The Pardon Paradox

Adam Perry*

1. Introduction

Almost every constitution in the world confers a power to pardon.[1] Pardon powers are found in the constitutions of old states and new states, Western states and non-Western states, states with a Christian tradition and states without one. Pardon powers are part

* Associate Professor, Faculty of Law, University of Oxford. My thanks to Tara Alberts and Swati Jhaveri for their helpful comments. This paper is a draft and some things are still missing, including some citations. Any comments or criticisms, no matter how small, are welcome at adam.perry@law.ox.ac.uk.

[1] Of the 193 United Nations member states' constitutions, 188 grant a power of pardon or amnesty. The exceptions are Andorra, Bosnia and Herzegovina, Syria, Saudi Arabia, and Yemen. Information drawn from Constitute Project <www.constituteproject.org> accessed 20 March 2019. See also: J Close, 'Amnesty Provisions in the Constitutions of the world: A Comparative Analysis' (*International Law Blog*, 5 January 2015) <https://bit.ly/2CEjWm7> accessed 20 March 2019.

of the constitutions of states as diverse as France, Indonesia, Peru, Russia, the United Kingdom, and the United States. Despite their ubiquity, pardon powers seem to conflict with two of the most basic principles of constitutionalism. Contrary to the separation of powers, a pardon power gives to a branch of government other than the judiciary a role in determining criminal liability in particular cases. Contrary to the rule of law, a pardon power is traditionally neither controlled by nor ruled by law. Hence the pardon paradox: pardon powers are everywhere but belong nowhere.

I aim to show that, despite appearances, neither the separation of powers nor the rule of law is a reason *against* an uncontrolled pardon power. This dissolves the paradox. I aim to show, further, that there is a reason - strong, though not necessarily conclusive - *for* an uncontrolled pardon power. To be clear, my interests are not explanatory. For all I know the reason *why* pardon powers are common is the nefarious one that rulers wish to aggrandise themselves, or the mundane one that powers in new constitutions are copied across from powers in old constitutions. I set aside this sort of historical or political inquiry to focus on the justificatory issues.

2. The separation of powers

A pardon lifts or lessens criminal liability. It does so by setting aside or dispensing with the law, not by changing it. And it is granted by someone other than a judge – the executive, normally, or the legislature, less commonly.[2]

So defined, the power to pardon seems to conflict with the principle of the separation of powers. That principle says among other things that judicial functions should be allocated to the judiciary. A "judicial function" includes 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,

[2] More exotic arrangements are possible, eg, the Constitution of Uzbekistan, Article 93, says that pardons and amnesties are proposed by the President and decided upon by the Senate.

highly developed rules for fact finding, procedural protections including the right to be heard, and so on.

We match judicial functions with the judiciary for reason of what in the jargon is called 'efficiency', meaning that the judiciary is well-suited to exercise judicial functions.[3] 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, too. But such tribunals are not independent of the prosecution in a criminal proceeding.

Pardon powers are at odds with the allocation of judicial functions to the judiciary. To grant or refuse a pardon is to make a decision about how to treat a specific case, on a disputed matter, with serious consequences. This is a judicial function which is granted to the executive or legislature. James Fitzjames Stephen said it well:

> [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 oaths ….[4]

This problem 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 judiciary is suited to the job, it seems hard to believe that such a different branch would *also* be suited to it.

---

[3] 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*).

[4] JF Stephen, *Selected Writings A General View of the Criminal Law* (KJM Smith, ed. OUP 2014) 217-8. See also J Sebba, 'The Pardoning Power – A World Survey' (1977) 68 Journal of Criminal Law & Criminology 83: 'This independence [of the judiciary] would appear to be threatened by vesting in a non-judicial authority the power to pardon offenders duly convicted and sentenced in the course of a judicial process'.

3. The rule of law

Traditionally, a pardon power had a further feature: it was wholly arbitrary. The power was 'grounded solely in the will of the dispenser'. The 'dispenser' could grant or refuse a pardon for good reason or bad reason or no reason. He or she could carefully consider every pardon petition or reject them all unread. A pardon was a gift, an act of grace.

One of the requirements of the rule of law is generality. The principle is not opposed to powers to create specific legal norms or to make specific legal changes. But it does insist that these powers are constrained by rules. Those rules must 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'[5]. The most important piece of this machinery is an independent judiciary with reviewing powers.

Pardon powers seem to violate the rule of law, so understood.[6] Pardon powers are not subject to any type of external review. As a result, they are not subject to rules backed by judicial enforcement mechanisms. Suppose you have been convicted of an ordinary federal offence in America. All your appeals are exhausted, so you ask for a pardon. The President's power is at its most expansive. He can do as he likes. And if you don't like what he does, you have no recourse. You are at the President's mercy. The law does not rule; the President does. The pardon power is a form of legally sanctioned arbitrariness – 'lawful lawlessness'[7], in Austin Sarat and Nadar Hussain's words.

To clarify, that pardon powers were wholly discretionary did not mean that they were entirely free from judicial review. Judges were willing to determine whether an act fell within the scope of a

---

[5] J Raz, *The Authority of Law* (OUP 1979) 218.

[6] 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. For a closely related and insightful discussion of public mercy and arbitrariness, see A Tuckness and J Parrish, *The Decline of Public Mercy* (CUP 2014) 274-81.

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

pardon power. They might ask whether a person who purported to grant a pardon was in fact empowered to do so, for example.[8] What judges traditionally were unwilling to do is review pardon decisions on substantive (merits-based) or procedural grounds. Thus, as Hugo Adam Bedau says, pardon decisions were 'standardless in procedure, discretionary in exercise, and unreviewable in result.'[9]

4. Judicial review

Although pardon powers are ubiquitous, they look to be at odds with two of the most fundamental constitutional principles. I find this paradoxical, because it seems to me there is something to be said for pardon powers in their traditional, uncontrolled form, an intuition I struggle to reconcile with the apparently objectionable features of such powers. But others will see no puzzle here, merely a reflection of the disappointing fact that we do not always create the powers we should. The threats posed to the separation of powers and the rule of law are real and require intervention. It would be best, on this way of thinking, were pardon powers externally constrained and exercisable in a way similar to how judges exercise their powers. This, broadly, is the view taken by courts around the work since the 1980s.

The pioneer was India. In its 1981 decision in *Maru Ram v Union of India*, the Supreme Court of India held that 'all power, whatever its source, must, in its exercise, anathematise arbitrariness'[10]. The pardon power is no exception. The power 'cannot run riot; for no legal power can run unruly like John Gilpin on the horse but must keep sensibly on a steady course'[11]. Specifically, the power cannot be 'exercis[ed] arbitrarily or mala fide'[12]. Reliance on 'capricious

[8] eg, *ex parte Crump*, 10 Okla. Crim. 133 (Okla. Crim. App. 1913) (determining whether during a vacancy in the of office of Governor the Lieutenant Governor was empowered to grant a pardon).
[9] HA Bedau, 'The Decline of Clemency in Capital Cases' (1990) 8 New York University Review of Law and Social Change 255, 257.
[10] *Maru Ram v Union of India and Ors* (1981) 1 SCC 107, [62] (Krishna Iyer J, for himself Chandrachud CJ and Bhagwati J).
[11] *Maru Ram* (n 10) [62]
[12] *Maru Ram* (n 10) [62]

criteria'[13] must render a pardon decision invalid. Anything else would amount to a rejection of reason and the rule of law. 'Every action of the executive Government must be informed with reason', it said, and so must be 'free from arbitrariness'[14]. Freedom from arbitrariness, according to the Court, is the 'bare minimal requirement'[15] and the 'very essence'[16] of the rule of law.

Since *Maru Ram*, Indian courts have said that they are willing to review pardon decisions on a wide range of grounds. In the 2006 case of *Epuru Sudhakar v Government of Andhra Pradesh*[17], for example, the Supreme Court said that pardon decisions were reviewable for (i) failure to exercise discretion, (ii) bad faith, (iii) failure to take into account relevant considerations or to exclude irrelevant considerations; (iv) arbitrariness.[18]

Where India led, courts elsewhere have followed. In England and Wales, the prerogative of mercy (the power under which pardons are granted) had for centuries been considered unreviewable, as all prerogative powers were. Lord Denning in 1971 said:

> The high prerogative of mercy was exercised … with the greatest conscience and care. The law would not inquire into the manner in which that prerogative was exercised.[19]

In 1993, however, the High Court held in *R v Secretary of State for the Home Department, ex p Bentley* that decisions under the prerogative are reviewable for errors of law[20]. Later it held that pardon decisions are reviewable on the basis of a failure to take into account relevant considerations[21]. Northern Irish courts have compared pardon decisions to see whether they are 'unequal'[22] or 'irrational'[23].

---

[13] *Maru Ram* (n 10) [65].
[14] *Maru Ram* (n 10) [63].
[15] *Maru Ram* (n 10) [63].
[16] *Maru Ram* (n 10) [63].
[17] *Epuru Sudhakar and Anor. v Government of Andhra Pradesh and Ors* (2006) 8 SCC 161.
[18] *Epuru Sudhakar* (n 17) [34] (Pasayat J).
[19] *Hanratty v Lord Butler of Saffron Walden* (1971) 115 SJ 386
[20] *R v Secretary of State for the Home Department, ex p Bentley*
[21] *R (Page) v Secretary of State for Justice* [2007] EWHC 2026 (Admin).
[22] *McGeough v Secretary of State for Northern Ireland* [2012] NICA 28, [16].
[23] *McGeough v Secretary of State for Northern Ireland* [2012] NICA 28, [10].

Over the same period, the Privy Council has undergone a remarkable change of perspective. In *de Freitas v Benny*, Lord Diplock in the Privy Council said that a pardon power is 'the exampler of a purely discretionary act'[24]. That case was decided in 1975. It was affirmed in 1996 in *Reckley v Minister of Public Safety (No 2)*, in which the pardon power was described as an unreviewable 'act of grace'[25]. But a mere two years later, the Privy Council reversed itself, holding in *Lewis v Attorney General*[26] that the pardon power *is* subject to judicial review on procedural grounds. Recently, in *Pitman v State*[27], the Privy Council opened the way to review on substantive grounds as well.

Even in the United States, where the presidential pardon power was long regarded as absolute, courts are now willing to review pardon decisions on limited grounds. In the 1998 case of *Ohio Adult Parole Authority v Woodard*, Rehnquist CJ echoed earlier cases when he said, 'the heart of executive clemency' is a 'matter of grace'[28]. But the majority of justices held that the pardon power was subject to minimal procedural safeguards. O'Connor J, who formed part of the majority, gave as an example of a ground of legitimate judicial intervention 'a state official flip[ing] a coin to determine whether to grant clemency'[29].

To the list of jurisdictions in which pardon decisions are newly amenable to judicial review we could add Canada[30], South Africa[31], Singapore[32], and many others.[33] I do not wish to over-generalise. There are still jurisdictions in which pardon decisions remain

[24] *de Freitas v Benny* [1976] AC 239, 247.
[25] *Reckley v Minister of Public Safety and Immigration (No 2)* [1997] 1 AC 527, 540.
[26] *Lewis v Attorney General of Jamaica* [2001] 2 AC 50.
[27] *Pitman v The State* [2017] UKPC 6, [50]. See also Lord Mance, 'Justiciability' (2018) 67 ICLQ 739, 755 (*Pitman* 'postulates its availability [ie, judicial review] on substance').
[28] *Ohio Adult Parole Authority v Woodard,* 523 U.S. 272 (1998), 281.
[29] *Woodard* (n 28) 289 (emphasis in original).
[30] *Thatcher v Canada (Attorney General)* [1997] 1 FC 289.
[31] *Albutt v Centre for the Study of Violence and Reconciliation* [2010] ZACC 4
[32] *Yong Vui Kong v Attorney-General* [2011] SGCA 9.
[33] For a good overview of the trend towards judicial review of pardon decisions, see A Novak, *Executive Clemency in Gobal Perspective* (Routledge 2016) ch 8.

unreviewable, for example, France[34]. And when courts do review pardon decisions, they do so on varying grounds. Indian courts are willing to review pardon decisions on many grounds on which American courts are not, for instance. But the trend is clear: greater judicial in a greater number of jurisdictions.

The expansion of judicial review addresses both objections to pardon powers. Through judicial review, pardon powers are subject to general standards of rationality and fairness. This defuses the threat to the rule of law. What these standards demand, it turns out, is a quasi-judicial process: a fair hearing, an unbiased decision-maker, consideration of all relevant factors, etc. Judges have not reallocated the pardon power to themselves. They have however forced other branches of government to act as they, the judges, would. In this way the threat to the separation of powers is reduced.

5. Results and codes

Judges are not wrong to worry about arbitrary power. In general, discretionary powers *should* be judicially reviewable. But I think that judges have tended to assume that pardon powers are like other discretionary powers. They have not stopped to ask *why* we have pardon powers in the first place, and whether, as a result, there is reason to treat them specially.

To set out the traditional justification for pardon powers and my alternative, it will help to have some definitions to hand, which I largely borrow from Jeffrey Brand-Ballard:

> *Superiority*. One result in a case is superior to another if and only if, absent the law, there is greater reason to reach the first result than the second.
>
> *Optimality*. A result in a case is optimal if and only if there is no result superior to it.
>
> *Suboptimality*. A result in a case is suboptimal if and only if there is a result superior to it. [35]

A result may be required, prohibited, or permitted by the legal rules of a system. The rules of a system are the system's 'code'.

---

[34] See R Levy, 'Pardons and Amnesties as Policy Instruments in Contemporary France' (2007) 36 Crime & Justice 551, 557.

[35] Definitions adapated from J Brand-Ballard, *Limits of Legality* (OUP 2010) 75.

I want to focus on arguments for pardon powers which cast them as a means of bringing about an optimal result when a code requires a suboptimal result. This means I will not be discussing some commonly mentioned reasons for pardon powers. For example, I will not be discussing whether pardons are acts of forgiveness. To forgive a wrong suggests that it was permissible and therefore not suboptimal to attach blame in the first place. Nor will I be discussing the potential to use pardons to correct for wrongful convictions or judicial errors, because neither type of result is required by law. I doubt these arguments succeed, which is why I am focusing my attentions elsewhere, but nothing I say depends on their failure.

In parallel with the definitions above, call one code 'superior' to another if and only if there is greater reason to adopt the first code than the second. Call a code *optimal* if there is no code superior to it and *suboptimal* if there is. A suboptimal result may be required by either a suboptimal code or an optimal code. Finally, call a case in which a suboptimal code requires a suboptimal result a *negative-closure case*, and a case in which an optimal code requires a suboptimal result a *negative-gap case*.[36]

6. Negative-closure cases

Why would a code be suboptimal? A code may be *culpably suboptimal*: legislators erred when they made the code and are blameworthy for doing so. The pardon power is not the right device to correct for the suboptimal results such a code produces. Most obviously, that is because the pardon power may be held by the legislature itself. The other possibility is that the code is *excusably suboptimal*: the code is not as good as it could be, but legislators 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. Equity

[36] Ibid.

When they make a code, legislators make a set of prospective rules. Because these rules are prospective, they must be based on reasons which legislators can foresee. Legislators 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. They are only human, and as Aristotle says, 'practical affairs [are] of this kind from the start'[37]. Negative-closure cases are practically unavoidable.[38]

How should officials respond in excusable negative-closure 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 '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.[39] In this way, we will accomplish a form of 'equity'[40].

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 negative-closure cases. It is no surprise therefore that many courts and commentators see the pardon power as a fitting vehicle for equity in the Aristotelian sense. The pardon power is a 'court of equity' in the monarch's own breast, said Blackstone.[41] 'The criminal code … takes of so much necessary severity', Hamilton wrote, 'that without easy access' to pardons 'justice would wear a

---

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

[38] 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.

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

[40] Ibid.

[41] 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.

countenance too sanguinary and cruel'[42]. And in *State v Leak* it was said that '[p]ardons and remissions are in derogation of the law, and should never be extended except in cases which, could the law have foreseen, it would have excepted from its operation'.[43]

B. Objections

The argument from equity is, as I said, the traditional argument for pardon powers. The argument's popularity is puzzling, though, because it does not justify any of the power's distinctive features.

Recall that the pardon power can only be used to set aside or dispense with the code, not alter it. If the 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 negative-closure cases, not its suspension in just one such case. And recall that a pardon power can only be used in the direction of leniency, to lift or lessen criminal liability. 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, further, that a pardon power is held by the executive or legislature. If the aim is to do justice in particular cases, it is judges who are better suited to the task. Judges are the ones 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, and so on. Finally, recall that a pardon power is wholly discretionary. To do equity is to do justice, and there is a duty to do justice. It is odd, then, to choose as a vehicle for equity a power which confers the discretion *not* to do equity in some or all negative-closure cases.

Some scholars would take issue with my final objection, 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

---

[42] A Hamilton, 'Federalist No. 74' in I Kramnick (ed), *The Federalist Papers* (Penguin 1987) 422.
[43] *State v Leak*, 5 Ind., 359, 363 (1854).

would need a 'superpardon'[44]. If we then think, as Rachel Barkow does, that judicial review occurs against a background of 'ex ante' rules, then we will believe that judicial review 'contradicts the reason for having … [a pardon power] in the first place'[45].

I accept that 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. Consider substantive review on what in Commonwealth circles is known as a 'correctness standard'. Here the judge asks whether a decision is, by his or her own lights, correct. If the correct decision is understood as 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. Equitable interpretation

Because legislators cannot see into the future, their code will require suboptimal results in some cases. We should have a device to correct for negative-closure cases, but the pardon power is not it. There is, however, a power which is perfectly suited to solve the problem: a power of equitable interpretation.[46] This power is held by judges. It can be used to alter the law to say what the legislature would have said had it been apprised of the all of the facts. It can be used to extend as well as limit liability. Finally, it is a power which must be used, in appropriate circumstnaces. It is everything Aristotle might want. Moreover, a power of equitable

---

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

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

[46] The varieties of equitable interpretation are discussed in J Evans, 'A Brief History of Equitable Interpretation' in J Goldsworthy and T Campbell (eds), *Legal Interpretation in Democratic States* (Ashgate/Dartmouth 2002).

interpretation does not pose particular problems for the separation of powers or rule of law.

The point is not lost on critcs of pardon powers. Ross Harrison, for example, acknowledges the 'need for flexibility'[47] 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.[48]

Consider, Harrison says, *R v Richard Bailey*[49]. 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'[50].

7. The case for pardon powers: a sketch

I have rejected the traditional argument for pardon powers. It will take me a while to develop my alternative, so let me indicate where I am going. My argument has three steps.

First, optimal codes require suboptimal results. Suppose that a code includes a blanket pohibition on assisted suicide. This prohibition yields some suboptimal results. Think of people who are terminally ill and suffering terribly, who sincerely wish to end their life at a time of their choosing with a doctor's help but cannot lawfully do so. But suppose – purely for the sake of argument – that were the code to be changed to include an exception for this type

[47] R Harrison, 'The Equality of Mercy' in H Gross and R Harrison (eds), *Jurisprudence: Cambridge Essays* (OUP 1992) 120.
[48] Ibid.
[49] *R v Richard Bailey* (1800) 168 Eng. Rep. 651.
[50] Harrison (n 47) 120.

of case, it would lead to people being pressured to commit suicide. Overall, let us imagine, it would be better to leave the code as is.[51]

How should we respond when an optimal code produces suboptimal results? One strategy is to always deviate from the code. Another strategy is to always adhere to the code. I reject both of these strategies. The strategy I favour is optimal deviation, which involves deviating in some suboptimal result cases and adhering in others. The most important thing about optimal deviation is that it leads to treating like cases differently. For example, in some cases when a person is terminally ill, suffering terribly, etc., we would set aside the code. In some like cases, we would do not. Treating like cases unalike may sound terribly unjust. But it is more defensible, I say, than the alternatives. This is the second step in the argument.

The third step is to show that pardon powers are a good way to achieve optimal deviation. If optimal deviation is the aim, then we need a power which can be used to set aside the law without changing it, and which can be used arbitrarily or inconsistently. The pardon power is just what we want. But it 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, it follows that there is good reason to give the pardon power to the executive or legislature instead. There is more to the argument than that, but this is the rough idea.

8. Negative-gap cases

I turn now to the first step in my argument, which is to that optimal codes require suboptimal results in some cases.

This might seem like an unpromising starting point. If a code were truly optimal, would it not only require optimal results? This was Cesare Beccaria's thought. It was the reason for his scepticism towards pardon powers. Beccaria wrote:

> [Pardons] are one of the noblest prerogatives of the throne, but, at the same time, a tacit disapprobation of the laws. Clemency

[51] This slippery slope argument is one of the main objections to the legalisation of assisted suicide in both policy and philosophical circles. For an overview, see MP Battin, 'Euthanasia and Physician-Assisted Suicide' in H LaFollette, *The Oxford Handbook of Practical Ethics* (OUP 2005) 682-6.

> 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.[52]

Pardons may be a legitimate means of correcting for suboptimal results. But an optimal code does not require suboptimal results. So, under an optimal code, there is no place for pardons. Or, as Andrew Novak puts it: 'If a pardon is just, the law must be wrong; if the law is just, a pardon must be wrong'[53].

A. The price of accuracy

To see why optimal codes will require suboptimal results, it helps to recall an old debate about rule consequentialism.

Rule consequentialists say that an act is right if and only if it is required by the set of rules with the best consequences. But consider a rule which requires an act which does not have the best consequences. It seems that there must be a better rule, namely, one just like the first except that it does not require the act in question. We can lay the same charge every time rule consequentialism prescribes an act without the best consequences. So, critics said, rule consequentialism collapses into act consequentialism, ie, the theory that an act is right if and only if it has the best consequences.[54]

Rule consequentialists responded by distinguishing two versions of their theory. A simple version assesses rules by the consequences of adherence to them. This version may be vulnerable to the criticism in the last paragraph. But a sophisticated version assesses rules based on the consequences of their acceptance or internalisation.[55] A set of rules extensionally equivalent to act consequentialism would be very complicated and thus very costly to accept or internalise. These costs mean that it would be better to

---

[52] C Beccaria, *An Essay on Crimes and Punishments* (James Donaldson 1788) 167-8.
[53] Novak (n 33) 3.
[54] For a detailed explanation and two variants of the criticism, see B Hooker, *Ideal Code, Real World* (OUP 2000) 93-99.
[55] Ibid.

accept a simpler set of rules. By hypothesis this simpler set is not extensionally equivalent to act consequentialism.

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 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, contra Beccaria, an optimal code requires a suboptimal result, or as I will also say, there is a negative-gap case.

B. The limits of equitable interpretation

Negative-closure cases can be remedied by equitable interpretation, as I said. But things are different when it comes to negative-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 market clutter the area and prevent delivery vehicles making deliveries.[56]

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.

[56] 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.

> When, precisely, does the market stop being in process? What is near the market? What counts as a delivery vehicle? [57]

What Evans calls 'endless argument' leads to all three of the costs I described in the last section. 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". [58]

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'[59]? These are 'not cases for an equitable exception'[60] 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'[61]. The law is optimal though it yields a suboptimal result. Narrowing the law through interpretation would produce a suboptimal code. 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'[62].

---

[57] Ibid.
[58] Ibid.
[59] Ibid.
[60] Ibid.
[61] Ibid.
[62] 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 ('It is a platitude that justifiable rules can sometimes have morally regrettable applications'); F Schauer (n 38) 132-33; Brand-Ballard (n 33) 83.

9. Deviation and adherence

How should we respond in negative-gap cases? Not by interpreting them 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 first strategy is to set aside or 'deviate from' the code in *all* negative-gap cases. The second is to adhere to the code in all negative-gap cases. And the third strategy, the one I favour, is to selectively deviate – to deviate in some cases and adhere in others.

A. Bentham's bold idea

Bentham proposed a version of the first strategy. He proposes a system of code-making and adjudication with two elements.[63] 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'[64]. 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; it is 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, ie, in every negative-gap case.

---

[63] The elements of Bentham's theory are found in …. My account of Bentham's theory is based on the account in G Postema, *Bentham and the Common Law Tradition* (OUP 1986).

[64] Postema (n 63) 448.

B. Postema's problem

I find Bentham's proposal fascinating, but Gerald Postema points out a fatal flaw.

A legal code serves 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'[65]. 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'[66]. Moreover, because Bentham insists that judicial 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'[67].

Could Bentham reply that because the decisions of courts have no law-making effect, they will not be treated as a source of guidance? Postema anticipates this response. His answer is that judges will inevitably strive for consistency and coherence in their decisions.[68] 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'[69]. 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.

C. Secrecy

Bentham insists on consistent and public deviation in negative-gap cases. His proposal fails, and if Postema is right, one reason is the publicity requirement. Could we improve on Bentham's idea by

---

[65] Ibid 454.
[66] Ibid.
[67] Ibid.
[68] Ibid 457.
[69] Ibid 454. Postema also says that people will inevitably interpret the code in light of judicial pronouncements on its meaning.

dropping the publicity element? The idea is that so long as we deviate in secret our deviations will not destabilise the code.[70]

It seems to me that there are two serious problems. 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 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.

Second, given secrecy, consistency is hard to maintain. A secret practice is by nature difficult 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.

D. Adherence

If we should not deviate from a code in *all* case of a suboptimal type, publicly or not, it might seem natural that we should adhere in all cases. But this is to swing too far the other way. As Alan

[70] 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.

Goldman says, a single decision 'on moral merits instead of law will have little effect on the stability and predictability of the legal system'[71]. 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.

E. Selective deviation

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

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 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 negative-gap cases is worse than deviating in none; and second, that deviating in no negative-gap cases is worse than deviating in some. Hence the optimal deviation rate is between 1 and 0.

This is a claim about deviation across all negative-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 *negative-gap case-type* is a case-type in which an optimal code requires a suboptimal result.[72] If Postema is right, then we do not want to deviate in all tokens of a negative-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 negative-gap case-type either, because deviation in a single case has a significant benefit and a trivial cost. This tells us

[71] A Goldman, *Practical Rules: When We Need Them and When We Don't* (CUP 2002) 42. See also Brand-Ballard (n 33) 130 ('a particular deviant decision compromises predictability values only to a minimal extent').
[72] Adapted from Brand-Ballard (n 35) c 13.

that the optimal deviation rate in any negative-gap case-type is between 1 and 0.[73]

That may sound innocuous, but it has an important implication. Call two cases *alike* when they are tokens of the same case-type, ie, when they have the same legally and morally relevant features. The conclusion in the last paragraph is that, in token cases of a given negative-gap case-type, we ought to deviate in some and adhere in others. It follows that we ought to treat like cases unalike. 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.[74] 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 negative-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.

---

[73] Paley may have favoured a similar idea. In a discussion of capital punishment, he says that the punishment ought to be administered selectively, even for the same offence. 'By this expedient, few actually suffer death, whilst the dread and danger of it hang over the crimes of many. The tenderness of the law cannot be taken advantage of. … [N]o one will adventure upon the commission of any enormous crime, from a knowledge that the laws have not provided for its punishment'. W Paley, *The Principles of Moral and Political Philosophy*, first published 1785 (Liberty Fund 2002) xx.

[74] Assume neither is more needy, deserving, etc.

10. Paradox dissolved

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

I said that a pardon power can only be used in the direction of leniency. If our aim is to selectively deviate from a code, due to its overbreadth, that is all we want to do. I said that a pardon sets aside or dispenses with the law, without changing it. Again, that serves our aims. For we do not want to change optimal rules; that would only make them worse. We want to set them aside in particular cases, while leaving them intact for use in future cases, which is precisely what a pardon does.

A. The rule of law

A pardon power in its traditional form is wholly discretionary. The power is not subject to general standards, and as a result, its exercise is arbitrary. At first this sounds like an undesirable characteristic, but it is perfect 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.

Judges have seen things differently, of course. Krishna Iyer J. in *Maru Ram* said that 'no legal power can run unruly', 'but must keep sensibly on a steady course'[75]. Suppose we listened to this warning and subjected a pardon power to a system of rules, laid down in advance, specifying when the power is to be used. The pardon power would no longer be arbitrary, and Krishna Iyer J would say it is a victory for the rule of law.

But this gets things exactly backwards: it would be a defeat for the rule of law. The rule of law favours rule-based decision-making because normally that leads to predictability and stability. With respect to pardon powers, it is the other way around. If citizens

[75] *Maru Ram* (n 10) 62.

could count on the law being set aside, it would undermine the code. So, if we are to uphold the code, pardons must be granted inconsistently. To be granted inconsistently, pardons must not be required under conditions set out in advance. Pardons should 'be like lightning bolts, relatively rare and in principle hard to predict'[76]. Ironically, then, the rule of law itself insists that a pardon power be not ruled by law.[77]

B. The separation of powers

There are two reasons to allocate a pardon power to a branch other than the judiciary. The first is Postema's point: 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.

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 negative-gap case depends on what other judges will do in *their* negative-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.[78]

I do not claim that this problem is *impossible* to solve. There may be ways for judges to coordinate their activities, with enough effort and ingenuity.[79] But there is a simpler way to coordinate: allocate

---

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

[77] This is not to say that the rule of law favours a pardon power (of any form) over no pardon power.

[78] Goldman argues that judges will struggle to solve exactly this sort of coordination problem: Goldman (n 71) 42-46.

[79] Brand-Ballard (n 35) c 10.

the power to pardon to a relatively unified, centralised authority. And this, of course, is what actual pardon powers do. They grant the power to the legislature or executive, both of which are designed to make collective decisions. "Coordination" is facilitated by reducing the effective number of actors.[80]

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. So, the separation of powers in fact *favours* giving pardon powers to the legislature or the executive.

C. Judicial review

If pardons are a mechanism for selective deviation, then there is no place for judicial review of pardon refusals. With respect to substantive (merits-based) review, the explanation is simple: there is no case in which it is irrational or incorrect to refuse a pardon. Even when the grant of a pardon is rational, a refusal is also rational in a like case. What about procedural review? O'Connor J said that it should be unlawful for 'a state official [to] flip a coin to determine whether to grant clemency'[81]. This was supposed to be the minimal procedural safeguard, the least that could be required. But, if a pardon power is a means of approximating the optimal deviation rate, then flipping a coin may be an excellent way to decide whether to grant a pardon in one negative-gap case versus another. We should not demand this minimal safeguard. A fortiori, we should not demand any more rigorous procedure.

So far, I have been discussing review of pardon *refusals*. Grants are different. If a power-holder purports to grant a pardon in a case which is *not* a negative-gap case, then it may be legitimate for judges to intervene. If this sounds strange, it may help to know that for

---

[80] See also: A Perry, 'Mercy and Caprice Under the Indian Constitution' (2017) 1 Indian Law Review 56, 65.
[81] *Woodard* (n 28) 281.

many centuries pardons in England were granted on something like this basis. Pardons were typically granted on the recommendation of a trial judge, who was in a better position to know whether the law had produced a bad result.[82] Judges determined which cases were suboptimal, leaving it for the executive to determine in which of those cases to grant a pardon.

Indeed, my analysis favours judicial review of pardon decisions on at least one ground. The "rule against fettering discretion" says, roughly, that decision-makers should not adopt rules as to the use of their discretionary powers.[83] Were a power-holder to adopt a rule setting out when a pardon would be granted, it would undermine the purpose of the power. Judicial intervention would be justified as a result. Thus, judges should intervene to make sure that pardons can never be relied upon

11. Battered women and self-defence

Let me end with an illustration of how my proposal might work in practice.

Jacqueline Sauvage was married to her husband, Bernhard Marot, for 40 years. During that time, Marot abused Sauvage, physically, emotionally, and sexually. The couple had three daughters and a son. Marot husband most likely raped the daughters and abused the son. In 2012, the son committed suicide. The next day, while Marot was sitting on their porch, Sauvage picked up a rifle and shot him three times in the back, killing him. Sauvage was convicted of murder and sentenced to 10 years in prison. But there was a public outcry at Sauvage's plight. Eventually a petition for a presidential pardon received 400,000 signatures. President Hollande, in 2016, granted Sauvage a full pardon, and she was freed shortly after.

---

[82] H Lacey, *The Royal Pardon: Access to Mercy in Fourteenth-Century England* (York Medieval Press 2009) xx; JM Beattie, *Crime and the Courts in England 1660-1800* (OUP 1986) xx. For an account of a similar practice in the United States and its formalisation in Justice Department guidance, see M Colgate Love, 'The Twilight of the Pardon Power' (2010) 100 Journal of Criminal Law and Criminology 1169, 1175-6.

[83] See A Perry, 'The Flexibility Rule in Administrative Law' (2017) 76 CLJ 375.

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, for example, meant to hurt her husband but not, the court found, to kill him. And so she was given a 5 year sentence initially, 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.[84]

These are like cases, treated differently. Is there any justification for this inconsistent treatment?

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. And the same was true of Dimet: the threat she defended herself against was not imminent. Let us 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,

[84] My account of the two cases is drawn from K Fitz-Gibbon and M Vannier, 'Domestic Violence and the Gendered Law of Self-Defence in France: The Case of Jacqueline Sauvage' (2017) 25 Feminist Legal Studies 313.

> 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 violence ....[85]

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'[86] which will end with sanctioning vigilantism.

Let me 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 let me 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.[87] It follows that we have suboptimal results required by an optimal rule. That is, Sauvage's case and Dimet's case are negative-gap cases. How, then, should we deal with such cases?

We should *not* adhere to the imminence requirement in every negative-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. Judicial review would limit the potential for this kind of inconsistency. So, it is a virtue that in France (unlike the jurisdictions I surveyed in section 4), the pardon power remains a wholly discretionary 'act of conscience'.

To be clear, my claim is that, on certain assumptions, we ought to selectively deviate from the imminence requirement in battered women's cases. These are, to repeat, assumptions. I do not think they are wild or crazy asssumptions, but they could well be wrong. So I do not claim that, in fact, selective deviation is the best way to deal with battered women's cases; for all I know, we would be

---

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

[86] Kaufman (n 85) 363.

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

better off avoiding these suboptimal results through some other means.

12. Conclusion

Pardon powers are ubiquitous, but in their traditional form they have seemed to conflict with both the rule of law and the separation of powers. But on reflection there is no conflict here. On the contrary, if we are to have a pardon power, the rule of law itself favours a power which is uncontrolled, so that its exercise is inconsistent. The judiciary is not well suited to exercise power inconsistently, so the separation of powers favours allocating the pardon power to the legislature or the executive. The features of pardon powers which seemed to be in tension with the rule of law and the separation of powers are in fact justified based on those same principles.