Appendix for Enforcement Policy in a Dynamic Model of Deterrence

Clifford Bekar* Kenneth I. Carlaw† B. Curtis Eaton‡

Revised January 30, 2026

*Associate Professor, Economics, Lewis and Clark College, bekar@lclark.edu.

†Professor, Economics, University of British Columbia - Okanagan, kenneth.carlaw@ubc.ca.

‡Emeritus Professor, Economics, University of Calgary and Honorary Professor, Economics, University of British Columbia - Okanagan, eaton@ucalgary.ca.

Computational Tools, Replication, and Robustness

In Section 1 of this appendix we show that our dynamic deterrence model is a regular, finite Markov chain and we describe and benchmark the main computational tools used to explore the model. In Section 2 we provide links to the code used to produce all of the figures and tables in the paper, allowing interested readers to replicate our results. In Section 3 we provide support for our assertion that "within limits, our results are robust to variations in the parameters and alternative versions of the model." In particular, we address two important criticisms of our model. Our main results are clearly not dependent on the assumption that the net utilities of violating the law for all potential violators in all periods are drawn from the same probability distribution. Nor are they dependent on the assumption that all potential violators use the same information to calculate their subjective probabilities of apprehension. In Section 4 we examine the robustness of our main policy results with respect to variation in the parameters of the model.

1 Analytical results and computational tools

This section first demonstrates that the time series of violations and apprehensions is generated by a regular finite Markov chain with state space $\boldsymbol{Z}$, where $\boldsymbol{Z}$ is the set of possible histories $((v^{t-z}, ..., v^{t-1}), (a^{t-z}, ..., a^{t-1}))$. Then the details for the benchmarking exercise for estimating stationary distributions of the Markov process, which is central for our convergence simulations, and the search algorithm for multi-bin policies are provided.

1.1 Deterrence as a Markov Process

In this subsection we show that the assumptions set out in Section 2 of the paper define a finite, regular Markov chain in the space of all possible histories of apprehensions and violations over the previous z periods, $\boldsymbol{Z}$. The history at time t is

$$h^t = ((v^{t-z}, ..., v^{t-1}), (a^{t-z}, ..., a^{t-1}))$$

where the v's and a's are the numbers of violations and apprehensions in the periods indicated. $\boldsymbol{Z}$ is the set of all histories $((v^{t-z},\ldots,v^{t-1}),(a^{t-z},\ldots,a^{t-1}))$ such that for all $k \in \{t-z,\ldots,t-1\}$, $0 \leq v^k \leq N$ and $0 \leq a^k \leq v^t$. Any history, h^t, includes z pairs (v^k, a^k). Each pair is an element of the set of possible pairs, which has $L = 1 + 2 + \ldots + (N+1)$ elements (the 1 in the sum corresponds to the case in which $v^k = 0$ and $a^k = 0$, the 2 to the case in which $v^k = 1$ and $a^k = 0$ or 1, and so on). So the cardinality of $\boldsymbol{Z}$ is L^z.

Before we prove the Markov chain result it is useful to recall the assumptions made in Section 2 of the paper. In any period t each person has the opportunity to engage in some form of antisocial behaviour (ASB). Person i's net utility of doing so, g_i^t , is a random draw from $\phi(g_i^t)$, a normal probability distribution with mean μ and standard deviation σ.

The deterrence mechanism consists of a law proscribing the ASB, a monetary sanction, F, on apprehended violators, and an apprehension technology. The objective probability that a violator will be apprehended in period t is the following function of R^t and v^t:

$$p(R^t, v^t) = \gamma \cdot min\left(1, \frac{R^t}{v^t}\right),$$

where v^t is the number of violations in period t, R^t is the quantity of enforcement resources deployed in that period, and $0 \leq \gamma \leq 1$.

The common subjective probability of apprehension is the following function of h^t:

$$q(h^t) = \frac{\alpha + a^{t-z} + \cdots + a^{t-1}}{\alpha + \beta + v^{t-z} + \cdots + v^{t-1}},$$

where $\alpha > 0$ and $\beta > 0$ are Bayesian priors.

Person i violates the law if and only if the net utility of doing so exceeds the expected cost:

$$g_i^t \geq q(h^t) \cdot F.$$

The enforcement policy, P, specifies the quantity of enforcement resources, R^t, to be

deployed in period t as a function of h^t:

$$R^t = P(h^t),$$

where $0 \leq R^t \leq N$.

The sequence of events in any period is the following: (i) $R^t = P(h^t)$ units of the enforcement resource are deployed; (ii) every person realizes their g_i^t and calculates the common subjective probability of apprehension, $q(h^t)$; (iii) everyone simultaneously chooses violate or comply; and, finally, (iv) each violator is apprehended and sanctioned with probability $p(R^t, v^t)$.

We denote the the probability of transitioning between any two states, h^t and h^{t+1}, in one period as $T(h^t, h^{t+1})$, and the entire transition matrix as $\boldsymbol{T} = [T(h^t, h^{t+1})]$. A Markov chain is completely described by its transition matrix, so we begin the proof that our model defines a finite regular Markov chain by calculating $T(h^t, h^{t+1})$ for any pair of arbitrary states.

Suppose the state at time t is,

$$\widehat{h}^t = ((\hat{v}^{t-z}, ..., \hat{v}^{t-1}), (\hat{a}^{t-z}, ..., \hat{a}^{t-1})).$$

At time $t+1$ the state will be

$$\widehat{h}^{t+1} = ((\hat{v}^{t-z+1}, ..., \hat{v}^{t-1}, v^t), (\hat{a}^{t-z+1}, ..., \hat{a}^{t-1}, a^t)),$$

where $v^t \in \{0, ..., N\}$ and $a^t \in \{0, ..., v^t\}$. Elements $((\hat{v}^{t-z+1}, ..., \hat{v}^{t-1}), (\hat{a}^{t-z+1}, ..., \hat{a}^{t-1}))$ in $\widehat{h}^{t+1}$ are predetermined by the history of the process; v^t is determined by the choices of potential violators, and a^t by the operation of the enforcement mechanism in period t. The probability of transitioning from $\widehat{h}^t$ to any state that does not satisfy the predetermination restrictions is, of course, 0.

The probability of transitioning from state $\widehat{h}^t$ to $\widehat{h}^{t+1}$, $T(\widehat{h}^t, \widehat{h}^{t+1})$, is the product of two

conditional probabilities. The first is the probability that, conditional on being in state $\widehat{h}^t$, exactly v^t violations occur, $prob(v^t|\widehat{h}^t)$. The second is the probability that, conditional on there being v^t violations in state $\widehat{h}^t$, exactly a^t of the violators are apprehended, $prob(a^t|v^t, \widehat{h}^t)$.

Person i will violate if and only if $g_i^t \geq q(\widehat{h}^t) \cdot F$. Since all $g_i^t s$ are random draws from the same normal probability distribution $\phi(g_i)$, all the $\pi_i s$ are the same function of $\widehat{h}^t$,

$$\pi(\widehat{h}^t) = \int_{q(\widehat{h}^t)\cdot F}^{\infty} \phi(g_i) dg_i > 0.$$

Further, since people share a common subjective probability of apprehension all N individual probabilities, $\pi_1(\widehat{h}^t), ...\pi_N(\widehat{h}^t)$ are equal. So $\pi_i(\widehat{h}^t) = \pi(\widehat{h}^t)$. Then $prob(v^t|\widehat{h}^t)$ is the multinomial probability that v^t of the N potential violators will choose violate, given that anyone of them will do so with probability $\pi(\widehat{h}^t)$.

Next, $prob(a^t|v^t, \widehat{h}^t)$ is the binomial probability that a^t of the v^t violators will be apprehended given that each of them is apprehended with probability $p(P(\widehat{h}^t), v^t)$. So, the probability of transitioning from $\widehat{h}^t$ to $\widehat{h}^{t+1}$ is,

$$T(\widehat{h}^t, \widehat{h}^{t+1}) = prob(v^t|\widehat{h}^t) \cdot prob(a^t|v^t, \widehat{h}^t) > 0.$$

Since both conditional probabilities are strictly positive, the probability of transitioning from any state $\widehat{h}^t$ to $\widehat{h}^{t+1}$ in one period is strictly positive.

To establish that the Markov chain is regular we show that starting from any arbitrary state $\widehat{h}^t$ in period t, the probability of being in any state in h^{t+z} at time $t+z$ is strictly positive. Notice that in period $t+z$ all of the predetermined elements of the original history, $\widehat{h}^t$, have been replaced by elements of the subsequent history of length z.

$$
\begin{aligned}
\widehat{h}^{t} &= ((\hat{v}^{t-z}, \cdots, \hat{v}^{t-1}), (\hat{a}^{t-z}, \cdots, \hat{a}^{t-1})) \\
\widehat{h}^{t+1} &= ((\hat{v}^{t-z+1}, \cdots, \hat{v}^{t-1}, v^{t}), (\hat{a}^{t-z+1}, \cdots, \hat{a}^{t-1}, a^{t})) \\
\widehat{h}^{t+2} &= ((\hat{v}^{t-z+2}, \cdots, \hat{v}^{t-1}, v^{t}, v^{t+1}), (\hat{a}^{t-z+2}, \cdots, \hat{a}^{t-1}, a^{t}, a^{t+1}) \\
&\vdots \\
h^{t+z} &= ((v^{t}, v^{t+1}, \cdots, v^{t+z}), (a^{t}, a^{t+1}, \cdots, a^{t+z}))
\end{aligned}
$$

The probability of transitioning from $\widehat{h}^{t}$ to h^{t+z} in z periods is which is just the joint probability of transitioning in one period through each of the relevant histories, $\widehat{h}^{t}, \widehat{h}^{t+1}, \cdots, h^{t+z}$:

$$
\begin{aligned}
&T(\widehat{h}^{t}, \widehat{h}^{t+1}) \cdot \\
&T(\widehat{h}^{t+1}, \widehat{h}^{t+2}) \cdot \\
&T(\widehat{h}^{t+2}, \widehat{h}^{t+3}) \cdot \\
&\vdots \\
&T(\widehat{h}^{t+z-1}, h^{t+z}).
\end{aligned}
$$

Since each of the above probabilities is strictly positive so too is the joint probability.

1.2 Convergence criteria and benchmarking

Initiating a simulation requires values for: the quantity of enforcement resources (R); the length of the relevant history (z); the number of potential violators (N); the mean (μ) and standard deviation (σ) of the PDF from which each persons private gain from committing a violation is drawn $\phi(g)$; the probability a violator is apprehended if investigated (γ); the utility penalty paid by apprehended violators (F); and, the Bayesian priors regarding apprehension (α, β). Unless noted, we set these parameters according to our baseline parameterization in both codebases.

For each step of the simulation:

- Each potential violator is assigned a $g_i^t, \forall i \in N$ drawn randomly from $\phi(g)$ with $\mu = 0.6$ and $\sigma = 0.2$.
- A common subjective probability of apprehension, q, is calculated using the history of violations and apprehensions, $q^t = \frac{\alpha + a^{t-z} + \ldots + a^{t-1}}{\alpha + \beta + v^{t-z} + \ldots + v^{t-1}}$, where $\alpha = 1$, $\beta = 0.25$.
- People choose to violate if and only if their $g_i^t \geq q^t \cdot F$, where $F = 1$.
- Each violator is apprehended with probability $p(P(h^t), v^t) = p^t = \gamma \cdot \min(1, \frac{R^t}{v^t})$, where v^t is the number of violators in the current period and $\gamma = 0.8$.
- The number of violations, apprehensions, and the sum of the gains from violating (i.e., the sum of the realized draws on $\phi(g)$ for those who violated) are recorded.
- The frequency distribution of violations and apprehensions is calculated for the current period and recorded.
- The history is updated and the simulation moves to the next period.

We calculate the cumulative frequency distribution of violations and apprehensions up to time t. The frequency distribution of violations is $\boldsymbol{FV}(nv_0^t, nv_1^t, \ldots, nv_N^t)$, where nv_j^t is the number of instances in the first t periods of the simulation in which there were exactly j violations.

The simulation proceeds in blocks of $50,000$ ticks. At the end of each block n we use $\boldsymbol{FV}$ to calculate the relative frequency distribution of violations $\boldsymbol{RFV}^n = (rv_0^n, \ldots, rv_N^n)$, where rv_j^n is the proportion of periods in which there were exactly j violations in the first $50000 \cdot n$ periods of the simulation. As n approaches infinity $\boldsymbol{RFV}^n$ approaches $\boldsymbol{D}$. The question is: How large must n be to ensure that $\boldsymbol{RFV}^n$ is a *good* estimate of $\boldsymbol{D}$? To answer this we need a convergence criterion.

At the end of each simulation block, for all $n \geq 2$, we calculate the following test statistic:

$$TEST^n = \sum_{j=0}^{j=N} |rv_j^n - rv_j^{n-1}|$$

$TEST^n$ is a measure of the *distance* between $\boldsymbol{RFV}^n$ and $\boldsymbol{RFV}^{n-1}$. $TEST^n$ does not decline monotonically, which complicates any test of convergence. After extensive experimentation we settled on the following criterion: we require that $TEST^n \leq 0.01$ for 5 consecutive blocks of n. Denoting the minimum n for which the convergence criterion is satisfied by n^* yields our approximation of the stationary distribution, $\boldsymbol{RFV}^{n^*} \sim \boldsymbol{D}$.

The convergence criteria represent a minimum requirement to produce a reliable estimate of $\boldsymbol{D}$. An important confounding issue for finding n^* is that the *shape* of $\boldsymbol{D}$ changes dramatically as R changes. For low values of R, $\boldsymbol{D}$ is unimodal and right skewed, for high values unimodal and left skewed, for intermediate values it becomes bimodal. This renders n^* non-monotonic in R (see the last row of Table A.1). In some cases order to produce "smooth" plots we run the simulation longer than required by the $TEST$ criteria alone.

Our convergence simulator has some limitations over certain ranges of some key parameter values, in particular those that comprise the composite parameter $\frac{\sigma}{(\gamma \cdot F)}$. If $\frac{\sigma}{(\gamma \cdot F)}$ is sufficiently small the n required to achieve a good approximation of the stationary distribution is actually much larger than the n^* generated from the convergence criteria. This is due to the high persistence of the q-attractors. For these parameterizations the simulator is highly sensitive to initial conditions, staying for a very long time (sometimes the entire simulation) in only one of the q-attractors and rarely (especially for some convergence simulations around the cliff) transitioning to the other q-attractor, even though the stationary distribution is bi-modal in these parameter values. This yields a poor estimate of the stationary distribution even though the convergence simulator may have taken a very long time to converge. More accurate estimations of the stationary distribution for such parameter values would take orders of magnitude longer, longer than is feasible to make simulation use-

ful. These limitations inform the choice of our baseline parameterization and our robustness testing. It is important to note that this is not a limitation of the model, but rather of our computational simulation capabilities.

We turn now to the issue of benchmarking our simulation to a known transition matrix. Given the number of states in $\boldsymbol{Z}$ for $z = 1$, it is only manageable to calculate $\boldsymbol{T}$ for $N \leq 50$. Doing so allows us to directly compute $\boldsymbol{D}$. But for $z > 1$ and larger N we must use numerical simulation methods to estimate $\boldsymbol{D}$. We benchmark our simulated estimates of $\boldsymbol{D}$ by ensuring they replicate $\boldsymbol{T}$ for the case of three one-bin policies in which $R = \{5, 21, 45\}$, $z = 1$ and $N = 50$.

We employ Monte Carlo simulations and the relationship between $\boldsymbol{D}$ and $\boldsymbol{RFV}^{n^*}$ to: (i) cross validate the output from our Python and Matlab codebases; and, (ii) benchmark the accuracy of the simulator against a known transition matrix. The computational intensity of producing both $\boldsymbol{D}$ and $\boldsymbol{RFV}^{n^*}$ increases geometrically with the size of the state space. It is only practical, therefore, to implement both approaches for a relatively small N ($N = 50$) and $z = 1$. For these parameters, and for each element of $R = \{5, 21, 45\}$, we generate 1000 instances of $\boldsymbol{RFV}^{n^*}$. We then construct the following distance metric:

$$ATEST = \boldsymbol{D} - \boldsymbol{RFV}^{n^*} = \sum_{j=0}^{j=N} |d_j - rv_j^{n^*}|$$

In Table 1 we report the mean, standard deviation, and maximal value of $ATEST$ along with the mean n^*. It is apparent that the two platforms deliver results that are essentially the same, and that $\boldsymbol{RFV}^{n^*}$ very closely approximates $\boldsymbol{D}$ as $n \to 400,000$.

Code for benchmarking of the simulator can be found at:

- Python file(s): Simulator.ipynb
- Matlab file(s): ASB BenchSD, ASB BenchTMSD,

Table 1: Benchmarking the simulator

	Matlab Codebase			Python Codebase		
	Value of R					
	5	**21**	**45**	**5**	**21**	**45**
Mean of ATEST	0.0017	0.0098	0.0078	0.0018	0.0115	0.0087
Std Dev of ATEST	0.0009	0.0017	0.0011	0.0009	0.0019	0.0011
Max value ATEST	0.0061	0.0193	0.0116	0.0057	0.0198	0.0128
Mean n^*	6.0010	7.8430	7.0870	6.0010	7.8120	6.0010

1.3 Search algorithm for multi-bin policies

We use a two stage procedure to identify the approximately optimal Two-Bin and Three-Bin policies for the baseline parameterization. In the first stage, we use a directed search routine with a loose convergence criterion to identify the neighborhood in which the optimal policy lies, and in the second stage we use Monte Carlo methods to generate tight estimates of the cost of policies in the neighborhood. In checking the robustness of our policy analysis in Section 4 of this appendix we use loser criteria to find the approximately optimal policy. This is done for two reasons. First, in our search for the optimal Multi-Bin policies in the baseline we discovered the optimal policy identified in the first stage of the process was invariably the optimal policy after the search of its neighborhood and that in the neighborhood all policies produced costs very close to each other (the cost surface was "flat" in the relevant neighborhood). Second, there is a high computational price for implementing the complete two stage procedure and our analysis in Section 4 does not require its precision. So instead, in the first stage of the procedure we identify the approximately optimal policy from a random seeding of 50 initial policies.

We describe the algorithm at the core of the directed search routine in terms of the v^{t-1} proxy. Required modifications for the AR^t proxy are indicted in the relevant codebase. For a given Two-Bin enforcement policy at stage s, $P^s = (BE^s, R_1^s, R_2^s)$, a modified policy $P^{s+1} = (BE^{s+1}, R_1^{s+1}, R_2^{s+1})$, is generated using a three step procedure: in the first step, BE^{s+1} is chosen from the set $BE^s - 5, BE^s - 4, \ldots, BE^s + 5\}$ to minimize the cost of ASB for policy

(BE^{s+1}, R_1^s, R_2^s); in the second step, R_1^{s+1} is chosen from the set $\{R_1^s - 5, R_1^s - 4, \ldots, R_1^s + 5\}$ to estimate the minimum cost of ASB for policy $(BE^{s+1}, R_1^{s+1}, R_2^s)$; in the third step, R_2^{s+1} is chosen from the set $\{R_2^s - 5, R_2^s - 4, \ldots, R_2^s + 5\}$ to estimate the minimum cost of ASB for policy $(BE^{s+1}, R_1^{s+1}, R_2^{s+1})$. For each of the 33 policies considered, we ran a convergence simulation and used it to estimate the expected cost of ASB. The search is terminated when the criteria $(EC(P^s) - EC(P^{s+1}))/EC(P^s) < .02$ is met. Each of the 1000 simulations was seeded with an initial policy randomly chosen from the set $\{(BE, R_1, R_2)|0 \leq BE \leq 100, 0 \leq R_1 \leq 100, R_1 \leq R_2 \leq 100\}$ and for each we used the directed search algorithm to a generate a terminal policy. From initial policy to terminal policy, the average number of policies evaluated was 165. To identify the neighborhood of the optimal policy, we ranked the 1000 terminal policies from lowest to highest cost. The first 78 policies in this ranking, and 98 of the first 100, were in the set $\{(BE, R_1, R_2)|30 \leq BE \leq 39, 29 \leq R_1 \leq 32, 53 \leq R_2 \leq 65\}$. This is the neighborhood of the optimal policy.

The search for the optimal Three-Bin enforcement policy was analogous to that for the Two-Bin enforcement policy policy, with the differences being the need for two additional steps in the procedure to search over the two additional parameters, BE_2 and R_3, of the Three-Bin enforcement policy. The algorithm for the Three-Bin enforcement policy search is, given a Three-Bin enforcement policy at stage s, $P^s = (BE_1^s, BE_2^s, R_1^s, R_2^s, R_3^s)$, a modified policy $P^{s+1} = (BE_1^{s+1}, BE_2^{s+1}, R_1^{s+1}, R_2^{s+1}, R_3^{s+1})$, is generated using a five step procedure analogous to that of the three step procedure of the multi-bin enforcement policy policy search. For each of the 55 policies considered, we again ran a convergence simulation and used it to estimate the cost. The termination criteria for the Three-Bin enforcement policy search was the same as that used for the Two-Bin policy search. The 1000 simulations were randomly seeded with refined policies chosen from the set $\{(BE_1, BE_2, R_1, R_2, R_3)|0 \leq BE_1 \leq 100, BE_1 \leq BE_2 \leq 100, 0 \leq R_1 \leq 100, R_1 \leq R_2 \leq R_3, R_2 \leq R_3 \leq 100\}$. From initial policy to terminal policy, the average number of policies evaluated was 440 for each of the 1000 simulations. The neighborhood of the optimal Three-Bin enforcement policy was

determined in the same way as for the Two-Bin policy.

The cost estimate produced by any one convergence simulation is, obviously, a random variable. In the neighborhood of the optimal policy, the cost function is so flat that in order to reliably identify the optimum policy, many independent cost estimates are needed. Accordingly, for every policy in the neighborhood of the optimum we ran 150 convergence simulations and calculated the cost of ASB in the steady state for each of them. Our cost estimate is the mean value of these 150 cost estimates. Typically the standard deviation of cost for the 150 simulations is close to 0.20, and hence the standard error of the estimate is close to $0.2/\sqrt{150} = 0.016$.

The code for the search algorithms and the Monte Carlo simulations over the set of policies in the neighborhood of the optimum can be found here:

Matlab file(s):

- Two-Bin enforcement policy policy v^{t-1}
- Three-Bin enforcement policy policy v^{t-1}
- Two-Bin enforcement policy policy Monte Carlo v^{t-1}
- Three-Bin enforcement policy policy Monte Carlo v^{t-1}
- Three-Bin enforcement policy policy AR^t
- Two-Bin enforcement policy policy AR^t

2 Replicating figures and baseline results

Here we layout and explain the code for replicating all of the figures in the paper.

2.1 Figure 1: Simulated time series of violations for chosen values of R

Figure 1 has five panels, each panel has two columns. The left panels of Figure 1 each plot a random sample of violations for $5,000$ periods. The right panels plot the resulting frequency distribution of violations. This is done for five one-bin policies, $R = \{39, 40, 41, 42, 43\}$ in our baseline parametrization.

The code for producing Figure 1 from Section 3 can be found here:

- Matlab file(s): Figure 1

2.2 Figure 2: Dynamics of quasi-attractors and the distribution of violations $R = 39$

Figure 2 plots $E(v^t|v^{t-1})$ for $R = 39$. The figure has two insets. The left inset plots the frequency distribution for $\boldsymbol{D}$. The right inset plots the autocorrelation coefficients for 1, 2, and 3 lags. The intersection of $E(v^t|v^{t-1})$ and the plotted 45-degree line from the origin reveal the system's q-attractors.

Code for producing Figure 2 from Section 3.1 can be found here:

- Matlab file(s): Figure 2

2.3 Figure 3: Persistence of GA and BA

Figure 3 has three panels: (i) Panel 1 reports the persistence of the q-attractors; (ii) Panel 2 the time spent in them; and, (iii) Panel 3 reports expected violations in a period for each q-attractor and in total. All are reported for one-bin policies in the range $15 \leq R \leq 55$.

Code for producing Figure 3 from Section 3.1 can be found here:

- Matlab file(s): Figure 3

2.4 Figure 4: Observable indices related to the cost of ASB

Figure 4 has three panels: (i) Panel 1 reports the compliance rate; (ii) Panel 2 the reserve capacity rate; and, (iii) Panel 3 the deficient capacity rate. All are reported for one-bin policies in the range $0 \leq R \leq N$.

Code for producing Figure 4 from Section 3.1 can be found here:

- Matlab file(s): Figure 4

2.5 Figure 5: $EC(h^t|R)$ for $\rho = 2$ and $\lambda = 5$

Figure 5 plots the expected cost of ASB for one-bin policies in the range $0 \leq R \leq N$.

Code for producing Figure 5 from Section 4.3 can be found here:

- Matlab file(s): Figure 5

2.6 Figure 6: Dynamics of quasi-attractors and the distribution of violations for the cost minimizing one-bin policy $R^* = 44$

Figure 2 plots $E(v^t|v^{t-1})$ and the associated q-attractors for the cost minimizing one-bin policy $R = 44$. The figure has two insets. The left inset plots the frequency distribution for $\boldsymbol{D}$. The right inset plots the autocorrelation coefficients for 1, 2, and 3 lags.

Code for producing Figure 6 from Section 4.3 can be found here:

- Matlab file(s): Figure 6

2.7 Figure 7: Extinguishing BA with a two-bin policy conditioned on v^{t-1}

Figure 7 has two panels: (i) Panel 1 plots $E(v^t|v^{t-1})$ and the q-attractor for a near optimal two bin policy; and, (ii) Panel 2 plots $E(v^t|v^{t-1})$ and the q-attractors for $R = 30$.Both panels

have two insets. In both cases the left inset plots the frequency distribution for $\boldsymbol{D}$; the right inset plots the autocorrelation coefficients for 1, 2, and 3 lags.

Code for producing Figure 7 from Section 4.3 can be found here:

- Matlab file(s): Figure 7

2.8 Table 1: Approximately optimal policies

Table 1 presents cost estimates for the (approximately) optimal multiple-bin policies.

Code for producing Table 1 from Section 4.3 can be found here:

- Matlab file(s): Table 1

3 Robustness of the model

In Section 2 of the paper, we asserted that "within limits, our results are robust to variations in the parameters and alternative versions of the model." In this section of the appendix we present some of the evidence that led us to this assertion.

We presented four salient features of the dynamic process in Section 3 of the paper. All four are apparent in Figures 1, 3 and 4 of the paper.

(i) **The cliff of expected violations:** *As R increases from 0 to N we observe three distinct regimes in the behavior of $E(v|R)$. Initially there is an unruly regime where $E(v|R)$ is relatively high and slowly decreasing in R, followed by a transitional regime in which $E(v|R)$ drops precipitously as R increases, and, finally, a compliant regime where $E(v|R)$ is relatively low and virtually constant in R. All three regimes exhibit random variation in the number of violations.*

(ii) **Q-attractors:** *The unruly regime is dominated by the bad attractor, BA, where the number of violations at its focal point is relatively high. As R increases violations fall modestly as BA's relative prominence declines. The compliant regime has a single good attractor, GA, where the number of violations at the focal point is relatively low and unresponsive to R.*

(iii) **The transition:** *Both BA and GA exist in the transitional regime. As R increases through the transitional regime, the persistence of BA diminishes from approximately* 1 *to* 0*, while the persistence of GA increases from approximately* 0.95 *to* 1*. The evolution of the relative persistence of BA and GA in the transitional regime leads to rapid change in the time spent in each.*

(iv) **Positive Autocorrelation:** *There is persistent positive autocorrelation in the time series of violations in all three regimes. The decay rate in autocorrelation in the unruly and compliant regimes is much higher than in the transitional regime producing a more significant lag structure than in the other two regimes. In the transitional regime positive autocorrelation may be significant out to* 50 *periods.*

In Section 3.1 of this appendix, we show that these features are robust to alternative ways of modeling key features of the dynamic process. In Section 3.2 we systematically explore robustness with respect to variation in the parameters of the model.

3.1 Robustness to alternative assumptions

In any period t, the choice of person i is based on a comparison of the net utility of violation, g_i^t, with the expected cost of violation, $q_i^t \cdot F$, where q_i^t is person i's subjective probability of apprehension. Person i chooses violate if and only if

$$g_i^t \geq q_i^t \cdot F.$$

We examine three robustness experiments. In the first two we examine alternative ways of modeling g_i^t and q_i^t. The information determining q_i^t is the recent history of violations and apprehensions, $((v^{t-z}, ..., v^{t-1}), (a^{t-z}, ..., a^{t-1}))$. This history is in part driven by the objective probabilities of apprehension, $(p^{t-z}, ..., p^{t-1})$. Our third experiment considers alternative ways of modeling p^t.

3.1.1 Person specific g-distributions

In the paper g_i^t is a random draw from a normal distribution, $\phi(g_i^t)$, with mean μ and standard deviation σ. In every period, every person's g_i^t is drawn from the the same PDF,

$\phi(g_i^t)$. Ex-ante, there are no differences over people in the net benefit of crime. Is this driving our results? We show it is not by demonstrating that a very different way of modeling net utilities produces a dynamic process that exhibits the four salient features above.

Assume g_i^t is a random draw from a person specific normal PDF with identical variances but different means. The N person specific means, $\{\mu_1, \mu_2, ..., \mu_N\}$, are chosen so that the population mean is approximately $\mu = 0.6$ and their standard deviation is approximately $\sigma = 0.008$. Specifically, the mean of person i's PDF is the value of g such that:

$$\Phi(g) = \frac{1}{2N} + \frac{(i-1)}{N} \quad \forall \quad i \in \{1, 2, ..., N\},$$

where N is the number of potential violators and $\Phi(g)$ is the cumulative distribution for $\phi(g)$. Now the ex ante net benefits of crime to people can be rank ordered. Person 1 has the lowest net benefit, person 2 the second lowest, and person N the highest.

In generating Figures 1 and 2, to ensure comparability with baseline results, we have used the baseline values for N, F, μ, σ, γ, z, α, and β.

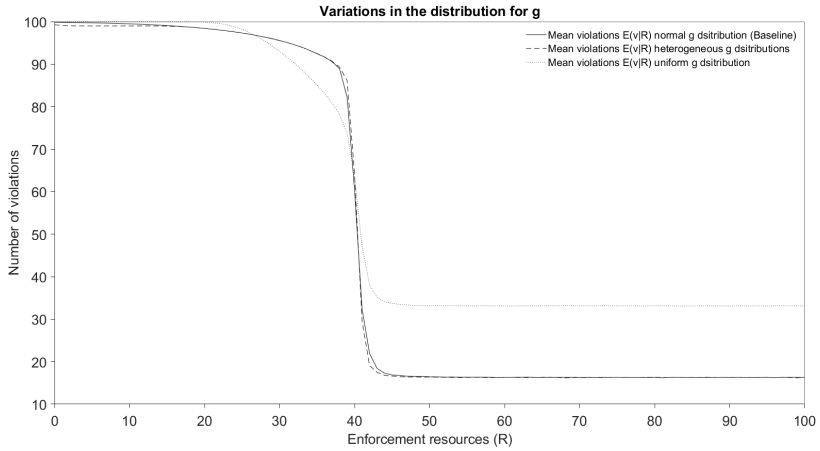

Figure 1: Cliff of expected violations for person specific g-distributions.

In Figure 1, the dashed line is $E(v|R)$ — the cliff of expected violations — for the variation in which g_i^t is a random draw from a person specific normal PDF. For comparison, the solid line is $E(v|R)$ for the baseline parameterization, the relationship we first saw in Figure 3.4. The two are virtually identical. We have also included the cliff of expected violations for a

variation of the basic model in which the g_i^ts are drawn from a rectangular distribution with support [0.2, 1.1]. The details are somewhat different, but the pattern is clearly the same.

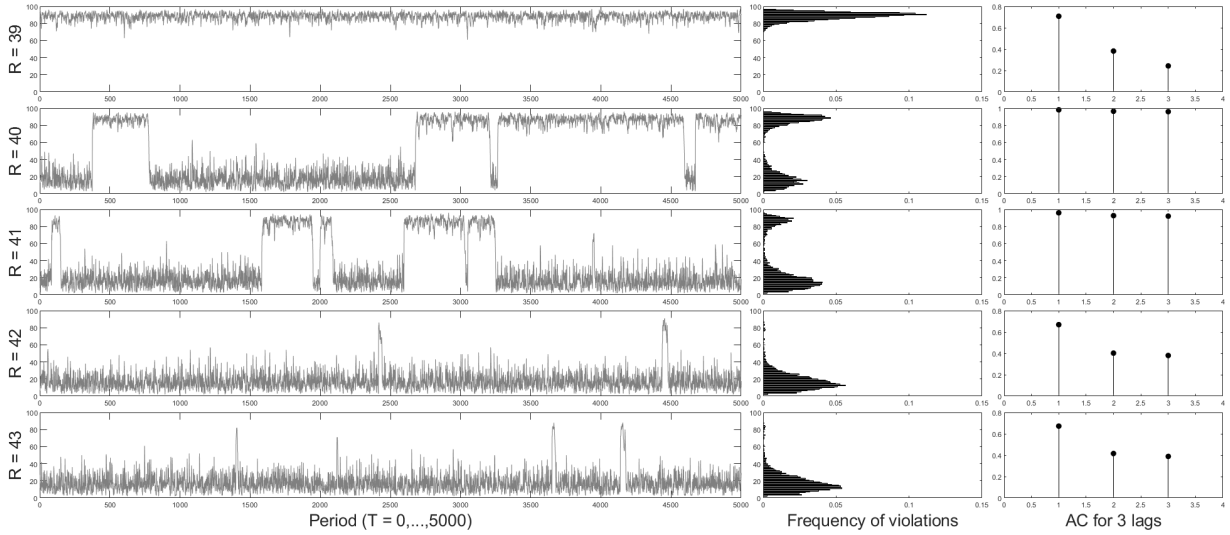

Figure 2: Simulated times series of violations for person specific g-distributions.

Figure 2 presents simulated violations data for five values of R spanning the transitional regime for the variation of the model in which the g_i^ts are drawn from person specific PDFs. The four salient features are apparent in this figure. There is considerable variation over time in violations. There are three regimes and, in all three, there is significant positive autocorrelation for lags of one, two and three periods.

The code that produces Figure 1 and Figure 2 can be found here:

- Matlab file: Figures 1 and 2

3.1.2 Person specific reference groups for subjective probabilities

In the baseline model it is assumed that q_i^t is determined by the history of violations and apprehensions in the z most recent periods, $((v^{t-z}, ..., v^{t-1}), (a^{t-z}, ..., a^{t-1}))$,

$$q_i^t = \frac{\alpha + a^{t-z} + ... + a^{t-1}}{\alpha + \beta + v^{t-z} + ... + v^{t-1}}, \ \forall \ i \in \{1, ..., N\}.$$

This requires that every person knows the complete history of violations and apprehen-

sions over the preceding z periods, and it ensures that they every person shares the same q. Is this driving our results? We show that the answer is no by examining the behavior of the model when every person knows only the history of a random sample of the entire population and uses that history to estimate their subjective probability of apprehension. More precisely, we assume people are organized on a *random network* with reference groups of size RG. Since each person's reference group includes themselves, they observe their own history as well as the $RG-1$ histories of people in their reference group. In each period, each person uses the recent history of their reference group to estimate an individual q_i^t. Clearly the subjective probability of apprehension for different reference groups may diverge, and there is a small probability that any two people from different reference groups will share precisely the same q.

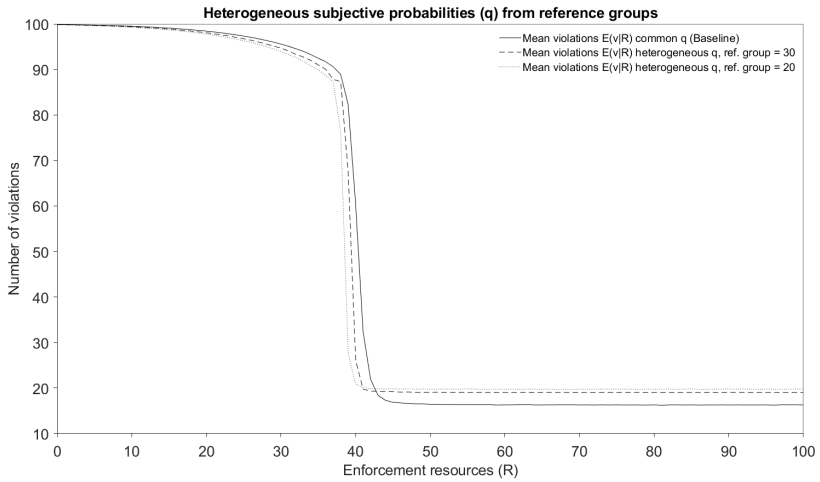

Figure 3: Cliff of expected violations for person specific reference groups.

In Figure 3, the solid line is the relationship $E(v|R)$ for the baseline parameterization of the model. The dashed and dotted lines are the relationships for the cases in which $RG = 30$ and $RG = 20$. The details are slightly different, but the first salient feature listed above is true of all three. As R ranges from 0 to 100 three regimes are apparent and as it transits the transitional regime there is a precipitous drop in $E(v|R)$.

Figure 4 presents simulated violations data for five values of R that span the transitional regime for the case in which $RG = 30$. Once again, the salient features we discovered in

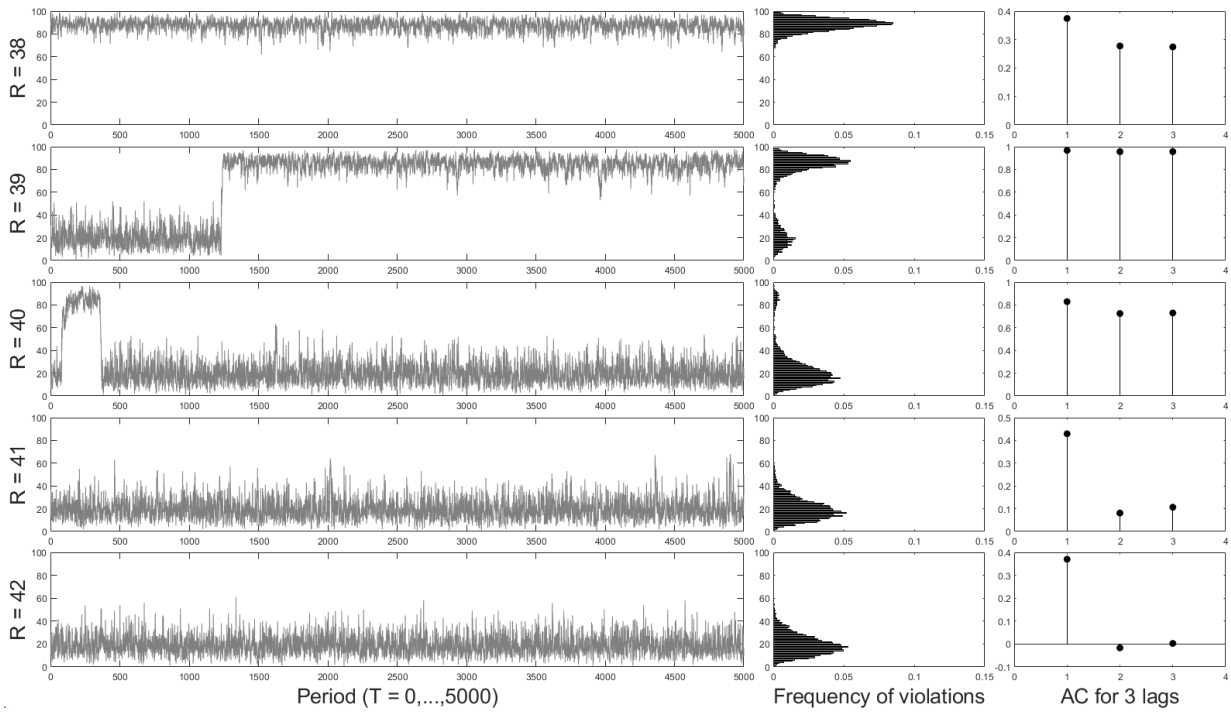

Figure 4: Simulated times series of violations for person specific reference groups.

Section 3 are characteristic of the data generated by this dynamic process.

The code that produces Figure 3 and Figure 4 can be found here:

- Matlab file: Figures 3 and 4

3.1.3 Different specification for the objective probability of apprehension

In the baseline model the objective probability that a violator will be apprehended in period t, p^t, is:

$$p^t = \gamma \cdot \min(1, \frac{R^t}{v^t}).$$

This formulation is based on three assumptions: (i) it takes one unit of enforcement resources, no more and no less, to investigate a violation; (ii) when a violation is investigated the violator is apprehended and punished with probability $0 < \gamma < 1$; and, (iii) when $v^t > R^t$, the probability that any one of the violations will be investigated is $\frac{R^t}{v^t}$ and when $v^t \leq R^t$ the probability is 1.

An alternative specification assumes that the probability that a violator is apprehended is a function of the quantity of deterrent resources devoted to investigating the violation $r = R/v$. Specifically we assume, $p = \gamma \cdot (1 - \frac{1}{\epsilon^r})$, where $\epsilon > 1$ and $0 < \gamma < 1$. This probability is increasing and concave in r and bounded below by 0 and above by γ. Given enforcement resources R^t and violations v^t, maximizing the expected number of apprehensions requires allocating $\frac{R^t}{v^t}$ to each violation. With this allocation, the probability that any violator is apprehended is,

$$p^t = \gamma \cdot (1 - \frac{1}{\epsilon^{R^t/v^t}}).$$

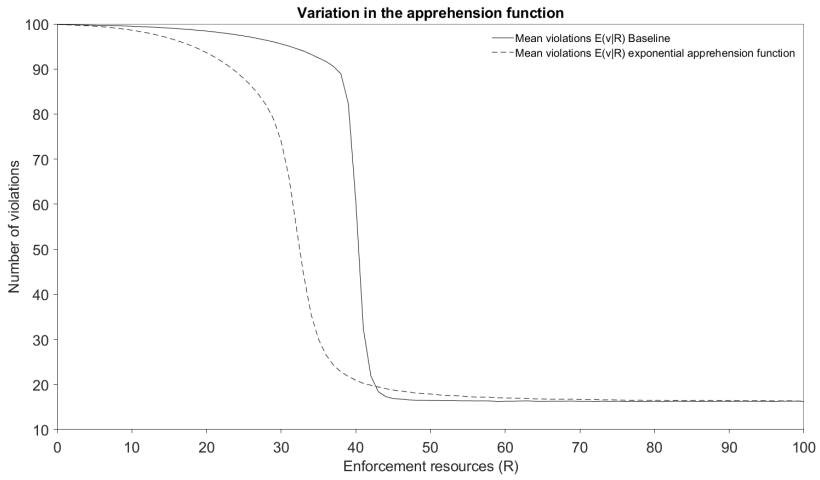

Figure 5: Cliff of expected violations for alternative apprehension function.

In Figure 5, the solid line is the relationship $E(v|R)$ for the baseline parameterization of the model and the dashed line is the relationships for the alternative specification of p^t. The details are different, but the first of the four prominent features we listed above is true of both. The transition from the unruly to the compliant regime is not as abrupt with the alternative specification.

Figure 6 presents simulated violations data for five values of R that span the transitional regime with the alternative specification of the apprehension function. The four salient features listed above are again apparent.

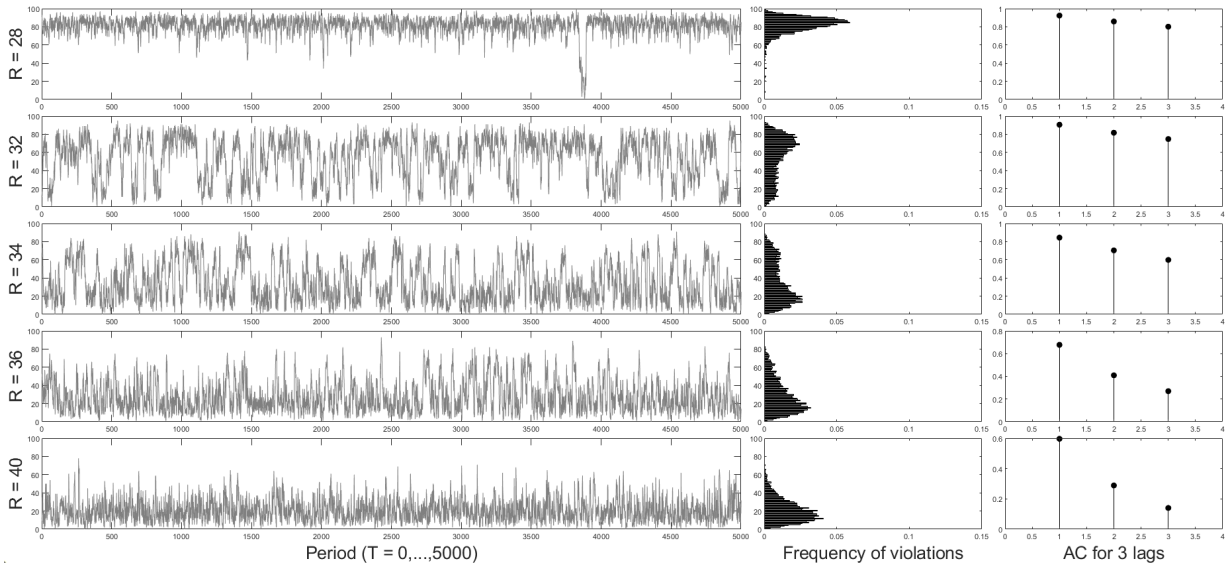

Figure 6: Simulated times series of violations for alternative apprehension function.

The code that produces Figure 5 and Figure 6 can be found here:

- Matlab file: Figures 5 and 6

3.2 Robustness to variation in parameters of the model

In examining robustness with respect to variation in the parameters of the model, for simplicity we focus on the first salient feature, the cliff of expected violations. We begin by examining the cliff over a range of alternative values of sanction, F, demonstrating the cliff's prevalence in a large space of possible sanctions. We then develop a measure of prominence or salience of the cliff that we call DROP and explore the the relationships between DROP and the parameters of the baseline model.

3.2.1 Variation in sanction, F

We consider one-bin policies that vary R and fix F. One important reason to treat F as a fixed parameter is motivated by what deterrence scholars term the *certainty principle*. Economists and criminologists have both empirically demonstrated that decisions to violate are much more sensitive to the perceived certainty of punishment than to the severity of

punishment. In fact, while the workhorse model treats F and p as perfect substitutes, the empirical literature finds that F and violations are only poorly correlated. The same literature has robustly demonstrated a negative correlation between violations and R. The finding for "the greater salience of sanction certainty is so common, in fact, that it has achieved recognition as the 'certainty principle'" (Apel, 2013, p. 73). There are other reasons to focus on R rather than F. Large fines may be morally unacceptable, hard to enforce (they create agency problems in enforcement), and violators may be unable to pay them.

Here we consider one-bin policies in which both R and F vary. We restrict our attention to policies (R, F) for which $R \in \{0, ..., N\}$ and $F \in \{0, 0.05, ..., 1.45, 1.5\}$. This results in a set of such polices with $101 \cdot 31 = 3131$ elements. Figure 7 is a surface plot of $E(v|R, F)$ for this set of polices. Two cross-sections are highlighted. The first is produced by holding the baseline sanction constant and varying R. This reproduces the "cliff of expected violations" detailed above. The second, producing another such cliff, is produced by holding enforcement spending constant and varying F.

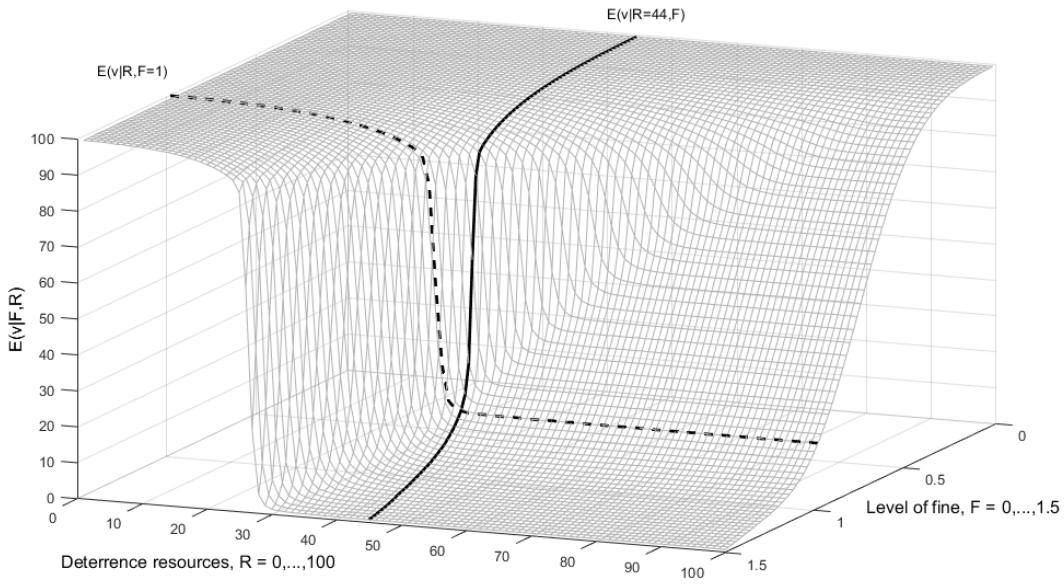

Figure 7: The space of one-bin policies.

The code that produces Figure 7 can be found here:

- Matlab file(s): Figure 7

Both cross-sections, as well as the surface itself, reveal the same regimes we detail above: (i) the unruly regime consists of a plateau around the top edge of the figure; (ii) the compliant regime is the rectangular plateau at the bottom; and, (iii) the transitional regime is a cliff separating the two.

3.2.2 Defining DROP and CLIFFSET

The geometric signature of the cliff is a small range of contiguous values of R such that as R is increased through that range there is a dramatic drop in $E(v|R)$. We fix the range of R at $RANGE = INT(0.05 \cdot N)$, that is $RANGE$ is the largest integer in $\frac{N}{20}$. For the baseline parameterization $RANGE = 5$.

For a given parameterization we first estimate the relationship $E(v|R)$ for all $R \in \{0, ..., N\}$. For every integer $X \in \{0, ..., N - RANGE)$ we then calculate the reduction in $E(v|R)$ when R is increased from X to $X + RANGE$. $DROP$ is the maximum value of this difference normalized as a fraction of its maximum possible value, N.

$$DROP = \frac{\max_{\{X\}}[E(v|R = X) - E(v|R = X + RANGE)]}{N}$$

Intuitively, $DROP$ measures the height of the cliff relative to the size of the population of potential violators. It is bounded below by 0 and above by 1. For example, $DROP = 0.5$ indicates that in the vicinity of the cliff an increase of $INT(0.05 \cdot N)$ in R reduces the expected number of violations in the stationary distribution by $0.5 \cdot N$, or that the slope of the cliff is approximately -10. In defining $CLIFFSET$, we set the threshold value of $DROP$ to 0.35, producing a slope of -7. The value of $DROP$ for the baseline parameterization is roughly 0.7, producing a cliff with a slope of -14.

Figure 8 plots three different cliffs: one for the baseline parameterization, one in which $DROP = 0.35$, and one in which $DROP = 0.1$. The baseline produces the most prominent cliff and the largest value of DROP.

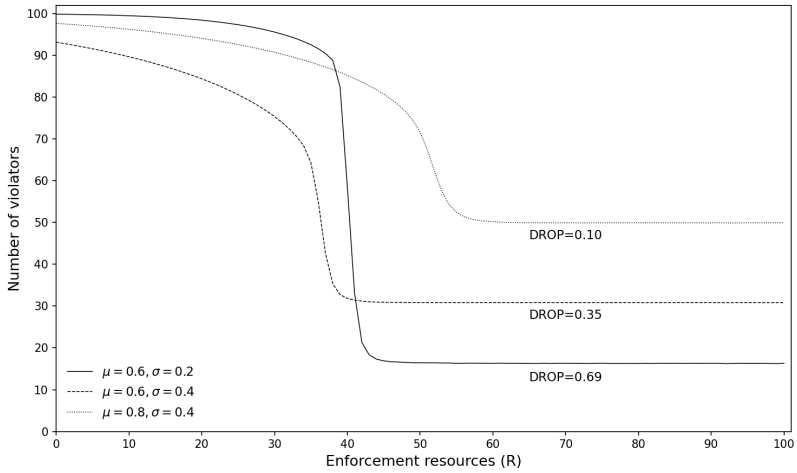

Figure 8: Three cliffs of expected violations and their measured $DROP$.

The code that produces Figure 8 can be found here:

- Matlab file(s): Figure 3

3.2.3 Characterizing DROP and CLIFFSET

$DROP$ is a function of six parameters, $\mu, \sigma, \gamma, F, N$ and z, so $CLIFFSET$ is an object in a six dimensional space.

$$CLIFFSET = \{\mu, \sigma, \gamma, F, N, z | DROP \geq 0.35\}$$

While a complete description is not feasible, we can describe some of the most important aspects of $CLIFFSET$. Since the underlying behavior of the model depends only on the value of $\gamma \cdot F$, and not independently on γ or F, we are able to reduce $CLIFFSET$'s dimensionality by treating $\gamma \cdot F$ as a composite parameter. Think of $\gamma \cdot F$ as the size of the

deterrence stick. We start by describing a subset of $CLIFFSET$,

$$CLIFFSET' = \{\mu, \sigma \mid DROP \geq 0.35,\ \gamma \cdot F = 0.80,\ N = 100,\ z = 2\}$$

In Figure 9, we have constructed a surface plot of $DROP$ for $\mu \in \{0, 0.05, 0.1, ..., 1.00\}$ and $\sigma \in \{0.15, 0.20, ..., 0.55\}$, so the set of (μ, σ) pairs has 189 elements.We have not included estimates of DROP for values of σ less than 0.10 because our convergence simulation routine is not reliable for small values of $\frac{\sigma}{\gamma \cdot F}$. We have included a shaded reference plane representing our threshold value of $DROP$. The outer boundary of $CLIFFSET'$ is defined by the locus of points where the surface plot and the reference plane intersect. We have projected the iso-DROP locus for $DROP = 0.35$ onto the baseplane, and $CLIFFSET'$ is the set of (μ, σ) pairs on and inside the iso-DROP locus.

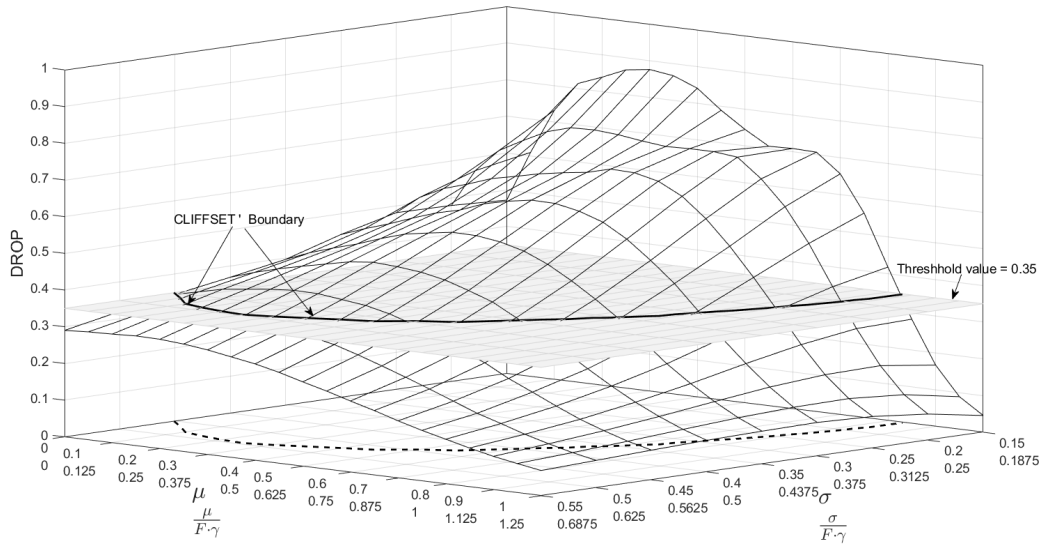

Figure 9: $DROP$ for a range of μ and σ.

The code that produces Figure 9 can be found here:

- Matlab file(s): Figure 9

There is another interpretation of $CLIFFSET'$. Setting $s > 0$ and applying this scalar to three key parameters of $CLIFFSET'$, $(s \cdot \mu, s \cdot \sigma, s \cdot \gamma \cdot F)$, the ratio of the new set of

values for $DROP'$ and the original $DROP$ values is, to a very close approximation, equal to 1. In other words, $DROP$ is homogeneous of degree 0 in μ, σ and $\gamma \cdot F$. This follows from the fact that the probability π that any person chooses violate in any period is homogeneous of degree 0 in these same parameters. In every period, every person has a well defined subjective probability of apprehension, q. When their opportunity g is revealed they choose violate if and only if $g > q \cdot \gamma \cdot F$. Prior to the realization of g, the probability π that they choose violate is $1 - \Phi(q \cdot \gamma \cdot F)$. Since $\phi(g)$ is normal, $\Phi(q \cdot \gamma \cdot F)$ and hence π, are themselves homogeneous of degree 0 in μ, σ and $\gamma \cdot F$.

A simple rescaling of the μ and σ axes and a redefinition of the variables the axes represent allows us to reinterpret Figure 9. The surface plot in the figure represents the DROP statistics for pairs $(\frac{\mu}{\gamma \cdot F}, \frac{\sigma}{\gamma \cdot F})$, and in the baseplane the set identified by the dashed line and the axes is $CLIFFSET'$ for any parameterization in which $N = 100$ and $z = 2$.

$DROP$ is increasing in both N and z, and hence for larger values of either or both of these parameters the associated subset of $CLIFFSET$ is somewhat larger than the one identified in Figure 9. To give a sense of how sensitive the subset is to these variables we have calculated DROP statistics for $z = 3$ and for $N = 160$ for three points on the outer boundary of $CLIFFSET'$.

Table 2: DROP statistic for N=160 and z=3.

Threshold tuples $(\frac{\mu}{\gamma \cdot \mathbf{F}}, \frac{\sigma}{\gamma \cdot \mathbf{F}})$	**DROP**$(\mathbf{N = 160})$	**DROP**$(\mathbf{z = 3})$
(0.9125, 0.4088)	0.3781	0.3650
(0.9981, 0.2993)	0.3870	0.3763
(0.7825, 0.4856)	0.3968	0.3711

The $DROP$ statistic is larger than 0.35 for all three of the points when $N = 160$ and for all three when $z = 3$, indicating that the manifold in Figure 9 shifts up for an increase in either N or z. When the manifold shifts up $CLIFFSET'$ increases.

4 Robustness of policy analysis

The estimates reported in this section are generated using simpler versions of the directed search algorithms discussed in Section 1. The number of randomly seeded searches was reduced to 50 in each case implying that the errors in the estimation of the optimal policy will be slightly higher. But, as we have observed, the space in which the optimum occurs is very flat and so there is very little difference between estimates in the neighborhood of the optimum. Reducing the number of searches increased computational efficiency at nearly zero cost to precision.

We examine the robustness of three salient features discussed in Section 4 of the paper using our procedure. The salient features are:

(i) **Productive reserve capacity:** *The reserve capacity rate associated with the optimal one-bin enforcement policy, and the approximately optimal two- and three-bin policies are at least* 0.40*, meaning that in a typical period more than 40% of available enforcement capacity is not deployed investigating a violation.*

(ii) **The efficacy of dynamic policy:** *Simple dynamic policies dramatically outperform static policies. Replacing the optimal one-bin enforcement policy with the optimal two-bin policy reduces the expected cost of ASB by 12% to 21%.*

(iii) **Diminishing returns to complexity:** *Simple dynamic policies do roughly as well as more complex dynamic policies. Replacing the optimal two-bin policy policy with the optimal three-bin policy reduces the expected cost of ASB by less than 1%.*

We focus on multi-bin policies that are conditioned on v^{t-1}. We fix N and z: $N = 100$ and $z = 2$. In Figures 10 and 11 we present evidence on the robustness of these three features as two composite parameters, $DROP$ and RAT, vary. RAT is eqaul to $= (\lambda - 1)/\rho$ – it is the net external cost that violations impose on society, $(\lambda - 1)$, relative to the price of one unit of the enforcement resource, ρ. $DROP$ is an index of the prominence of the cliff of expected violations that we introduced in Section 2 of Appendix 1. Having fixed N and z, $DROP$ is determined by μ, σ, γ, and F. Think of $DROP$ as a composite parameter that captures the effects of variation in these four parameters.

In Panels 1 and 3 of Figure 10, we examine the relationship between the reserve capacity

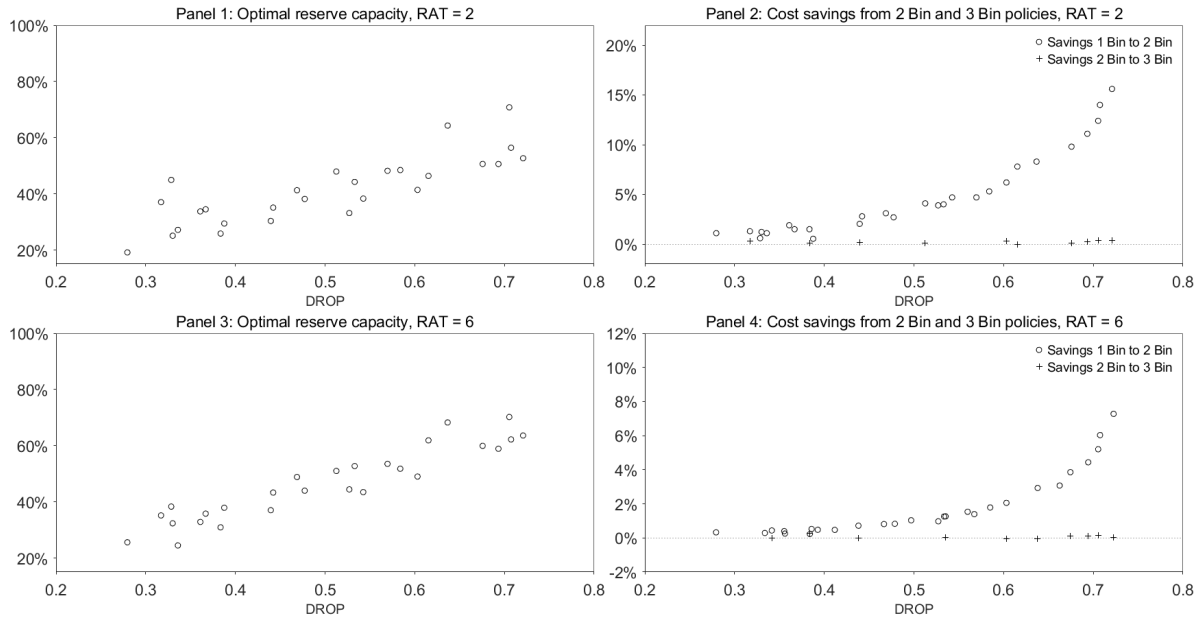

Figure 10: Reserve capacity and cost savings as a function of $DROP$.

associated with the optimal Two-Bin policy and $DROP$ for two values of RAT. In Panel 1, RAT is 2 and in Panel 3 it is 6. Reserve capacity tends to increase as DROP increases, and it is significant over the entire range of $DROP$ values. At its lowest it is roughly 20% and at its highest 70%. In Panel 1 of Figure 11 we examine the relationship between reserve capacity associated with the optimal Two-Bin policy and RAT, for three values of $DROP$. For all three values, there is massive discontinuity in the relationship between reserve capacity and RAT. (Given the coarse grid of RAT values, we have not identified any of the three discontinuities precisely.) When RAT is very small the best enforcement policy is to do nothing. As RAT passes through the point of discontinuity the reserve capacity rate jumps up, and beyond the discontinuity it increases slowly with RAT. For the two larger $DROP$ values to the right of the discontinuity the rate exceeds 60% and for the lowest value it is about 20%.

In Panels 2 and 4 of Figure 10, we examine the relationships between $DROP$ and the cost savings that arise: (i) when we replace the optimal two-bin policy with the optimal one-bin policy; and, (ii) when we replace the optimal two-bin policy with the optimal three-bin policy. In Panel 3, RAT is 2 and in Panel 4 it is 6. The "o"s indicate savings for replacing

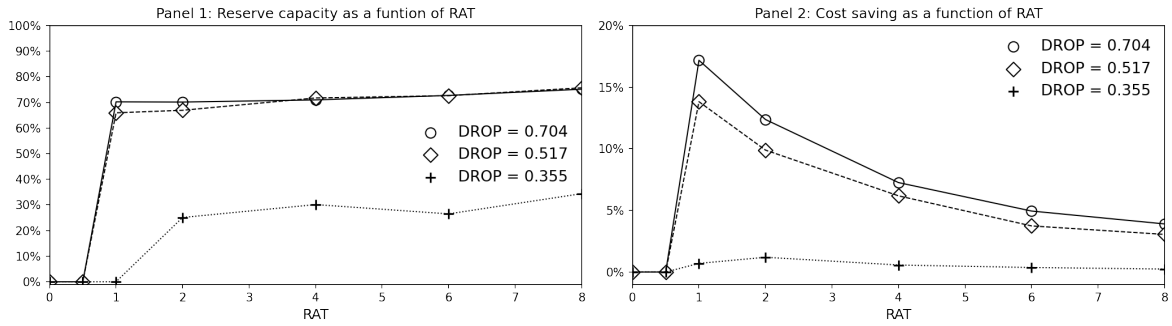

Figure 11: Reserve capacity and cost savings as a function of RAT.

one-bin with two-bin policies; the "+"s indicate savings for replacing two-bin with three-bin policies. In the first instance the cost savings are significantly positive and increasing in $DROP$. In contrast, further complexity results in savings close to 0. In Panel 2 of Figure 11 we examine the relationship between RAT and the cost savings associated with replacing the optimal one-bin enforcement policy with the optimal two-bin policy, for three values of $DROP$. For all three there is massive discontinuity in the relationship. When RAT is small doing nothing is the optimal enforcement policy regardless of whether we are using a one-bin or multi-bin enforcement policy, so the cost reduction is 0. Beyond the discontinuity, for the two larger values of $DROP$ the cost saving is significant and inversely related to RAT. For the smallest value of $DROP$, the cost saving is insignificant.

The results reported in Figures 10 and 11 are based on parameterizations in which $N = 100$ and $z = 2$. Because simulations with larger values of these parameters are very time intensive, we have run just a few of them. Based on those simulations it seems clear that the patterns seen in these figures and the magnitudes of the numbers are robust to increases in the values of the parameters. Variations of $N = 160$ and $z = 3$ were run on two values each of $DROP \approx \{0.7, 0.4\}$ and $RAT = \{2, 10\}$. The pattern of cost savings for two-bin and three-bin policies was preserved. Costs savings for two-bin over one-bin policies increased slightly for both the larger N and z values, with the increases ranging from 0.1% to 1.8% over the baseline values with the maximum cost savings being 15.2%. The cost savings in moving from two- to three-bin policies was again negligible with the total savings ranging from 0%

to 0.4%. The reserve capacity values were almost identical for each for each of the four cases. For example, for $DROP \approx 0.7$ and $RAT = 2$, reserve capacity $= (0.4260, 0.4240, 0.4302)$ for the (N, z) tuples of $(100, 2)$, $(160, 2)$, and $(100, 3)$ respectively.

The code that produces Figure 10 can be found here:

- Matlab file(s): Figure 10

The code that produces Figure 11 can be found here:

- Matlab file(s): Figure 11