriser, tillits- og forskyvningsrisiko;
- **Menneskelig beslutning:** ansvarlig myndighet vedtar tiltak; avvik fra anbefalingen begrunnes;
- **(g) Iverksett og evaluer:** iverksett første periode; logg prognosefeil; re-estimer parameterne periodisk; gjenta fra (a) når skranken har flyttet seg;

endringer, forsøkshjemler og solnedgangsklausuler (avsnitt 7.1), og fastsetter de lovbestemte grensene T_i^{lov} innenfor de konstitusjonelle.

Arbeidsdelingen svarer til sløyfene i figur 3. Metoden krever ikke at én instans fatter alle vedtak, men at sløyfene deler ett bilde av hvilken skranke som binder, og at hver dimensjonerer med de andres tiltak og med prognosen for den gjenopprettede likevekten; avsnitt 8.5 viser at det er denne delingen, ikke reglenes forsiktighet, som hindrer overutbygging og svingninger i kapasitet og reaksjonsform. Modellens bidrag er å vise *hvor* i kjeden en endring gir størst virkning, og å skille mellom tiltak som krever penger, tiltak som krever organisering, og tiltak som krever lovendring. Med domstolskøen koblet til dynamikken (avsnitt 7.3) kan tiltakene dessuten vurderes i sammenheng med styringen av politiinnsatsen: hva en planlagt opptrapping betyr for restanse, saksbehandlingstid, frafall og tillit, hvor stor andel av sakene som må ledes til forenklede reaksjoner for at køen skal holdes stabil, og hvor stor og hvor rask en kapasitetsutvidelse må være for at restansen skal avvikles (avsnitt 8.2). Rekkefølgen har betydning: underordning virker raskt, men stabiliserer ikke kjeden alene; utvidelse gjenoppretter balansen, men først når kapasiteten er på plass; og å begrense politiinnsatsen til det domstolen kan ta unna, gir det dårligste utfallet for kriminaliteten.

9.6 Hva beslutningstakeren får

I tillegg til en anbefalt fordeling av innsatsen gir modellen fem typer innsikt som sjelden er tilgjengelige i dag: *stabilitetsmarginen*, hvor nær dagens praksis er grensen for pendelsvingninger, om det er reaksjonsstyrken eller forsinkelsen som bør endres, og der selvforsterkende mekanismer er plausible, avstanden til et mulig vippepunkt (avsnitt 8.4); *skyggeprisen på budsjettet*, $-\partial\mathcal{J}^*/\partial\mathcal{B}$, som direkte innspill til budsjettforhandlinger; *prisen på rettsstaten*, hvor mye høyere kriminalitet eller utrygghet de konstitusjonelle skrankene innebærer i en gitt situasjon, et tall som gjør avveiningen synlig uten å være et argument for å flytte skrankene (avsnitt 6.3); *flaskehals og skyggepris per ledd*, $-\partial\mathcal{J}^*/\partial Z_k$, verdien av det lovbestemte handlingsrommet (avsnitt 7.5) og prognoser for restanse og saksbehandlingstid ved planlagt innsats, som skiller tiltak som krever bevilgning, fra tiltak som krever lovendring; og *risikoen for forskyvning og tillitstap*, hvilke kategorier som vil øke når innsatsen flyttes, og når planlagt innsats nærmer seg grensen for det som oppleves som forholdsmessig, T_{prop}.

9.7 Skranker for forsvarlig bruk

Sjekkliste for forsvarlig bruk

- ☐ **Aggregert nivå:** modellen brukes på kategorier og områder, aldri til å forutsi enkeltpersoners atferd. EUs KI-forordning (Europaparlamentet og Rådet, 2024) forbyr risikovurderinger av enkeltpersoner for å forutsi lovbrudd når de bare bygger på profilering (artikkel 5), og klassifiserer annen bruk i rettshåndhevelse som høyrisiko. Forordningen er per november 2026 ikke gjennomført i norsk rett; personopplysningsloven og politiregisterloven gjelder uansett.
- ☐ **Menneskelig beslutning:** modellen anbefaler; ansvarlige myndigheter beslutter og begrunner eventuelle avvik.
- ☐ **Åpenhet:** vekter (A_i, S_i, θ, ω), kostnadsfunksjoner, parametere og antakelser dokumenteres og kan etterprøves.
- ☐ **Låste rettsstatsskranker:** $T_i^{\max}$ og menneskerettighetene er absolutte i optimeringen og kan ikke byttes mot måloppnåelse.
- ☐ **Kontroll av tilbakekobling:** for kontrollavhengige lovbrudd vurderes jevnlig om økt registrering skyldes økt innsats og ikke økt kriminalitet (Ensign mfl., 2018).
- ☐ **Flere indikatorer:** registrerte tall skal ikke bli mål i seg selv (Campbell, 1979). Offer- og innbyggerundersøkelser brukes som korrektiv.
- ☐ **Usikkerhet og robusthet:** anbefalinger skal tåle rimelige feil i parameterne. Når usikkerheten er stor, bør man velge den robuste løsningen framfor det beregnede optimum.
- ☐ **Personvern:** bare aggregerte og anonymiserte data brukes.

9.8 Kom i gang

Fullskala MPC er ikke nødvendig for å få nytte av modellen. Et realistisk første prosjekt er en *stabilitetsdiagnose* for ett politidistrikt med tre til fire kategorier; den krever bare historiske serier for innsats og registrert kriminalitet, et anslag på forsinkelsen og et anslag på effektparameteren fra litteraturen, og simuleringskoden bak figurene 4–6 og 9 kan brukes som utgangspunkt. Neste steg er scenarioanalyse med makrodrivere; først når modellen predikerer godt utenfor estimeringsperioden, bør den brukes som løpende beslutningsstøtte. Tre tiltak kan iverksettes uavhengig av modellen: korte ned forsinkelsen (hyppigere statistikk, fullmakt til omprioritering innenfor budsjettåret) og dempe reaksjonen på enkeltårs utslag, som begge følger av proposisjon 1 alene, og en flaskehalsdiagnose av straffesakskjeden (avsnitt 9.5), som bare krever restanse- og saksbehandlingsstatistikk som allerede finnes, og avdekker om økt politiinnsats i dag omsettes i reaksjoner eller i kø. Den koblede simuleringen bør i første omgang brukes til å rangere tiltak, ikke til å tallfeste nivåer.

10 Diskusjon og begrensninger

Strukturell stabilitet. Lovbrytere tilpasser seg ny politikk, og parametere som er estimert under ett regime, gjelder ikke nødvendigvis under et annet (Lucas, 1976). Dette taler for jevnlig re-estimering, robuste styringsregler og forsiktighet med store politikkskift som bare begrunnes med modellen.

Målstyring og indikatorpress. Når en indikator brukes til styring, øker presset på den, og den kan miste informasjonsverdien sin (Campbell, 1979). Måleligningen (4) og bruken av flere uavhengige datakilder er modellens svar på dette, men løser ikke problemet fullt ut. Observerbarheten er den dypere begrensningen (avsnitt 2.4): registertall alene kan ikke skille faktisk kriminalitet fra registreringstilbøyelighet, og mellom offerundersøkelsene hviler estimatet på antakelser om hvor langsomt r_i endrer seg. Antakelsene kan gjøres eksplisitte og testes, men ikke unngås, og usikkerheten de gir, bør møtes med lavere reaktivitet og med at raske omprioriteringer bygger på de kategoriene der registreringsgraden er best kjent.

Forenklinger. Modellen er aggregert og ser bort fra romlig struktur, heterogenitet i befolkningen og fordelingsvirkninger. To politikkløp med samme $\mathcal{J}$ kan fordele byrdene svært ulikt mellom områder og grupper; totypemodellen i avsnitt 8.3 viser at også kapasitetstapet i domstolene har en fordelingsside, der prioritering verner de alvorlige sakene på de øvriges bekostning. Hensynet til likebehandling kan legges inn som ekstra skranker, for eksempel grenser for hvor mye innsatsen per innbygger kan variere mellom områder. Nettverk mellom lovbrytere er bare representert gjennom det aggregerte rekrutteringsleddet i avsnitt 8.4, som viser at selvforsterkende mekanismer kan gi vippepunkter, men ikke hvor de ligger; det krever en eksplisitt nettverks- eller markedsmodell og empirisk identifikasjon av terskelen. Politiets kapasitet er modellert med én metningsparameter, uten den sakstype- og distriktsvise strukturen som avgjør hvor presset faktisk oppstår. Funksjonsformene for p_i, Π_i, Λ, F og $\mathcal{H}$ er valgt for å være enkle og tolkbare. De må testes empirisk.

Beslutningsreglene. Reglene i avsnitt 8.5 er stiliserte representasjoner av hvordan påtale- og budsjettmyndighet kan svare på det de observerer, ikke beskrivelser av dagens praksis. Nivåene i tabell 6 avhenger av terskler, forsterkninger, dødsoner og ledetider; det som er robust, er mekanismene: at innstrømningen til en overbelastet domstol er et oppblåst signal, at to aktører som svarer fullt ut på samme signal overdimensjonerer, og at et felles skrankebilde med prognose fjerner begge deler. Den koordinerte regelen forutsetter at prognosen for den gjenopprettede likevekten er omtrent riktig; med en feilspesifisert intern modell arver den svakheten til MPC i avsnitt 8.2. Andelen saker som kan avgjøres forenklet, er i praksis bestemt av sakenes art og av hjemlene i straffeprosessloven, og er et rettspolitisk valg, ikke en styringsparameter. Kvalitetssløyfen bygger på årlige målinger og virker derfor på nivået, ikke på raske fall. Lovgivers sløyfe er ikke simulert; den inngår i metoden som et handlingsrom gitt på forhånd (avsnitt 7.1).

Normative valg. Vektene i K_{obj} og $\tilde{K}$, kostnadsfunksjonene og nivåene K^*, U^* og Λ_{krit} er verdivalg, ikke tekniske størrelser. Modellens fremste styrke er at slike valg blir eksplisitte og etterprøvbare.

Dynamikk og forutberegnelighet. Kravet om et lovbestemt handlingsrom (avsnitt 7.1) står i et reelt spenningsforhold til rettsstatens krav om forutberegnelighet, likebehandling og vern mot tilbakevirkning. Spenningen løses ikke ved å velge det ene framfor det andre, men ved å plassere dynamikken riktig: i prioritering, organisering, prosessformer og reaksjonsformer, som kan justeres innenfor rammelover og forsøkshjemler, og ikke i straffebudene og strafferammene,

som må være kjent på forhånd. Også innenfor dette rommet har dynamikken en pris. Hyppige endringer i prosessregler øker opplæringsbehovet, svekker rettspraksisens veiledende verdi og kan i seg selv bli en kilde til feil. Forsøksordninger reiser dessuten spørsmål om likebehandling mellom dem som omfattes, og dem som ikke gjør det. Kjedemodellen i avsnittene 7.2 og 7.3 er på sin side en forenkling: grunnmodellen har én kø uten prioritering, mens domstolene prioriterer blant annet varetektssaker og saker med unge lovbrytere, og totypemodellen i avsnitt 8.3 viser at dette betyr mye for hvem som rammes, men ikke for rangeringen av tiltakene. Ikke modellert er at sakstyper har ulik behandlingstid, at ledd deler ressurser, og at ankeadgangen gir tilbakeløp. Parameterne ϑ, χ_{D}, μ_4 og $\vartheta_{\mathrm{A}}, \vartheta_{\mathrm{B}}$ er stiliserte; rangeringen av tiltakene er robust, men nivåene er det ikke. Frafallet ς_{D} begrenser køen, men skjuler at henleggelse og foreldelse er beslutninger med egne rettssikkerhetsmessige kostnader, som med prioritering konsentreres i de sakene som prioriteres ned.

Empirisk forankring. Modellen er ikke kalibrert eller validert, og tre av premissene er ikke påvist. Det er ikke dokumentert at norsk rettshåndhevelse pendler mellom over- og underhåndhevelse; med realistisk reaktivitet kan sløyfen ligge trygt i det stabile området (avsnitt 8.1). Det er ikke dokumentert at domstolen er kjedens bindende ledd; dagens tall peker andre steder (avsnitt 1), og etterforskning, påtale og straffegjennomføring har verken kø, frafall eller kapasitet i simuleringene. Og måleproblemet, som er den første forutsetningen, er ikke prøvd: alle simuleringene antar et feilfritt estimat, og resultatene er enkeltbaner uten parameterusikkerhet. Kriminalitetsmålet blander dessuten en bestand av pågående virksomhet med strømmer av registrerte lovbrudd (avsnitt 2), og de ytre driverne er makrovariabler, mens utviklingen i norske tall i stor grad forklares av praksis- og regelendringer, anmeldelsesmønstre og nye kriminalitetsformer.

Institusjonelle rammer. Metoden forutsetter samordning mellom aktører som ikke kan instruere hverandre. Påtalemyndigheten er uavhengig i den enkelte sak, domstolene og dommerne er uavhengige, og leddene i kjeden har hver sine bevilgninger under ettårsprinsippet. Artikkelen beskriver ikke hvem som kan beslutte hva, med hvilken hjemmel, eller hvordan sløyfene i figur 3 passer inn i budsjettåret, tildelingsbrevene og den ordinære etatsstyringen. Modellens størrelser er normaliserte og er ikke oversatt til årsverk, saker eller kroner, og innsatsen behandles som fritt flyttbar, noe den ikke er. Prinsippet har også en insentivside som ikke er behandlet: en regel som gir mer ressurser der tallene er dårligst, krever måling som er uavhengig av dem som får ressursene. Forebygging er ikke en beslutningsvariabel, slik at modellen bare kan anbefale tiltak i straffesakskjeden, også der det mest virkningsfulle tiltaket ligger utenfor den.

Samfunnsøkonomi. Målfunksjonen (22) er stilt opp, men ikke brukt. Ingen kostnader eller gevinster er tallfestet, regimene er ikke sammenlignet på samfunnstap, og skyggeprisene er ikke beregnet. Gevinstene modellen peker på — mindre stopp-og-start, mindre overkapasitet, kortere køer — er vist som forløp, ikke som verdier, og kostnadssiden dekker bare politiets innsats. En samfunnsøkonomisk analyse etter statens rammeverk, med nullalternativ, enhetskostnader per ledd, kalkulasjonsrente, skattefinansieringskostnad, usikkerhet og fordelingsvirkninger, gjenstår (avsnitt 6).

Veiplan for videre arbeid. Artikkelen er ment som et utgangspunkt. Det som gjenstår, er i prioritert rekkefølge:

(1) en flaskehalsdiagnose av hele kjeden på norske tall, som avgjør hvilket ledd modellen bør bygges rundt;

(2) estimering av faktisk reaktivitet og forsinkelse i budsjett- og prioriteringsprosessene, og av stabilitetsgrensen for en inkrementell regel (avsnitt 4.4);

(3) en kjedemodell med kø og frafall i alle fire ledd, flere sakstyper, ulike behandlingstider og tilbakeløp;

(4) simuleringer med målestøy, skjevhet og parameterusikkerhet;

(5) kalibrering på norske data, herunder av politiets metningsnivå K_{sat} per sakstype, av virkningen av ulike reaksjonsformer og av om rekrutteringsterskler finnes i bestemte kategorier;

(6) kostnadsfunksjoner i kroner og en nytte-kostnadsvurdering av tiltakene, med forebygging som alternativ;

(7) en beskrivelse av beslutningsrettigheter, hjemler og budsjettmekanikk for samordningen;

(8) en romlig versjon med forskyvning mellom distrikter, og en modellprediktiv styring som optimerer hele tiltaksvektoren (7) på hver komponents tidsskala.

Simuleringskoden som gjenskaper tallene og figurene i artikkelen, gjøres åpent tilgjengelig sammen med den, slik at andre kan etterprøve, kritisere og bygge videre.

11 Konklusjon

Innsatsen for å håndheve lover og regler bør løpende stå i forhold til hvor mye, og hvor alvorlig, de faktisk brytes. Artikkelen har gitt dette prinsippet et presist innhold — et alvorlighetsvektet mål på de faktiske bruddene og en styringsregel som lar innsatsen følge avviket fra et akseptabelt nivå — og vist at gjennomføringen er et dynamisk styringsproblem med forsinkelser, strategiske motspillere, målefeil og absolutte normative grenser. Prinsippet gjelder alle tiltak, ikke bare politiets innsats: reaksjonsformer, kapasitet, prosessuell kvalitet og lovbestemt handlingsrom er bevisste responser på det observerte bildet, formulert som én beslutningsregel, $T \leftarrow f(K, \mathcal{O})$, fra kriminalitetsbildet og observasjonene til en samlet tiltaksvektor. Analysen gir fem hovedinnsikter; de fire første svarer til forutsetningene for å realisere prinsippet, den femte til hvordan tiltakene bør velges.

For det første må bruddene måles riktig. Registrert kriminalitet avhenger av både tillit og innsats, og en innsats som følger registreringen i stedet for de faktiske bruddene, kan skape selvforsterkende sløyfer; styringen bør derfor bygge på estimater som kombinerer registerdata med offer- og innbyggerundersøkelser. *For det andre* må reaksjonen være avdempet nok. En regel der man reagerer kraftig når kriminaliteten øker, kan gi vedvarende pendelsvingninger når forsinkelsen er lang. Svingningene er ikke tilsiktet: de oppstår selv når hver beslutning er bevisst og rimelig ut fra de tallene som foreligger, fordi tallene beskriver en situasjon som tiltak som ennå ikke har virket, allerede har endret. Stabilitetsgrensen kan beregnes, og moderat reaktivitet er robust uansett forsinkelse når responsen er proporsjonal. I simuleringene fjerner en avdempet regel svingningene og verner tilliten; prognosebasert styring fjerner i tillegg det varige avviket. Om norsk rettshåndhevelse faktisk ligger nær stabilitetsgrensen, er ikke undersøkt. *For det tredje* må tilpasningen skje innenfor legitimitetens grenser. Tillit er både et mål og et virkemiddel, og den svekkes av både for lite og uforholdsmessig mye håndhevelse. Tillit bør derfor behandles som en skranke og fremmes gjennom prosessuell kvalitet, ikke gjennom volum. *For det fjerde* kan styringen av politiinnsatsen ikke være mer dynamisk enn det lovverket og straffesakskjeden tillater. Lovgivningen er sløyfens langsomste ledd, og den bør ikke gjøres raskere gjennom hendelsesdrevet hastelovgivning, men gis et regelstyrt handlingsrom på forhånd — rammelover, forsøkshjemler, solnedgangsklausuler og innebygd evaluering — innenfor Grunnlovens og EMKs faste grenser. Et ledd nedstrøms kan være kjedens flaskehals; med domstolen som gjennomregnet eksempel omsettes da økt politiinnsats i restanser og lengre forsinkelse, og *Theory of Constraints* gir en metode

for å utnytte, underordne og utvide kapasiteten uten å svekke rettssikkerhetsgarantiene. Simuleringene viser at opptrapping ved en mettet domstol gir høyere kriminalitet og lavere tillit, at underordning virker raskt mens bare utvidelse gjenoppretter balansen varig, og at en mettet domstol kan maskere en ustabil styringssløyfe til kapasiteten utvides. Resultatene står seg med politimetning, som gjør sløyfen mer følsom, og med prioriterte domstolskøer, som verner de alvorlige sakene og flytter kostnaden til de øvrige. Der kriminalitet avler kriminalitet, kan systemet ha vippepunkter; da må den avdempede responsen også være tidlig og tilstrekkelig, og stabilitetsmarginen må måles som avstand til vippepunktet, ikke bare til svingningsgrensen. *For det femte* må tiltakene velges etter en metode som ser hele kjeden under ett: observer og estimer, sammenlign alvorlighetsvektet med målene, identifiser den bindende skranken, velg type tiltak etter skranken, dimensjoner avdempet og med prognose, kontroller de rettsstatlige skrankene, og gjenta. Når underordning, kapasitet og kvalitet gjøres til responser på observasjoner, gir desentraliserte regler, der hver aktør svarer rimelig på sitt eget signal, en kapasitet om lag 30 prosent over behovet og en dempet kapasitetssvingning over seks til sju år, fordi innstrømningen til en overbelastet domstol er et signal flaskehalsen selv har blåst opp; en koordinert regel med felles skrankebilde og prognose gir nesten samme utfall som vedtak fattet med full kjennskap til likevekten, med mindre kapasitet og uten svingninger.

Alt dette er resultater i en teoretisk modell med stiliserte parametere. Hvilket ledd som binder i dag, hvor kraftig styringen faktisk reagerer, og hva tiltakene koster og er verdt, er empiriske spørsmål som artikkelen stiller, men ikke besvarer (avsnitt 10).

Den viktigste praktiske anbefalingen er derfor tredelt: *mål forsinkelsen, demp reaksjonen og styr etter prognoser, ikke etter gårsdagens tall; dimensjoner lovverkets handlingsrom og straffesakskjedens kapasitet slik at styringen har et rom å virke i; og la alle som rår over et virkemiddel, svare på det samme bildet av hvor skranken binder.* Utviklingen i henleggelser, saksbehandlingstid og soningskø (avsnitt 1) understreker at dette er en helhetlig, systemisk oppgave: politi, påtalemyndighet, domstoler, kriminalomsorg og lovgiver må ses og styres som ett system. Straffebud og strafferammer ligger fast; regelverket må gi fleksibiliteten. Når disse forutsetningene er på plass, kan innsatsen for å håndheve lover og regler løpende stå i forhold til hvor mye, og hvor alvorlig, de faktisk brytes — uten at tilpasningen selv blir en kilde til ustabilitet, restanser eller svekket tillit.

Referanser

Altman, E. (1999). *Constrained Markov Decision Processes*. Chapman & Hall/CRC, Boca Raton.

Åström, K. J. og Murray, R. M. (2021). *Feedback Systems: An Introduction for Scientists and Engineers* (2. utg.). Princeton University Press, Princeton.

Ayres, I. og Braithwaite, J. (1992). *Responsive Regulation: Transcending the Deregulation Debate*. Oxford University Press, New York.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2):169–217.

Bhuller, M., Dahl, G. B., Løken, K. V. og Mogstad, M. (2020). Incarceration, recidivism, and employment. *Journal of Political Economy*, 128(4):1269–1324.

Blumstein, A. og Larson, R. (1969). Models of a total criminal justice system. *Operations Research*, 17(2):199–232.

Braga, A. A., Turchan, B. S., Papachristos, A. V. og Hureau, D. M. (2019). Hot spots policing and crime reduction: An update of an ongoing systematic review and meta-analysis. *Journal of Experimental Criminology*, 15(3):289–311.

Campbell, D. T. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*, 2(1):67–90.

Chalfin, A. og McCrary, J. (2017). Criminal deterrence: A review of the literature. *Journal of Economic Literature*, 55(1):5–48.

Conitzer, V. og Sandholm, T. (2006). Computing the optimal strategy to commit to. I *Proceedings of the 7th ACM Conference on Electronic Commerce (EC'06)*, s. 82–90.

Cook, P. J. og Zarkin, G. A. (1985). Crime and the business cycle. *Journal of Legal Studies*, 14(1):115–128.

Domstoladministrasjonen (2026). *Årsrapport 2025*. Trondheim.

Den norske dommerforening (2025). *Innspill til justiskomiteen om statsbudsjettet for 2026*. Oslo.

Ensign, D., Friedler, S. A., Neville, S., Scheidegger, C. og Venkatasubramanian, S. (2018). Runaway feedback loops in predictive policing. I *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, PMLR 81:160–171.

Europaparlamentet og Rådet (2024). Forordning (EU) 2024/1689 om harmoniserte regler for kunstig intelligens (KI-forordningen). *Den europeiske unions tidende*, L, 12.7.2024.

Fajnzylber, P., Lederman, D. og Loayza, N. (2002). Inequality and violent crime. *Journal of Law and Economics*, 45(1):1–40.

Forrester, J. W. (1961). *Industrial Dynamics*. MIT Press, Cambridge, MA.

Freeman, R. B. (1999). The economics of crime. I Ashenfelter, O. og Card, D. (red.), *Handbook of Labor Economics*, bd. 3C, s. 3529–3571. Elsevier, Amsterdam.

Goldratt, E. M. (1990). *What Is This Thing Called Theory of Constraints and How Should It Be Implemented?* North River Press, Croton-on-Hudson, NY.

Goldratt, E. M. og Cox, J. (1984). *The Goal: A Process of Ongoing Improvement*. North River Press, Croton-on-Hudson, NY.

Guerette, R. T. og Bowers, K. J. (2009). Assessing the extent of crime displacement and diffusion of benefits: A review of situational crime prevention evaluations. *Criminology*, 47(4):1331–1368.

Hale, J. K. og Verduyn Lunel, S. M. (1993). *Introduction to Functional Differential Equations*. Springer, New York.

Hayek, F. A. (1945). The use of knowledge in society. *American Economic Review*, 35(4):519–530.

Hayes, N. D. (1950). Roots of the transcendental equation associated with a certain difference-differential equation. *Journal of the London Mathematical Society*, s1-25(3):226–232.

Holling, C. S. (1959). Some characteristics of simple types of predation and parasitism. *The Canadian Entomologist*, 91(7):385–398.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45.

Kelly, M. (2000). Inequality and crime. *Review of Economics and Statistics*, 82(4):530–539.

Kingman, J. F. C. (1961). The single server queue in heavy traffic. *Mathematical Proceedings of the Cambridge Philosophical Society*, 57(4):902–904.

Kleiman, M. A. R. (1993). Enforcement swamping: A positive-feedback mechanism in rates of illicit activity. *Mathematical and Computer Modelling*, 17(2):65–75.

Kleiman, M. og Kilmer, B. (2009). The dynamics of deterrence. *Proceedings of the National Academy of Sciences*, 106(34):14230–14235.

Kleinrock, L. (1975). *Queueing Systems, Volume 1: Theory*. Wiley, New York.

Kriminalomsorgen (2026). *Nøkkeltall fra kriminalomsorgen for oktober 2026*. Pressemelding, november 2026.

Korzhyk, D., Conitzer, V. og Parr, R. (2010). Complexity of computing optimal Stackelberg strategies in security resource allocation games. I *Proceedings of the 24th AAAI Conference on Artificial Intelligence*, s. 805–810.

Levitt, S. D. (1997). Using electoral cycles in police hiring to estimate the effect of police on crime. *American Economic Review*, 87(3):270–290.

Little, J. D. C. (1961). A proof for the queuing formula: $L = \lambda W$. *Operations Research*, 9(3):383–387.

Løvgren, M., Skjervheim, Ø., Høgestøl, A., Kotsadam, A., Bjørnshagen, V., Hyggen, C. og Lid, S. (2023). *Nasjonal trygghetsundersøkelse 2022*. NOVA Rapport 14/23. NOVA, OsloMet, Oslo.

Lucas, R. E. (1976). Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy*, 1:19–46.

Lum, K. og Isaac, W. (2016). To predict and serve? *Significance*, 13(5):14–19.

McCrary, J. (2002). Using electoral cycles in police hiring to estimate the effect of police on crime: Comment. *American Economic Review*, 92(4):1236–1243.

Minorsky, N. (1922). Directional stability of automatically steered bodies. *Journal of the American Society for Naval Engineers*, 34(2):280–309.

Nagin, D. S. (2013). Deterrence in the twenty-first century. *Crime and Justice*, 42(1):199–263.

Nagin, D. S. og Telep, C. W. (2017). Procedural justice and legal compliance. *Annual Review of Law and Social Science*, 13:5–28.

Phillips, A. W. (1954). Stabilisation policy in a closed economy. *Economic Journal*, 64(254):290–323.
Politidirektoratet (2024). *Politiets innbyggerundersøkelse 2024: Hovedrapport*. Oslo.
Polinsky, A. M. og Shavell, S. (2000). The economic theory of public enforcement of law. *Journal of Economic Literature*, 38(1):45–76.
Ranchordás, S. (2014). *Constitutional Sunsets and Experimental Legislation: A Comparative Perspective*. Edward Elgar, Cheltenham.
Raphael, S. og Winter-Ebmer, R. (2001). Identifying the effect of unemployment on crime. *Journal of Law and Economics*, 44(1):259–283.
Rawlings, J. B., Mayne, D. Q. og Diehl, M. M. (2017). *Model Predictive Control: Theory, Computation, and Design* (2. utg.). Nob Hill Publishing, Madison, WI.
Rossi, F., van Beek, P. og Walsh, T. (red.) (2006). *Handbook of Constraint Programming*. Elsevier, Amsterdam.
Scattolini, R. (2009). Architectures for distributed and hierarchical Model Predictive Control – A review. *Journal of Process Control*, 19(5):723–731.
Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., Held, H., van Nes, E. H., Rietkerk, M. og Sugihara, G. (2009). Early-warning signals for critical transitions. *Nature*, 461:53–59.
Sherman, L., Neyroud, P. W. og Neyroud, E. (2016). The Cambridge Crime Harm Index: Measuring total harm from crime based on sentencing guidelines. *Policing*, 10(3):171–183.
Sherman, L. W. og Weisburd, D. (1995). General deterrent effects of police patrol in crime «hot spots»: A randomized, controlled trial. *Justice Quarterly*, 12(4):625–648.
Smith, O. J. M. (1957). Closer control of loops with dead time. *Chemical Engineering Progress*, 53(5):217–219.
Statistisk sentralbyrå (2026). *Flere henleggelser og færre gjerningspersoner*. Artikkel i statistikken Etterforskede lovbrudd, publisert 2. november 2026. Statistisk sentralbyrå, Oslo.
Sterman, J. D. (1989). Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Management Science*, 35(3):321–339.
Sterman, J. D. (2000). *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin/McGraw-Hill, Boston.
Sunshine, J. og Tyler, T. R. (2003). The role of procedural justice and legitimacy in shaping public support for policing. *Law & Society Review*, 37(3):513–548.
Tambe, M. (2011). *Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned*. Cambridge University Press, Cambridge.
Tsebelis, G. (1989). The abuse of probability in political analysis: The Robinson Crusoe fallacy. *American Political Science Review*, 83(1):77–91.
Tyler, T. R. (1990). *Why People Obey the Law*. Yale University Press, New Haven.
Wiener, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*. Technology Press, Cambridge, MA, og Wiley, New York.

A Symboloversikt

Tabell 9. Modellens variabler, parametere og teoretiske forankring.

Symbol	Begrep	Måleindikator / datakilde	Forankring
N_i	Aktivitetsnivå i kategori i	Estimert fra register og offerundersøkelser	—
A_i, S_i, θ	Alvorlighet, skadekostnad, aversjon	Strafferammer, kostnadsanslag	Becker (1968)
K_{obj}, K_{ref}	Vektet objektiv kriminalitet; referansenivå	Avledet fra (2)	Becker (1968)
U, Λ	Opplevd trygghet, tillit	Innbygger- og trygghetsundersøkelser	Tyler (1990)
$\tilde{K}$, ω, Φ	Effektivt kriminalitetspress	Avledet fra (3)	—
r_i, ε_i	Registreringsgrad, målefeil	Offerundersøkelser	Ensign mfl. (2018)

fortsetter på neste side

Symbol	Begrep	Måleindikator / datakilde	Forankring
$\mathbf{E},\ \Omega_i,\ \beta_k,\ \ell_k$	Makrovektor, strukturelt pådrag, koeffisienter, etterslep	SSB, NAV, Norges Bank	Freeman (1999)
τ	Systemisk tidsforsinkelse	Tidsstempler, budsjettkalender	Forrester (1961)
$T_i,\ T_0,\ T^{\max}$; $K^*,\ e$	Innsats, grunninnsats, kapasitetstak; målnivå, styringsavvik	Ressursregnskap; politisk vedtak	—
$k_p,\ k_i,\ k_d,\ W$	Styringsparametere (PID)	Estimert fra historisk praksis	Åström og Murray (2021)
$\gamma_i,\ \eta,\ \delta_i$	Effektivitet, tillitselastisitet, avgang	Kvasieksperimenter	Chalfin og McCrary (2017)
$\mathcal{U}_i^C,\ B_i,\ F_i$	Forventet nytte, gevinst, sanksjon	Estimert	Becker (1968)
$p_i,\ h_i$	Oppdagelsesrisiko og -intensitet	Kvasieksperimenter; ikke oppklaringsprosent alene	Nagin (2013)
$\Pi_i,\ M_i,\ s_i$	Deltakelse, potensial, heterogenitet	Estimert	Becker (1968)
σ_{ij}	Forskyvning (andel av fortrengt aktivitet)	Kontrollområdestudier	Guerette og Bowers (2009)
$\kappa_U, \psi, \kappa_\Lambda, \mu_k$	Trygghets- og tillitsdynamikk	Gjentatte undersøkelser	Sunshine og Tyler (2003)
$q,\ T_{\text{prop}}$	Prosessuell kvalitet; forholdsmessighet	Undersøkelser, tilsyn	Tyler (1990)
$c,\ a,\ b,\ \tau_c$	Tillitsjustert effektivitet; dempning, reaktivitet, kritisk forsinkelse	Avledet, ligningene (13) og (14)	Hayes (1950)
$\mathcal{J},\ \rho,\ C_{(\cdot)}$	Samfunnstap, diskonteringsrate, kostnadsfunksjoner	Valgt, ligning (22)	Polinsky og Shavell (2000)
$\langle \mathcal{X}, \mathcal{D}, \mathcal{C} \rangle$	Skrankeproblem	Budsjett $\mathcal{B}$, w_i, $T_i^{\max}$, Λ_{krit}, K^*, U^*	Rossi mfl. (2006)
$H,\ g,\ V_f,\ \boldsymbol{x}$	MPC-horisont, trinnkostnad, sluttkostnad, tilstand	Valgt	Rawlings mfl. (2017)
$T_i^{\text{lov}},\ T_i^{\text{grl}},\ \tau_{\text{lov}}$	Lovbestemt og konstitusjonell grense; lovgivningsforsinkelse	Lov og forskrift; Grunnloven, EMK	Ranchordás (2014)
$J_k,\ Z_k,\ Q_k$	Innstrømning, kapasitet og kø i ledd k av straffesakskjeden	Saksbehandlingsstatistikk, restanser	Goldratt og Cox (1984)
$\upsilon_k,\ \upsilon_k^{\max},\ \tau_k^{\text{kø}},\ \xi_{\text{D}}$	Utnyttelse, tillatt utnyttelse, køtid; andel saker til retten	Avledet fra (26); påtalestatistikk	Kleinrock (1975)
$R_{\text{D}},\ m_{\text{D}},\ \varsigma_{\text{D}}$	Avgjorte saker; ren behandlingstid; frafall fra køen	Domstolstatistikk, ligning (28)	Little (1961)
$\vartheta,\ \chi_{\text{D}},\ \mu_4,$ $\tau_{\text{h}},\ \tau_{\text{D}}^{\text{ref}}$	Andel av virkning som krever dom; celerity; tillitsvirkning; halvmetnings- og referansetid	Stiliserte, ligningene (29) og (30)	Nagin (2013)
$F,\ K_{\text{sat}},\ Z_{\text{E}},\ s$	Fjerning med metning; politiets halvmetningsnivå og kapasitet; metningsfaktor	Henleggelser av kapasitetsgrunner; ligningene (12′) og (13′)	Holling (1959)
$\pi_{\text{A}},\ \vartheta_{\text{A}},\ \vartheta_{\text{B}},$ $Q_k,\ R_k,\ \tau_k^{\text{kø}},\ c_{\text{D}}$	Alvorlige sakers andel; domsavhengig virkning per sakstype; kø, avgjørelser og ventetid per type; antall spor	Domstolstatistikk per sakstype, ligning (33)	Kleinrock (1975); Kingman (1961)
$\epsilon,\ K_{\text{h}},\ \varrho,\ \mathcal{H}$	Rekruttering: styrke, halvmetning, terskelbratthet, metningsfunksjon	Ikke estimert, ligning (34)	Scheffer mfl. (2009)
$\boldsymbol{v},\ \Psi,\ \mathcal{O}$	Tiltaksvektor, beslutningsregel, mengden av observasjoner	Avledet, ligningene (7)–(9)	Åström og Murray (2021)
$\tau_{\text{D}}^{\text{obs}},\ \xi_{\min},\ \tau_{\text{D}}^{\text{gr}},$ $k_\xi,\ \tau_\xi;$ $\Lambda^*,\ k_q,\ q_{\max}$	Påtalemyndighetens regel: observasjonsforsinkelse, gulv for andel til domstol, terskel for saksbehandlingstid, forsterkning, tidskonstant; kvalitetssløyfen: tillitsmål, forsterkning, øvre grense	Påtale- og domstolstatistikk, innbyggerundersøkelse; ligning (35), avsnitt 8.5	Goldratt (1990)

fortsetter på neste side

Symbol	Begrep	Måleindikator / datakilde	Forankring
υ^*, k_Z, τ_Z, τ_{ned}, Z^{bev}, H_Q, H_{p}, $\hat{F}^{\text{p}}$	Budsjettregel: målutnyttelse, forsterkning, ledetid, nedbyggingstid, bevilget kapasitet; horisont for restanseavvikling, prognosehorisont, prognostisert fjerningsstrøm	Budsjett, restansestatistikk, ligningene (36) og (37)	Sterman (1989, 2000)

B Bevis for proposisjon 1

Linearisering. Med $T(t) = T_0 + k_p(K_{\text{obj}}(t-\tau) - K^*)$ innenfor metningsgrensene og konstante $\bar{\Lambda}, \bar{\Omega}$ gir (12) $\dot{K}(t) = \bar{\Omega} - cT(t)K(t) - \delta K(t)$ med $c = \gamma\bar{\Lambda}^\eta$. Setter vi $K(t) = \bar{K} + x(t)$ og $T(t) = \bar{T} + k_p x(t-\tau)$, bruker likevektsbetingelsen $\bar{\Omega} = (c\bar{T} + \delta)\bar{K}$ og ser bort fra annenordensleddet $c\,k_p\,x(t)x(t-\tau)$, får vi $\dot{x}(t) = -(c\bar{T} + \delta)\,x(t) - c\bar{K}k_p\,x(t-\tau) = -a\,x(t) - b\,x(t-\tau)$, med $a > 0$ og $b > 0$.

Karakteristisk ligning. Innsetting av $x = e^{\lambda t}$ gir $\chi(\lambda) := \lambda + a + b\,e^{-\lambda\tau} = 0$. For $\tau = 0$ er eneste rot $\lambda = -(a+b) < 0$. Røttene avhenger kontinuerlig av τ, og bare endelig mange ligger i ethvert halvplan $\operatorname{Re}\lambda > -\epsilon$, så stabilitet kan bare tapes ved at en rot krysser den imaginære aksen (Hale og Verduyn Lunel, 1993). $\lambda = 0$ er ikke en rot, siden $a + b > 0$. Sett derfor $\lambda = i\varpi$ med $\varpi > 0$ (komplekskonjugerte røtter opptrer parvis). Real- og imaginærdelen gir

$$a + b\cos(\varpi\tau) = 0, \qquad \varpi - b\sin(\varpi\tau) = 0.$$

Kvadrering og addisjon gir $\varpi^2 = b^2 - a^2$.

Del (a). Hvis $b \le a$, finnes ingen $\varpi > 0$ som oppfyller dette. Ingen rot kan da krysse den imaginære aksen, og likevekten er asymptotisk stabil for alle $\tau \ge 0$.

Del (b). Hvis $b > a$, er $\varpi = \sqrt{b^2 - a^2}$, $\cos(\varpi\tau) = -a/b$ og $\sin(\varpi\tau) = \varpi/b > 0$. Kryssinger skjer dermed ved $\tau_m = \big(\arccos(-a/b) + 2\pi m\big)/\varpi$, $m = 0, 1, 2, \ldots$, og den minste er $\tau_c = \tau_0 = \arccos(-a/b)/\varpi$. For kryssingsretningen deriverer vi $\chi(\lambda(\tau)) = 0$ implisitt og bruker $be^{-\lambda\tau} = -(\lambda + a)$, som gir $(\mathrm{d}\lambda/\mathrm{d}\tau)^{-1} = -1/\big(\lambda(\lambda + a)\big) - \tau/\lambda$. For $\lambda = i\varpi$ er siste ledd rent imaginært, mens $\operatorname{Re}\big[-1/(i\varpi(i\varpi + a))\big] = \varpi^2/(\varpi^4 + a^2\varpi^2) > 0$. Siden et komplekst tall og dets inverse har realdel med samme fortegn, er $\operatorname{Re}(\mathrm{d}\lambda/\mathrm{d}\tau) > 0$ ved hver kryssing τ_m: røttene krysser alltid fra venstre mot høyre halvplan, og stabiliteten gjenvinnes ikke for større τ. Likevekten er altså asymptotisk stabil hvis og bare hvis $\tau < \tau_c$, og ved $\tau = \tau_c$ har den kritiske roten frekvens ϖ og periode $2\pi/\varpi$.

Korollar 1. For $0 < a < b$ er $\arccos(-a/b) \in (\pi/2, \pi)$ og $\sqrt{b^2 - a^2} < b$, så $\tau_c > \pi/(2b)$. Betingelsen $\tau < \pi/(2b) = \pi/(2c\bar{K}k_p)$ er derfor tilstrekkelig for stabilitet i del (b), og i del (a) er sløyfen stabil uansett. □

Merknad 2: linearisering med metning. La $F(K,T) = c'T\varphi(K)$ med $\varphi(K) = K/(1+K/K_{\text{sat}})$ og $c' = \gamma'\bar{\Lambda}^\eta$. Da er $\varphi'(K) = \varphi(K)\,s(K)/K$ med $s(K) = 1/(1 + K/K_{\text{sat}})$. Linearisering av $\dot{K} = \bar{\Omega} - c'T\varphi(K) - \delta K$ rundt $(\bar{K}, \bar{T})$ med $T = \bar{T} + k_p x(t-\tau)$ gir $\dot{x} = -\big(c'\bar{T}\varphi'(\bar{K}) + \delta\big)x(t) - c'\varphi(\bar{K})k_p\,x(t-\tau)$. Med $\bar{r} := c'\bar{T}\varphi(\bar{K})/\bar{K}$ er $c'\bar{T}\varphi'(\bar{K}) = \bar{r}\,s(\bar{K})$ og $c'\varphi(\bar{K})k_p = \bar{r}\bar{K}k_p$, som er (13′). Kalibreringen $\gamma' = \gamma(1 + K_{\text{ref}}/K_{\text{sat}})$ gir $\varphi(K_{\text{ref}})\gamma' = \gamma K_{\text{ref}}$, så $\bar{r} = c\bar{T}$ i referanselikevekten uavhengig av K_{sat}. For retningen på τ_c: ved fast b er $\partial\tau_c/\partial a = \big[1 + a\arccos(-a/b)/\sqrt{b^2 - a^2}\,\big]/(b^2 - a^2) > 0$ (siden $\mathrm{d}\arccos(-a/b)/\mathrm{d}a = 1/\sqrt{b^2 - a^2}$), så en lavere dempning gir en kortere kritisk forsinkelse. Grensen for stabilitet uansett forsinkelse, $b \le a$, blir $k_p \le (\bar{r}s(\bar{K}) + \delta)/(\bar{r}\bar{K})$. □

Merknad 4: kapasitetssløyfen. La innstrømningen J være gitt, og la budsjettet hvert år av lengde Δ korrigere andelen k_Z av gapet mellom nødvendig kapasitet J/υ^* og realisert kapasitet, med realisering τ_Z etter vedtaket. Erstattes de årlige vedtakene med en kontinuerlig rate, er $\dot{Z}(t) = (k_Z/\Delta)\big(J/\upsilon^* - Z(t-\tau_Z)\big)$, så $z = Z - J/\upsilon^*$ følger $\dot{z}(t) = -b\,z(t-\tau_Z)$ med $b = k_Z/\Delta$: proposisjon 1 med $a = 0$, der $\tau_c = \arccos(0)/b = \pi/(2b)$, så sløyfen er stabil hvis og bare hvis $k_Z\tau_Z/\Delta < \pi/2$. Regnes bevilget, men ikke realisert kapasitet med, korrigeres gapet mot den bevilgede kapasiteten $C(t)$, som er realisert når det nye vedtaket virker; da er $\dot{C}(t) = (k_Z/\Delta)\big(J/\upsilon^* - C(t)\big)$ uten forsinkelse, C går monotont mot behovet for alle $k_Z > 0$, og Z følger med ledetiden τ_Z. Nedbygging med tidskonstant τ_{ned} gir $a = 1/\tau_{\text{ned}} > 0$ og øker τ_c. □

Bistabilitet i (34). Ved fast T og Λ er likevektene røttene i $f(K) = \Omega + \epsilon\mathcal{H}(K) - aK$ med $a = c\bar{T} + \delta$. Siden $f(0) = \Omega > 0$ og $f(K) \to -\infty$, finnes minst én rot, og antallet er odde. $\mathcal{H}$ er for $\varrho > 1$ konveks på $[0, K^\dagger]$ og konkav deretter, med $K^\dagger = K_{\text{h}}\big((\varrho-1)/(\varrho+1)\big)^{1/\varrho}$ og største stigning $\mathcal{H}'(K^\dagger)$. Er $\epsilon\mathcal{H}'(K) \le a$ for alle K, er f strengt avtagende og roten entydig; tre røtter krever derfor $\epsilon\mathcal{H}'(K^\dagger) > a$. De to ytre røttene er stabile ($f' < 0$), den midterste (separatrisen) ustabil; røttene oppstår eller forsvinner parvis der $f = f' = 0$ (tangering), som er sadelnode-bifurkasjonen i teksten. Rundt den lave likevekten er den lokale dempningen $-f'(\bar{K}) = a - \epsilon\mathcal{H}'(\bar{K})$, som erstatter a i proposisjon 1 og går mot null nær tangeringen. Med proporsjonalregelen erstattes aK av $\big(c\,\mathrm{sat}_{[0,T^{\max}]}[T_0 + k_p(K - K^*)] + \delta\big)K$, som vokser raskere enn lineært inntil taket nås og lineært deretter; en høy likevekt finnes da bare når $T^{\max}$ er lite, eller, med metning (12′), når $Z_{\text{E}}(T^{\max})$ ikke overstiger $\Omega + \epsilon$ med tilstrekkelig margin. □