

1

While Norbert Marot was sitting on his porch one day in 2012, his wife picked up a rifle and shot him three times in the back, killing him. Jacqueline Sauvage thereby ended 47 years of physical, emotional, and sexual abuse at Marot's hands. Not only had Marot abused Sauvage; he had raped their three daughters. Probably, he had also abused the couple's son, who had committed suicide the day before Sauvage killed Marot. Although Sauvage had been abused by Marot, French law provides that self-defence is permitted only against an imminent or immediate threat. Because Marot was sitting with his back turned to Sauvage, he posed no immediate threat to Sauvage. Sauvage was accordingly convicted of murder and sentenced to 10 years in prison.

The result was widely criticised. How could it be just, many wondered, to impose additional suffering on someone who had already suffered so much? A petition for a presidential pardon for Sauvage obtained over 400,000 signatures. President Hollande, who had previously opposed the use of the pardon power, eventually relented, and pardoned Sauvage. She was released having served only three years of her sentence.

Also in 2012, Bernadette Dimet shot and killed her husband. The two cases were very similar. Dimet, like Sauvage, had been abused by her husband for many years. Dimet, like Sauvage, was not in immediate physical danger from her husband when she shot him. But Dimet, unlike Sauvage, was not granted a pardon. There were some differences between the cases. Dimet meant to hurt but not to kill her husband, for example, which is why she was given a 5 year sentence, rather than a 10 year sentence. But if these differences favour granting a pardon in one of these cases but not the other, they would seem to favour pardoning Dimet over Sauvage.¹

It is hard not to feel troubled by these two cases. If it was just to pardon Sauvage, it would seem unjust not to pardon Dimet. If it was just not to pardon Dimet, then it would seem unjust to pardon Sauvage. It appears that there was a failure of both non-comparative and comparative justice. Doubts

*Associate Professor, Faculty of Law, University of Oxford. For comments I thank Max Harris, Swati Jhaveri, Mattias Kumm, and Nicola Lacey. I presented drafts of this article at the LSE, NUS, and Oxford, and I thank the audience members.

¹My account of these cases is drawn from Kate Fitz-Gibbon and Marion Vannier, 'Domestic Violence and the Gendered Law of Self-Defence in France: The Case of Jacqueline Sauvage' (2017) 25 *Feminist Legal Studies* 313.

like these lead naturally to doubts about the power under which pardons can be granted. Why give anyone the power to set aside a punishment which is just, according to law? Why give that power to someone, like President Hollande, who has no judicial experience and does not follow judicial procedures? Why make the power arbitrary, such that it hinges on one person's will and whims?

I aim to answer these questions in the form of a new argument for pardon powers. Here is how I proceed. In section 2, I define a pardon power. So defined, pardon powers face three longstanding objections, which I set out in section 3. Section 4 introduces some basic distinctions, which I use to locate the standard justification for pardon powers and my preferred alternative. Section 5 describes the standard justification, according to which pardons are a vehicle of Aristotelian equity. I show that this argument fails comprehensively. It does not justify any of a pardon power's distinctive features or overcome any of the traditional justifications to pardon powers. Section 6 sets out a new justification for pardon powers, according to which pardons are a means of deviating from optimal rules when they require suboptimal results. I apply my argument in section 7, to show why Hollande might have had reason to pardon Sauvage but not Dimet. Section 8 concludes.

2

A pardon is an exercise of a power to pardon. A power to pardon, as I use the phrase, is any legal power with these features:

Result. The power's use cancels a person's legal liability to a criminal sanction.

Mechanism. The power's use does not alter any legal rule.

Power-holder. The power is possessed by a non-judicial official.

Control. The power's use is not constrained by any legal rule.

The first feature distinguishes a pardon from an exercise of prosecutorial discretion. Prosecutors may have the freedom not to ask a court to make a finding of legal liability, but liability is not thereby cancelled. The second feature distinguishes a pardon power from a legislature's power to cancel legal liability by abolishing or otherwise altering a legal rule imposing criminal liability. Had French lawmakers narrowed the offence of murder with retroactive effect, for example, they might have cancelled Sauvage's legal liability. They would still not have pardoned her. The third feature distinguishes a pardon power from an appellate court's power to reverse the initial imposition of a sanction at trial. The final feature distinguishes a pardon power from nearly every other legal power in a modern legal system.

The power to pardon granted under the French Constitution has all of these features. Article 17 states: ‘The President of the Republic is vested with the power to grant individual pardons.’ This power is used to cancel an individual’s liability for criminal offences. Using the power does not amend or abolish the law under which a person is liable. Hollande’s pardon of Sauvage, for instance, cancelled her liability to imprisonment, without altering the law of murder or of self-defence. The power is, of course, held by the President. The power is also unchecked. The reasons for a pardon decision ‘are not communicated’, and the decision itself is ‘not subject to judicial control with regard to its legality or constitutionality’². Within French law, a pardon is an unreviewable ‘act of conscience’³.

The French pardon power is far from unique. In the United States, for example, Article II, Section 2, Clause 1 of the Constitution provides the President with ‘the power to grant reprieves and pardons for offenses against the United States, except in cases of impeachment’. There are explicit limits: the President may not grant a pardon for state offences, or in cases of impeachment. Within these bounds, though, the President’s power to pardon is ‘granted without limit’⁴. Its exercise is therefore an ‘act of grace’⁵. In the United Kingdom, in law the monarch (but by convention the Home Secretary) has a prerogative power of mercy, under which pardons are granted. Traditionally, this was a power which was not subject to any external constraint. ‘The law would not inquire into the manner in which ... [the] prerogative [of mercy] was exercised’⁶. Other states have broadly similar powers.⁷

3

So characterised, pardon powers are open to several longstanding objections.

Start with the first two features of pardon powers: that the exercise of the power cancels a person’s legal liability to a criminal sanction, without altering any legal rule. The worry is that the combination of these features

² R Levy, ‘Pardons and Amnesties as Policy Instruments in Contemporary France’ (2007) 36 *Crime & Justice* 551, 557.

³ M-H Renaut, ‘Le droit de grâce doit-il disparaître?’ (1996) 3 *Revue de Science Criminelle* 575, 598

⁴ *Ex parte Garland*, 71 U.S. 333, 336 (1866).

⁵ *United States v Wilson*, 2 U.S. (7 Pet.) 150 (1883).

⁶ *Hanratty v Lord Butler of Saffron Walden* (1971) 115 SJ 386. Judicial review is available if the Home Secretary misunderstood the scope of the prerogative of mercy: *R v Secretary of State for the Home Department, ex p Bentley* [1994] QB 349.

⁷ See generally A Novak, *Executive Clemency in Global Perspective* (Routledge 2016) ch 8.

gives rise to a dilemma. Consider Cesare Beccaria's scepticism about pardon powers:

[The power to pardon] is one of the noblest prerogatives of the throne, but, at the same time, a tacit disapprobation of the laws. Clemency is a virtue which belongs to the legislator, and not to the executor of the laws, a virtue which ought to shine in the code, and not in private judgment. ... Let, then, the executors of the laws be inexorable, but let the legislator be tender, indulgent and humane.³

Suppose that the legislator has been appropriately 'tender, indulgent and humane' when making the code, and that it nonetheless chose to impose liability in a class of case. It would seem that the code ought not to be set aside. Instead, *pace* Beccaria, its application ought to be 'inexorable'. Conversely – and here I am going beyond Beccaria – suppose the legislator has *not* been appropriately 'tender, indulgent and humane'. It has imposed liability in a class of case which it ought not to have. Then the code ought to be changed, not merely set aside in a single case of this class but retained for future cases of that class.

With some license, we can attribute the following objection to Beccaria:

Beccaria's objection. A pardon power is justified only if it could be true that (1) a legal rule ought to be set aside, but (2) that legal rule ought not to be changed. But, if (1) is true, then (2) is not; and if (2) is true, then (1) is not. So, a pardon power is not justified.

To see the force of this objection, consider Sauvage's case again. If French lawmakers were right to impose an imminence requirement on self-defence, then it was wrong for Hollande to set that requirement aside in Sauvage's case. If lawmakers were wrong to impose an imminence requirement, then it ought to be set aside altogether, not simply in Sauvage's case. Whether lawmakers were right or wrong, it seems that there is no call for a pardon, which sets aside the law without changing it.

Now turn to the second feature of pardon powers: that they are allocated to non-judicial officials or institutions. We tend to think that judicial functions ought to be allocated to the judiciary. Judicial functions include the retrospective resolution of specific disputes, especially with contested legal or factual issues. The judiciary is characterised by small panels of decision-makers, legal expertise, independence from the parties, highly developed rules for fact finding, procedural protections including the right to be heard, and so on. We match judicial functions for reason of what constitutional scholars term 'efficiency', meaning that the judiciary is well-suited to

³ C Beccaria, *An Essay on Crimes and Punishments* (James Donaldson 1788) 167-8.

exercise judicial functions.⁹ Think of criminal matters. The resolution of a criminal dispute does not require the open-ended deliberative capacities or large membership of a legislature. What it demands is careful fact-finding, attention to the legal issues, and a fair hearing, all of which the judiciary can provide. True, there are also executive tribunals that can provide these things. But such tribunals are not independent of the prosecution in a criminal proceeding.

To grant or refuse a pardon is to make a decision about how to treat a specific case, on a disputed matter, in a criminal proceeding. This is a judicial function, which is nonetheless granted to a non-judicial entity. As James Fitzjames Stephen says:

[U]nder ... [a pardon power] a function which is really judicial is discharged by an irregular, irresponsible and secret tribunal consisting of a single statesman who has no special acquaintance with law and no judicial experience, who can neither examine witnesses nor administer oaths¹⁰

Hollande, for example, was empowered to set aside Sauvage and Dimet's liability to punishment, even though he had little legal training, did not hear from witnesses, and did not follow any of the trial procedures designed to promote truth or fairness.

We can put the objection this way:

The separation of powers objection. A pardon power grants a non-judicial body the power to cancel legal liability in particular criminal cases. But this type of power is justified only if it is held by a judicial body. So, a pardon power is unjustified.

The objection is compounded by an apparent inconsistency. Judges (and juries) are given responsibility at trial for imposing a punishment. But the final say is given to someone else, with a very different set of capacities, resources, and skills. If the French Cour d' Assises is a suitable body to decide Sauvage's liability, how could the President, with such a different institutional profile, *also* be suitable?

Finally, consider the most obviously worrying aspect of a pardon power: that it is unconstrained by law. The rule of law demands that government act according to law, that is, according to general legal norms, announced in advance. The rule of law is not opposed to discretionary powers. But these

⁹ See N Barber, 'Prelude to the Separation of Powers' (2001) 60 CLJ 59, 64-6 (tracing this rationale to Locke and the authors of *The Federalist Papers*).

¹⁰ JF Stephen, *Selected Writings A General View of the Criminal Law* (KJM Smith, ed. OUP 2014) 217-8.

powers need to be constrained by general standards, for example, standards of reasonableness and fairness. These standards must also be enforceable by a ‘legal machinery’, as Joseph Raz says, one which is capable of ‘supervising conformity to the rule of law and provid[ing] effective remedies’¹¹. The most important piece of this machinery is an independent judiciary with reviewing powers.

Pardon powers seem to threaten the rule of law in two respects.¹² First, they are not subject to any legal constraints. Second, and *a fortiori*, they are not subject to legal constraints backed by judicial enforcement mechanisms. Let us formulate the objection as follows:

The rule of law objection. A pardon power grants an unconstrained discretionary power. But discretionary power is justified only if it is constrained by general, externally enforceable legal norms. So, a pardon power is unjustified.

Suppose that Dimet, her legal appeals exhausted, turns to Hollande. She asks for a pardon. The President’s power is now at its most expansive. He can do as he likes. And she is at the President’s mercy. The law does not rule, in this scenario; the President does. The pardon power is thus a form of ‘legally sanctioned arbitrariness – ‘lawful lawlessness’¹³, in Austin Sarat and Nadar Hussain’s words.

To clarify, pardon decisions are reviewable, in the sense that judges are willing to determine whether an act falls within the scope of a power. They might ask whether a person who purported to grant a pardon was in fact empowered to do so, for example.¹⁴ What judges do not do is review acts that fall within the scope of the power. As Hugo Adam Bedau says, pardon decisions are ‘standardless in procedure, discretionary in exercise, and unreviewable in result.’¹⁵

4

¹¹ J Raz, *The Authority of Law* (OUP 1979) 218.

¹² For claims that uncontrolled pardon powers threaten the rule of law, see, eg, BV Harris, ‘Judicial Review, Justiciability and the Prerogative of Mercy’ (2003) 62 CLJ 631; D Markel, ‘Against Mercy’ (2004) 88 Minnesota Law Review 1421, 1428; A Sarat, *Mercy on Trial* (Princeton University Press 2005) 90-92, 160.

¹³ A Sarat and N Hussain, ‘On Lawful Lawlessness: George Ryan, Executive Clemency, and the Rhetoric of Sparing Life’ (2004) 56 Stanford Law Review 1307.

¹⁴ eg, *ex parte Crump*, 10 Okla. Crim. 133 (Okla. Crim. App. 1913) (determining whether during a vacancy in the office of Governor the Lieutenant Governor was empowered to grant a pardon).

¹⁵ HA Bedau, ‘The Decline of Clemency in Capital Cases’ (1990) 8 New York University Review of Law and Social Change 255, 257.

In light of these objections, what if anything can be said for a pardon power? I take up this question squarely in the next two sections. In this section I lay the groundwork by setting out some terminology and distinctions.

Let me distinguish two types of results in particular cases:

Suboptimal result. A result in a case is suboptimal iff there is an alternative result which there would be greater reason to achieve, absent the law.

Optimal result. A result in a case is optimal iff it is not suboptimal.

No one needs a pardon power to overturn optimal results; what is unobjectionable does not need correction. So, if pardon powers are justified, it is as a means of correcting for suboptimal results.¹⁶

A result can be required, prohibited, or permitted by the rules of a legal system. Historically, pardons were often granted to correct for suboptimal results which were also legal errors (e.g., wrongful convictions, judicial mistakes). However, there are now other institutions clearly better suited to the correction of this sort of suboptimal result (e.g., appellate courts). Accordingly, I will focus, as Beccaria does, on arguments for pardon powers which cast them as a means of correcting for suboptimal results which are required by law.

Let us call the set of rules of a legal system its *code*. In parallel with the definitions above, there are two types of codes:

Suboptimal code. A legal code is suboptimal iff there is an alternative code which there is greater reason to adopt.

Optimal code. A legal code is optimal if and only if it is not suboptimal.

It is obvious that suboptimal results may be required by suboptimal codes. Later, I will explain that suboptimal results may also be required by optimal codes. So let me introduce a final pair of definitions:

Closure case. A case is a closure case iff it is a case in which a suboptimal result is required by a suboptimal code.

Gap case. A case is a gap case iff it is a case in which a suboptimal result is required by an optimal code.

The idea is that in gap cases there is a divergence or difference between what is required by the code we should have and what we should (absent the code)

¹⁶ On the definition of suboptimal and optimal results, see J Brand-Ballard, *Limits of Legality* (OUP 2010) 75.

do. In closure cases that difference does not exist, because it is neither true that we should have the code we do nor that we should (absent the code) do as it requires us to do.¹⁷

Drawing things together, if pardon powers are justified, it is as a means of correcting for closure cases or gap cases. Ultimately, I will argue that the best defence of pardon powers casts them as a device for correcting for gap cases. First, though, I want to consider whether pardon powers can be defended as a means of addressing closure cases.

5

Closure cases arise only when a code is suboptimal. Suboptimal codes come in two varieties. Some codes are *culpably suboptimal*: lawmakers erred when they made the code and are blameworthy for doing so. Pardon powers are not the right device to correct for these suboptimal results. Most obviously, that is because pardon powers are often held by the same body which made the code (e.g., the legislature), and one body cannot both hold lawmaking power and check that power. Other codes are *excusably suboptimal*: the code is not as good as it could be, but lawmakers are not to blame for its shortcomings. This second possibility takes us to the oldest and most common argument for the pardon power: the argument from equity.

A

When they make a code, lawmakers make a set of prospective rules. Because these rules are prospective, they must be based on reasons which lawmakers can foresee. Lawmakers cannot know every twist and turn of the cases to which the code will apply, and inevitably the code will require some suboptimal results. A code based on perfect knowledge would not suffer from these flaws, but legislators are not to blame for their code's suboptimality. As Aristotle says, 'practical affairs [are] of this kind from the start'¹⁸. Closure cases are practically unavoidable.¹⁹

How should officials respond in such cases? Aristotle distinguishes two types of justice. There is justice according to positive law and absolute justice. Because 'about some things it is not possible to make a universal statement which shall be correct', legal justice sometimes falls short of absolute justice. When 'the legislator fails us and has erred by over-simplicity', we should

¹⁷ For a similar usage, see L Alexander, 'The Gap' (1991) 14 Harvard Journal of Law and Public Policy 695.

¹⁸ Aristotle, *The Nicomachean Ethics* (David Ross trans., OUP 2009) book V, ch 10.

¹⁹ The point is a common one. See eg F Schauer, *Playing by the Rules* (OUP 1991), ch 2; L Alexander and E Sherwin, *The Rule of Rules* (Duke University Press 2001) ch 4.

‘correct the omission’ and ‘say what the legislator would have said had he been present, and would have put into his law if he had known’ the facts of the case before us.²⁰ In this way, we will accomplish a form of ‘equity’²¹.

A pardon power is particularistic and retrospective. It can be used in light of all relevant considerations in a case, including ones not available to a legislature. Thus, it can be used to correct for some suboptimal results required by suboptimal codes. And this is indeed the purpose to which pardon powers were traditionally put. Consider the history of pardons in England. Through the 14th century, pardons in homicide cases were normally granted because of circumstances relevant to the offender’s culpability.²² Between the 16th and 18th centuries, the death penalty was vastly expanded, eventually attaching to more than 300 offences. Again, pardons were granted for reasons we would find familiar: given the nature of the offence and the character of the offender, death was considered too severe a penalty.²³ Pardons were used to fit punishments to crimes, and so to do justice in particular cases. No wonder that Blackstone characterised the English power to pardon as a ‘court of equity’ in the monarch’s own breast.²⁴

B

The argument from equity is, as I said, the traditional argument for pardon powers. The argument’s popularity is puzzling, though. It does not justify any of the power’s distinctive features. Nor does it overcome any of the objections in section 3.

Recall that a pardon power can only be used in the direction of leniency, to cancel liability to a sanction. But legislators may, due to their limited perspective, be too lenient as well as too harsh. It would seem that we should want a power which can be used to extend as well as narrow liability.

Recall that the pardon power can only be used to set aside or dispense with the code, not alter it. If a code requires a suboptimal result, then it will also require the suboptimal result in like cases. The remedy should be a general exception in the law, to cover a class of closure cases, not its suspension in

²⁰ Aristotle, *The Nicomachean Ethics* (David Ross trans., OUP 2009) book V, ch 10.

²¹ Ibid.

²² N Hunard, *The King’s Pardon for Homicide Before AD1307* (OUP 1969); H Lacey, *The Royal Pardon: Access to Mercy in Fourteenth-Century England* (York Medieval Press, 2009).

²³ JM Beattie, “The Royal Pardon and Criminal Procedure in Early Modern England” (1987) 221 *Historical Papers* 9.

²⁴ W Blackstone, *Commentaries on the Laws of England: A Facsimile of the First Edition of 1765-1769* (University of Chicago Press 1979), vol. 4, 397.

just one such case. So, the argument from equity does not overcome Beccaria's objection.

Recall that a pardon power is held by the executive or legislature. If the aim is to do justice in particular cases, it would seem to be judges who are better suited to the task. It is judges, after all, who are experienced in deciding particular cases, who are able to compare the treatment of particular cases, who have privileged access to what has been said at a person's trial, etc. It follows that the objection from the separation of powers is not overcome.

Finally, recall that a pardon power is wholly discretionary. To do equity is to do justice. There is a duty to do justice. So, why is there no legal duty to grant a pardon, when justice demands it? It would seem that the objection from the rule of law remains outstanding as well.

Some legal scholars might take issue with this final claim, about discretion. The response would go like this. Pardon powers, we are supposing, are meant to correct for the imperfections of rules. Were such powers to be constrained by rules, the original problem would recur. The rules constraining the power would be imperfect; the imperfections would need correction; and to correct for them we would need a 'superpardon'²⁵. If judicial review occurs against a background of 'ex ante' rules, then judicial review 'contradicts the reason for having ... [a pardon power] in the first place'²⁶.

I accept that, if our aim is equity, it would be counterproductive to specify in advance the circumstances in which a pardon should be issued. But it is a mistake to think that this is what judicial review must involve. Suppose that a judge asks whether a decision is, by his or her own lights, correct. If the "correct" decision is the equitable decision, then review on a correctness standard is compatible with the power being used in light of all the reasons in a case, including those not specifiable in advance. Indeed, if judges are better able to assess what is just in particular criminal cases, judicial review on a correctness standard would promote rather than inhibit equity. Or consider review on procedural grounds. To require a fair hearing is compatible with – indeed, conducive to – the exercise of the pardon power on all relevant considerations, and thus equitably. Judicial review is if anything a friend not a foe to equity.

C

Because legislators cannot see into the future, their code will require suboptimal results in some cases. We should have a device to correct for these

²⁵ LR Meyer, 'The Merciful State' in A Sarat and N Hussain (eds), *Forgiveness, Mercy, and Clemency* (Stanford University Press 2007) 86.

²⁶ R Barkow, 'The Ascent of the Administrative State and the Demise of Mercy' (2008) 121 *Harvard Law Review* 1333, 1364.

results, but the pardon power is not it. There is, however, a power which is suited to solve the problem: a power of equitable interpretation.²⁷ This power is used to interpret the law to say what the legislature would have said had it been apprised of all of the facts. It can be used to extend, as well as limit liability. The power is held by judges. And it must be used in appropriate circumstances, a requirement which is enforceable through appellate review. It avoids all of the objections that pardon powers face.

The point is not lost on critics of pardon powers. Ross Harrison, for example, acknowledges the ‘need for flexibility’²⁸ in the law, but denies that a pardon power is called for. He says:

[N]o legal system should be so rigid that it is forced into doing things which are manifestly and absurdly unjust in particular unforeseen cases. But this is an argument for providing a means for dealing with such cases inside the system, not for reaching outside it to arbitrary judgment. It is an argument for the final courts of appeal to have sufficient confidence on occasion to make the law, and themselves prevent manifest injustice.²⁹

Consider, Harrison says, *R v Richard Bailey*³⁰. When Bailey was at sea, a criminal offence was created, which he broke. There was no way Bailey could have known what he did was against the law, and ultimately, he was granted a pardon. But, Harrison says, ‘this is something which the law should be able to decide for itself, without intervention of an extraneous political official’³¹.

6

I have rejected the traditional justification for pardon powers. Now I want to develop a better argument, one which focuses on gap cases.

A

It might seem that gap cases are an impossibility. If a code is truly optimal, does that not suggest that it only requires optimal results?

²⁷ See J Evans, ‘A Brief History of Equitable Interpretation’ in J Goldsworthy and T Campbell (eds), *Legal Interpretation in Democratic States* (Ashgate/Dartmouth 2002).

²⁸ R Harrison, ‘The Equality of Mercy’ in H Gross and R Harrison (eds), *Jurisprudence: Cambridge Essays* (OUP 1992) 120.

²⁹ *Ibid.*

³⁰ *R v Richard Bailey* (1800) 168 Eng. Rep. 651.

³¹ Harrison (n 27) 120.

Consider a familiar exchange about rule versus act consequentialism. It might seem that the set of rules with the best consequences will require all and only the acts with the best consequences. In that case, rule consequentialism is at risk of collapsing into act consequentialism. Rule consequentialists may respond by distinguishing two versions of their theory. One version assesses rules by the consequences of adherence to them. Perhaps this version collapses into act consequentialism. But another version assesses rules based on the consequences of their acceptance or internalisation. A set of rules extensionally equivalent to act consequentialism would be complicated and thus costly to accept or internalise. In light of these costs, it would be better to accept a simpler set of rules. By hypothesis, the demands of this simpler set of rules are not extensionally equivalent to the demands of act consequentialism.³²

I do not wish to endorse rule consequentialism, but I do want to take two points from this exchange. One is that increasing the accuracy of a code comes with several costs. There are *deliberation costs*, in the form of the time, effort, and argument it takes to figure out what the complex code demands. There are *error costs*, in the form of the accidental violations that result from misunderstanding what the complex code requires. There are also *uncertainty costs*: it is hard to predict what others will do, given the difficulty in understanding the complex code and predicting others' mistaken understandings of it. The second point is that the costs may outweigh the benefits of greater accuracy. A crude-and-simple code may beat an accurate-but-complex one.

The lesson is that code-makers may blamelessly require suboptimal results for either of two reasons. One is that they make the code in advance with limited foreknowledge. The other possibility is that they justifiably choose guidance and stability over accuracy. In this second scenario, an optimal code requires a suboptimal result, resulting in a gap case.

B

Closure cases can be remedied by equitable interpretation, as I said. But things are different when it comes to gap cases. Jim Evans has a good example:

A market is held most days Monday to Friday from 6am to around 11am. Sometimes it finishes a little earlier. Vehicles parked near the

³² For a detailed explanation and two variants of the criticism, see B Hooker, *Ideal Code, Real World* (OUP 2000) 93-99.

market clutter the area and prevent delivery vehicles making deliveries.³³

How should the local authority respond?

The local authority could make a by-law saying, “No parking near the market while the market is in process, except for delivery vehicles”. But this would invite endless argument. When, precisely, does the market stop being in process? What is near the market? What counts as a delivery vehicle?³⁴

What Evans calls ‘endless argument’ leads to all three of the costs I described above. It takes time, effort, and perhaps litigation to figure out what the by-law demands. The possibility for accidental error is high. And not knowing where others will park may lead to costs for others (e.g., a delivery company not knowing whether it will be able to find parking). To avoid these costs, the authority takes a different approach:

[The authority] paints a yellow line alongside certain section of the road, and posts a notice backed by a by-law alongside these sections saying, “No Parking Monday to Friday 6am to 11am, without a delivery sticker”.³⁵

This by-law is less accurate but much simpler.

Suppose this simplicity is worth the inaccuracy. What happens when someone parks over the yellow line ‘at 10:30am when the market has finished early, or on a public holiday’³⁶? These are ‘not cases for an equitable exception’³⁷ to the by-law. That is because these are ‘among the cases that it was judged necessary to prohibit in order to have a workable means of stopping parked vehicles interfering with the market’³⁸. The law is optimal though it yields a suboptimal result. Narrowing the law through interpretation would simply worsen the law. In Burke’s words, this is a law which ‘may in some instances be a just subject of censure without being at all an object of repeal’³⁹.

³³ J Evans, ‘Aristotle’s Theory of Equity’ in W Krawietz, N MacCormick, GH von Wright (eds), *Prescriptive Formality and Normative Rationality in Modern Legal Systems* (Duncker & Humblot 1994) xx.

³⁴ Ibid.

³⁵ Ibid.

³⁶ Ibid.

³⁷ Ibid.

³⁸ Ibid.

³⁹ E Burke, *The Works of Edmund Burke* (Samuel Holdsworth 1842), vol. II, 436. See also D Lyons, ‘Derivability, Defensibility, and the Justification of Judicial Decisions’ (1985) 68 *The Monist* 325, 334; F Schauer (n 18) 132-33.

C.

So, optimal codes sometimes require suboptimal results. The question is how to respond in such cases. Not by interpreting gap cases out of existence, or by otherwise altering the code. That much we know. There are broadly three strategies that remain, all of which take the code as given. The *full deviation* strategy is to set aside the code in all gap cases. The *full adherence* strategy is to adhere to the code in all gap cases. The third strategy is *selective deviation* (or *selective adherence*), which is to deviate in some gap cases and adhere in others.

Jeremy Bentham put forward a version of the full deviation strategy. He proposes a system of code-making and adjudication with two elements.⁴⁰ First, all law-making authority is vested in the legislature, which makes a code based on the ‘utilitarian principle’. The ‘primary virtues’ of the code are ‘certainty, stability, and efficiency’⁴¹. The legislature’s code will sometimes yield suboptimal results. But Bentham does not think that judges should “correct” the code in such cases, because this would undermine the virtues of the code. Second, though, judges decide cases by directly applying the utilitarian principle. If the application of that principle favours a result different than that required by the code, then the judge will deviate from the code. The code is a guide to the relevant utilities, and a valuable rule of thumb for a judge faced with a complex set of considerations; but it does not settle what is to be done.

So, the idea is that different considerations are relevant at different stages. At the law-making stage the aim is to guide subjects. Accuracy is balanced against simplicity. The law-applying stage is different. Accuracy is all that matters at this stage. Bentham makes clear that judges should deviate from the code whenever it is at odds with the balance of underlying reasons, that is, in every gap case.

Gerald Postema points out a fatal flaw in Bentham’s proposal. A legal code promotes predictability and stability ‘only so long as citizens are convinced and have good reason to believe that others are convinced, that the laws on the books correspond closely to the laws applied and enforced by the courts’⁴². This will not be the case, on Bentham’s scheme, because judges sometimes deviate from the code. Indeed, ‘Bentham seems to believe that there will be a significant number of occasions on which deviation ... will be justified on utilitarian grounds’⁴³. Moreover, because Bentham insists that judicial

⁴⁰ My account of Bentham’s theory is based on G Postema, *Bentham and the Common Law Tradition* (OUP 1986) 395ff.

⁴¹ Postema (n 39) 448.

⁴² *Ibid* 454.

⁴³ *Ibid*.

proceedings should be public, and that judges give public reasons for their decisions, judicial deviations from the code will also be public. As a result, Postema says, ‘Bentham’s strategy must fail, for no matter how well *initially* public expectations are fixed on the code, they will inevitably shift back to focus on the activities of the courts and the patterns that emerge from them’⁴⁴.

Could Bentham reply that because the decisions of courts have no law-making effect, they will not be treated as a source of guidance? The difficulty is that judges will inevitably strive for consistency and coherence in their decisions.⁴⁵ As a result, ‘the attempt formally to deny precedential effect to judicial decisions will be futile’, and ‘a system of precedent-based case-law will inevitably arise’⁴⁶. Postema's claim, as I take it, is that public and consistent deviations from a code effectively alter the code, even if they do not do so officially. So, Bentham’s strategy effectively eliminates gap cases. By hypothesis, eliminating gap cases from a code would make the code worse, so the strategy should be rejected.

D

Bentham insists on consistent and public deviation in gap cases. His proposal fails, and if Postema is right, one reason is the publicity requirement. Could we improve on Bentham’s idea by dropping the publicity element? The idea is that so long as our deviations remain secret, they will not destabilise the code.⁴⁷

It seems to me that there are two serious problems with this revision to Bentham’s proposal. First, given consistency, secrecy is hard to maintain. For a judge to secretly deviate from the code would require both for the judge not to apply the code and for the judge to make it appear that he or she had applied the code. The second part is the hard part. The judge could lie about the facts of the case, make up a defence, or otherwise nullify the law. But this is not workable as a general solution. Eventually people will notice that code-breakers are walking around free even though under the code they should be incarcerated. Certainly, the code-breakers will start to catch on.

⁴⁴ Ibid.

⁴⁵ Ibid 457.

⁴⁶ Ibid 454. Postema also says that people will inevitably interpret the code in light of judicial pronouncements on its meaning.

⁴⁷ For a related proposal, see M Dan-Cohen, ‘Decision Rules and Conduct Rules: On Acoustic Separation in Criminal Law’ (1984) 97 Harvard Law Review 625. I take Cohen to argue that officials can systematically deviate from the rules applicable to citizens in the direction of leniency without undermining those rules if their deviations are hidden from citizens. But Dan-Cohen is careful not to claim that it is possible to keep any deviation perfectly secret, or to keep all deviations secret to even some degree.

Second, given secrecy, consistency is hard to maintain. A secret practice is by nature hard to monitor. Because it is hard to monitor, it is hard to enforce. There are two worries: that the code will be set aside when it should not be, and that it will not be set aside when it should be. I do not see how to avoid these worries. Legal accountability through judicial review is a non-starter: judicial review requires a claimant who can identify mistakes, which depends on knowledge and therefore publicity. The same is true of political accountability: we cannot hold the government to account for what it has done if we do not know what it has done.

So, it might *look* attractive to always deviate from an optimal code when it requires a suboptimal result, because it seems to yield all of the benefits of the code and none of its costs. But we should not deceive ourselves: we cannot have our cake and eat it too, even if our cake-eating happens in secret.

E

If we should not deviate from a code in *all* gap cases, publicly or not, it might seem natural that we should endorse the full adherence strategy and keep to the code in all gap cases. But this is to swing too far the other way. As Alan Goldman says, a single decision ‘on moral merits instead of law will have little effect on the stability and predictability of the legal system’⁴⁸. Given that it has little negative effect, while also avoiding a suboptimal result, a single deviation is justified. And if a single deviation is justified, then adherence in every case is not.

F

So far, I have rejected two extremes: deviation or adherence in all gap cases. Now I want to set out my preferred alternative: selective deviation.

Call the *deviation rate* the frequency (expressed as a probability) of deviations from an optimal code to avoid a suboptimal result. Over at least some range, each deviation has two consequences. One is the benefit of avoiding a suboptimal result. The other is the cost of undermining an optimal code. As the deviation rate increases, both the benefits and the costs increase. Call the rate at which the marginal benefit of deviation equals the marginal cost the *optimal deviation rate* or *optimal rate* for short. From the discussion so far, we know, first, that deviating in all gap cases is worse than deviating in none; and second, that deviating in no gap cases is worse than deviating in some. Hence the optimal deviation rate is between 1 and 0.

⁴⁸ A Goldman, *Practical Rules: When We Need Them and When We Don't* (CUP 2002) 42. See also Brand-Ballard (n 15) 130.

This is a claim about deviation across all gap cases in a legal system. But we can make the same claim at a smaller scale. Let us say that a *case-type* supervenes on the features which are legally and morally relevant to a case's treatment. A *gap case-type* is a case-type in which an optimal code requires a suboptimal result.⁴⁹ If Postema is right, then we do not want to deviate in all tokens of a gap case-type, because that would effectively amend the code to include an exception for that case-type. If Goldman is right, then we do not want to adhere in all tokens of a gap case-type either, because deviation in a single case has a significant benefit and a trivial cost. This tells us that the optimal deviation rate in any gap case-type is between 1 and 0.

That may sound innocuous, but it has an important implication. Call two cases *alike* when they are tokens of the same case-type, i.e., when they have the same legally and morally relevant features. The conclusion in the last paragraph is that, in token cases of a given gap case-type, we ought to deviate in some and adhere in others. It follows that we ought to treat like cases unlike. We ought to act inconsistently.

I imagine some readers will see this conclusion as a *reductio* of my analysis. Is it not irrational to treat like cases differently? Normally, yes; but not always. Imagine that a ship sinks and you have the only life raft. Walt and Wade are in the water. You can rescue one of them, but if you rescue both of them the raft will capsize, and everyone will drown. Assume that neither is more deserving, needy, etc. I take it that the relevant rational requirements are as follows:

- You are required to rescue Walt *or* Wade.
- You are required not to rescue Walt *and* Wade.
- You are not required to rescue Walt.
- You are not required to rescue Wade.

On this basis, if you rescue Walt but not Wade, then you do everything you are rationally required to do, even though you treat people alike situated differently.

Similarly, we are required to deviate in some token cases of a given gap case-type, required not to deviate in all such token cases, and not required to deviate in any particular token case. By deviating in one case and adhering in another, we treat like cases differently but do not act irrationally.

G

Pardon powers are not the right tool to correct for closure cases; equitable interpretation is more suitable. But gap cases are different. These cases call

⁴⁹ Adapted from Brand-Ballard (n 15) c 13.

for a strategy of selective deviation. Equitable interpretation, which creates prospective and general exceptions to the code, cannot help us here.

Pardon powers on the other hand are perfect for the task. I said that a pardon sets aside the law in a particular case, without changing it. If selective deviation is our aim, this is a suitable means. For we do not want to change optimal rules; that would only make them worse. We want to dispense with them in some gap cases, while leaving them intact for other gap cases. Earlier I set out Beccaria's objection:

Beccaria's objection. A pardon power is justified only if it could be true that (1) the law ought to be set aside in some case, but (2) the law ought not to be changed. But, if (1) is true, then (2) is not; and if (2) is true, then (1) is not. So, a pardon power is unjustified.

Now we can see that, contrary to the objection, (1) and (2) can both be true. Moreover, (1) and (2) are in fact both true in some gap cases. In such cases, there is a role for a power which cancels the liability imposed by a rule, without changing that rule – that is, for a pardon power.

The second objection to pardon powers was from the rule of law:

The rule of law objection. A pardon power grants an unconstrained discretionary power. But discretionary power can be justified only if it is constrained by general, externally enforceable laws. So, a pardon power is unjustified.

I grant that arbitrariness is normally an undesirable feature of a power. But it is ideal for the purpose of selective deviation. When we selectively deviate, we treat like cases differently. To make selective deviation possible, we therefore need a power which can be used arbitrarily. We need a power which can be exercised in one case and not in another case for no reason.

Further, and more strongly, once we introduce a power to set aside the law, the rule of law *favours* of leaving that power unconstrained. The rule of law demands rule-based decision-making because normally that leads to predictability and stability. With respect to powers to set aside the law, it is the other way around. If citizens could count on the law being set aside, it would undermine the code. So, if we are to uphold the code, dispensations must be granted inconsistently. To be granted inconsistently, they must not be required under conditions set out in advance. They should 'be like lightning bolts, relatively rare and in principle hard to predict'⁵⁰. Ironically,

⁵⁰ J Harrison, 'Pardon as Prerogative' (2000-01) 13 Federal Sentencing Reporter 147, 148.

the rule of law itself insists that a power to set aside the law be not ruled by law.

To whom should the responsibility for selective deviation be given? There are two reasons not to give it to judges. The first is the one Postema mentions: judges aim for consistency in their judgments. Left to their own devices, judges will inevitably develop a pattern of rule-like precedents for the use of a pardon power. But a system of rules for when to set aside an optimal code would undermine the code, which is what we wish to avoid. By contrast, neither the executive nor the legislature is institutionally committed to consistency or coherence in its decision-making.⁵¹

Second, were the responsibility to deviate from an optimal code left to judges, it would be up to them to arrange their decisions so that, collectively, they adhered to the code in enough cases and deviated in enough cases. But whether a particular judge should deviate in a particular gap case depends on what other judges will do in *their* gap cases. This is a type of coordination problem, where the right decision for any one judge depends on the decisions other judges make. It is a difficult problem for judges to solve, given the size of the judiciary, the large number of judicial decisions, and judges' lack of information about each other's decisions.⁵²

Think back to the objection from the separation of powers:

The separation of powers objection. A pardon power grants a non-judicial body the power to cancel legal liability in particular criminal cases. But this type of power could be justified only if it is held by a judicial body. So, a pardon power is unjustified.

Earlier I said that it seems puzzling that pardon powers are given to a branch other than the judiciary, given that the judiciary is best-suited to determining what is just in particular cases. But we can see now that being well-suited to wielding a pardon power is about more than particularistic justice. It is also about *not* acting consistently and about being able to anticipate and contain the systemic effects of deviating from a code. On these measures, 'efficiency' favours giving the pardon power to a branch other than the judiciary. Again there is an irony: far from opposing a pardon power, the separation of powers *favours* giving the power to set aside the law to the legislature or the executive.

⁵¹ I do not claim that the executive and legislature will develop *no* patterns or guidelines as to when to grant pardons. The history of pardon powers suggest that the process is often bureaucratized. See eg G Chadwick, *Bureaucratic Mercy* (Garland 1992). My point is comparative: the executive and legislature are less disposed to develop binding constraints on decision-making than judges.

⁵² Goldman (n 47) 42-46.

My argument identifies a pro tanto reason for pardon powers. The reason is not of trivial weight. The potential to do justice in particular cases while preserving a code as a source of stability and guidance is important enough to weigh heavily in the balance of reasons. But it is a reason which could be outweighed. In particular, if it is very likely that a pardon power would be seriously abused – for example, by being used for a purpose other than selective deviation – then there may be a reason against a pardon power all things considered.

7

To show how my proposal might work in practice, let me return to Sauvage and Dimet’s cases. As I said, these two cases are alike, yet they were treated differently: Sauvage was granted a pardon, Dimet was not. Is the inconsistency justified?

Under French law, self-defence is available as a legal defence only if the threat defended against is imminent or immediate. When Sauvage shot Marot, he was sitting with his back turned to her. Marot posed no immediate threat to Sauvage. In the law’s eyes, Sauvage did not shoot Marot in self-defence. The same was true of Dimet: the threat she defended herself against was not imminent. Assume that the results in these cases are suboptimal. Should the law be reformed, to avoid requiring these results? Perhaps. But it matters a great deal *how* we would go about reforming the law.

Suppose that the rationale for the imminence requirement is that, if a threat is not imminent, a person should seek help from the state, rather than taking matters into their own hands. For many battered women, seeking aid from the state is not a feasible or effective option. We could therefore avoid requiring results like in Sauvage’s and Dimet’s cases by introducing an exception to the law, so that the imminence requirement does not apply when state-aid is not a feasible or “effective” alternative to violence.

Introducing an “effectiveness exception” will meet with an immediate objection, however. Whitley Kaufman writes:

The problem is that the notion of “effectiveness” is so vague and open-ended, it would exceedingly complicate jury trials, resulting in lengthy, complex debates over how to define “effective”, whether the state was effective, and just how effective it had to be before force was justified. It is doubtful that such a standard could constrain the danger of the resort to vigilant violence⁵³

⁵³ W Kaufman, ‘Self-Defense, Imminence, and the Battered Woman’ (2007) 10 *New Criminal Law Review* 342, 363. For critical commentary, see P Robinson, S Garvey, and K Ferzan (eds), *Criminal Law Conversations* (OUP 2011) 415-26.

An effectiveness exception would take time and effort to apply. It would create uncertainty. And it would, Kaufman thinks, start us on a ‘slippery slope’⁵⁴ which will end with sanctioning vigilantism.

Assume, purely for the sake of argument, that the costs Kaufman identifies are real and significant enough that we should *not* qualify the imminence requirement. And assume, again for the sake of argument, that there is no other satisfactory way of avoiding requiring the results in Sauvage’s case and Dimet’s case.⁵⁵ It follows that we have suboptimal results required by an optimal rule. That is, Sauvage’s case and Dimet’s case are gap cases. How, then, should we deal with such cases?

We should *not* adhere to the imminence requirement in every gap case. Nor should we deviate from the requirement in every such case: that would introduce an effectiveness exception through the back door. Instead, we should selectively deviate. On these assumptions, then, it was not irrational to grant Sauvage but not Dimet a pardon. Nor would it have been irrational to grant Dimet but not Sauvage a pardon. Judicial review would limit the potential for this kind of inconsistency. So, it is a virtue that in France the pardon power remains a wholly discretionary act of conscience.

To be clear, I do not assume that the French legal code ought to include a pardon power, all things considered. Suppose that all things considered it ought *not* to grant a pardon power. The fact is that it does. In this non-ideal scenario, the President faces a choice as to whether to use the power, and if so how. My argument is that, *given* the existence of the power, all-things-considered justified or not, the President ought to use that power in some gap cases but not in other like cases.

8

I said that a legal power is a pardon power if and only if (1) its exercise sets aside a legal rule imposing criminal liability, (2) its exercise does not alter a legal rule, (3) its exercise is not constrained by any legal rule, and (4) it is possessed by a non-judicial official.

I defended pardon powers in three broad stages. First, even an optimal legal rule sometimes requires suboptimal results in particular cases. We can respond to these “gap” cases in several ways. One strategy is to always deviate from the legal rule to achieve the optimal result. Another strategy is to never deviate from the legal rule. The strategy I favoured is selective deviation, according to which we should deviate in some gap cases but adhere

⁵⁴ Kaufman (n 52) 363.

⁵⁵ Some jurisdictions do without an imminence requirement (eg Canada) and some include other defences like loss of control (eg England and Wales).

in others. If selective deviation is the end, then the means should be a power which sets aside a legal rule while leaving it intact for application in other gap cases. This calls for power with features (1) and (2). Second, selective deviation requires that like cases be treated differently. As a result, the means of selective deviation must be a power which can be used arbitrarily and is therefore unconstrained. This favours (3). Third, it is not enough that the power can be used inconsistently; it must in fact be used inconsistently. Because judges tend towards consistency in their decision-making, the means of selective deviation ought to be given to a non-judicial official, which accounts for (4).

This argument overcomes all three objections to pardon powers. Contrary to Beccaria's objection, it is sometimes true both that a legal rule is justified and that it ought to be set aside in a particular case. Contrary to the objection from the separation of powers, selective deviation is a task better carried out by non-judicial officials. Contrary to the objection from the rule of law, a power of selective deviation must be unconstrained to preserve the law's guiding function.