# The Performance of Legitimacy:

## Moral Authority, Institutional Corruption, and the Manufacture of Public Trust

Flyxion
Independent Researcher

November 2026

**Abstract**

Political and corporate authority frequently depends upon representations of public benefit that are not themselves evidence of beneficial conduct. The appearance of decisiveness may be mistaken for competence, philanthropy for accountability, and the successful management of a crisis for proof that its apparent resolution was undertaken in good faith. This essay examines the distinction between legitimate authority and the performance of legitimacy through three cultural and institutional settings: the competing forms of leadership depicted in the second season of *Gotham*, the corporate paternalism of the *Fallout* franchise, and the financial narratives surrounding cryptocurrency promotion. These cases illuminate a recurring mechanism in which an actor's declared moral purpose becomes a substitute for independent evaluation of the actor's conduct. The argument is not that displays of authority or professions of benevolence are necessarily deceptive, but that neither establishes the legitimacy it purports to communicate. A formal result is offered to this effect: when a genuine and a manufactured mechanism induce the same public record, no test applied to that record can exceed chance accuracy, and when the two records merely resemble one another, the best available test exceeds chance only by an amount fixed by the distance between them. The essay then observes that it is itself a performance of the kind it describes, possessing a theorem, a bibliography, a table, and a measured tone, none of which establishes that it is correct. It accordingly does not support the claim that any particular institution is illegitimate, does not support the claim that any particular institution is legitimate, and does not support the claim that its own central proposition applies to any real case. These limitations are restated at intervals.

**Keywords:** legitimacy; institutional corruption; signaling; non-identifiability; pooling equilibrium; replacement test; goal displacement

# 1 Introduction: The Appearance of Moral Authority

A familiar scene in political and institutional fiction begins with catastrophe. An organization has failed, its personnel have been compromised, and its members are demoralized. A new leader arrives, denounces the failures of the previous administration, imposes discipline, and promises to restore integrity. The audience is invited to interpret the performance as evidence that someone has finally taken responsibility. The invitation is accepted so reliably that it has become a structural convention of the genre, and the convention is worth examining precisely because it also operates, with fewer musical cues, outside the genre.

Yet the performance itself establishes very little. Anger does not demonstrate competence, decisiveness does not establish correctness, and the willingness to punish subordinates does not show that the underlying causes of institutional failure have been understood. The sociological literature has long distinguished the formal structure an organization displays from the activity it conducts, and has noted that conformity to the expectations of an audience can be pursued as an end in itself, sometimes at the expense of the work the expectations were meant to secure [9]. Goffman's treatment of the presentation of self supplies the individual-scale counterpart: an impression of competence is an accomplishment of staging, and the staging may or may not be accompanied by the competence [3].

The distinction becomes especially apparent in the second season of *Gotham*, where three figures present contrasting but structurally related performances of authority. Captain Nathaniel Barnes arrives at the Gotham City Police Department in the aftermath of a massacre carried out there by Jerome Valeska and the Maniax, and introduces a disciplinary regime in which the department is run as a military unit. Oswald Cobblepot consolidates his criminal organization through intimidation. Theo Galavan cultivates public standing by appearing to save the city at a charity gala, while privately pursuing the mayoralty and a program of large redevelopment projects. Their stated objectives differ, but their methods invite comparison. The criminal leader, the police reformer, and the aspiring civic benefactor each attempt to establish a relationship in which personal authority appears indispensable to collective survival.

The resulting question is not whether these characters possess identical moral standing. They plainly do not. It is whether the mechanisms through which authority becomes publicly credible can operate independently of the conditions that would actually justify that authority. The sections that follow argue that they can, state the claim in a form that admits proof, and then turn the apparatus of proof upon the essay that contains it.

*Statement of Non-Support.* This section does not establish that any leader, fictional or real, lacks legitimate authority. It establishes only that the question is open in each case considered.

## 2 Three Performances of Leadership

The juxtaposition of Penguin, Barnes, and Galavan permits an analysis of three distinct mechanisms: coercive compliance, disciplinary legitimacy, and manufactured public confidence. Penguin's authority rests partly upon the credible threat of retaliation. His associates need not believe that his decisions are morally justified; they need only believe that disobedience carries consequences. His organization exposes coercion relatively directly, and for that reason it serves as the baseline against which the other two are read, since an audience that is afraid is not thereby deceived, whereas the other two performances operate on belief and not on fear.

Barnes presents a more ambiguous case. He possesses formal institutional authority, and his stated opposition to police corruption is not intrinsically objectionable. Nevertheless, his confrontational manner raises the possibility that the outward performance of discipline can be confused with the substantive work of institutional repair. Removing personnel and issuing commands may be necessary in some circumstances, but neither action establishes that corruption's enabling conditions have been eliminated. A department from which the corrupt have been removed is, by that fact alone, a department from which the corrupt have been removed, and the further claim that it is a department in which corruption is now difficult requires a separate demonstration. The department's difficulties, as the series presents them, include organized crime, political interference, intimidation, and repeated failures of accountability, none of which is obviously susceptible to repair by a commanding voice. It is also worth noting the sequence of priorities on Barnes's arrival. The city has just suffered a catastrophic failure of public security, and the first act of the new captain is to establish dominance over the surviving personnel, so that the spectacle of command precedes any inquiry into how the catastrophe occurred. Penguin and Barnes thereby exhibit a structural parallel that their opposite professed purposes conceal. Penguin need not persuade his associates that his decisions are reasonable, only that disagreement is costly, and Barnes, who holds the institutional legitimacy of a police captain, opens with a performance that likewise emphasizes submission over deliberation.

Galavan introduces a further mechanism. He seeks authority through the public interpretation of events, cultivating the appearance of heroism and civic responsibility while concealing his involvement in the circumstances that produce his reputation. In the series he is responsible for the crisis involving Jerome and the Maniax, is then received as the man who stopped Jerome at the charity gala, and subsequently takes Penguin under his wing as an instrument of political influence. The phrase is offered without comment on its ornithological properties, except to record that the recipient of the protection is a flightless bird and the protector is a man whose ambitions are about to take off, and that the essay does not support any inference from this coincidence. Penguin attempts to secure obedience, Barnes attempts to impose order, and Galavan attempts to produce consent. These are different relationships between power and its audience, even when their theatrical characteristics

overlap. Weber's classification of authority by the grounds on which obedience is rendered is useful here, since each of the three figures draws upon a different combination of fear, office, and charisma, and the combinations can be recognized without any prior conclusion about whether the grounds are warranted [15].

A methodological caution is required at this point. Fiction supplies its audience with the one datum that an audience in real institutions almost never possesses, namely the hidden state of the principal characters. The viewer of *Gotham* is told, by the structure of the narrative, which of the three is plotting and which is merely severe. The real observer of a real reformer receives no such disclosure. The television series is therefore a useful source of clear examples and a poor model of the evidentiary situation, because it resolves in the first act the very uncertainty whose persistence constitutes the problem. The cases are used below as illustrations of mechanisms and not as evidence about their prevalence.

*Statement of Non-Support.* Nothing in this section supports the proposition that real reformers resemble Captain Barnes or that real benefactors resemble Theo Galavan. The resemblance, where it exists, is formal and was selected by the author.

## 3 The Manufacture of Necessity

Galavan's political ambitions illustrate a particularly consequential possibility: an actor may acquire legitimacy by appearing to solve a problem whose existence or severity the actor has helped produce. This mechanism differs from ordinary hypocrisy. A hypocrite fails to practice a professed principle. An actor who manufactures necessity instead manipulates the conditions under which others evaluate the need for that actor's authority. The general pattern has been discussed in the political economy of disaster, where a period of public disorientation is described as an opportunity for measures that would not otherwise command assent [4].

The sequence can be expressed as a causal pattern. A crisis creates public uncertainty; a conspicuous intervention provides a recognizable protagonist; the intervention becomes evidence of exceptional competence; and the resulting reputation supports an expansion of authority. In Galavan's case the resulting authority is sought through urban redevelopment, which is instructive because physical construction supplies an ostensibly concrete and public-spirited justification for the acquisition of political power while the underlying arrangements involve coercion and private interest. Each step is individually unremarkable, and each can occur without any wrongdoing. Crises do occur naturally, interventions are sometimes excellent, and reputations are sometimes earned. The sequence is therefore a template that honest and dishonest actors can both follow, which is exactly the property that makes it dangerous as a basis for inference.

The inference from successful intervention to legitimate authority is invalid without additional evidence. Observers must establish whether the intervention was effective, whether the actor contributed to the crisis, whether alternatives existed, and whether the proposed

expansion of authority is proportionate to the demonstrated achievement. This is the central epistemic vulnerability of crisis-based leadership. The most visible solution may not be independent of the problem, and the more visible the solution, the more attention is withdrawn from the question of its origin.

*Statement of Non-Support.* This section describes a possibility. It does not support the claim that the possibility is frequently realized, and it supplies no method for determining when it has been.

# 4 Corporate Benevolence and the Administration of Humanity

The *Fallout* franchise, in both its games and its 2024 television adaptation, extends the problem from individual leadership to institutional continuity. Vault-Tec's public purpose concerns human survival, yet its fictional institutional arrangements repeatedly subordinate individual welfare to organizational experimentation, secrecy, and control. Preservation becomes an administrative objective whose beneficiaries may have little opportunity to contest the means by which they are supposedly being preserved. The important distinction is between the welfare of an institution and the welfare of those whom the institution claims to serve.

An organization can survive while its beneficiaries suffer. It can expand its administrative capacity while diminishing individual autonomy. It can preserve its declared mission while systematically frustrating the interests that originally justified that mission. This suggests a general institutional failure mode: the means of pursuing a public purpose become objects of preservation in their own right, eventually acquiring priority over the purpose itself. The mechanism is the one Merton described under the heading of goal displacement, in which adherence to the procedures that serve an end gradually replaces the end [8]. Scott's study of high-modernist planning describes the same pattern at the scale of states, where legibility to the administrator is purchased at the expense of the welfare of the administered [12].

A difference between the two series deserves emphasis. In *Gotham* the contradiction between professed purpose and conduct is embodied in individual personalities, namely Penguin, Barnes, and Galavan, each of whom can be removed from the story. In *Fallout* the contradiction is embedded in the organization, so that the system continues to reproduce its ideology when individual executives disappear or are replaced. The second arrangement is the one that bears most directly on the replacement test discussed in Section 8, because an institution whose outcomes persist when its executives are replaced is exactly the case the test is designed to detect, and the Fallout material is included for that reason and not merely as a second illustration. The fictional case has a further feature that sharpens the point. Vault-Tec's customers purchase a promise about a future catastrophe. Guarantees of this kind can in principle be supported before the event by engineering tests, inspection, and independent certification, so verification is not logically impossible. The difficulty lies in

access. The beneficiaries have no adequate means of obtaining such verification, and the credibility of the promise is therefore carried largely by presentation. The brochure is a substantial part of the product.

*Statement of Non-Support.* This section does not support the claim that any actual corporation resembles Vault-Tec. The corporation in question does not exist, and its documented conduct is that of a writers' room.

## 5 Financial Narratives and the Production of Credibility

The promotional culture surrounding cryptocurrency offers a non-fictional setting in which similar questions arise, although the empirical claims require independent documentation that this essay does not undertake. Promoters may associate financial products with decentralization, individual liberty, technological progress, or economic inclusion. Such aspirations are not evidence that any particular product delivers the corresponding benefits. Celebrity endorsements, institutional affiliations, philanthropic claims, and technological sophistication can all contribute to the appearance of credibility without establishing financial soundness or accountability.

Ben McKenzie's transition from portraying Jim Gordon in *Gotham* to writing, with Jacob Silverman, a book-length examination of cryptocurrency promotion, and then to directing the 2026 documentary *Everyone Is Lying to You for Money*, provides an incidental cultural connection between the fictional and documentary cases [6, 7]. The film, which includes appearances by figures from the industry it examines, is concerned with how promoters build credibility through celebrity endorsement, claims of technological progress, and narratives of financial liberation, which is the same family of devices that Galavan assembles from heroism and redevelopment. The coincidence of casting is amusing, since an actor known for a character whose defining struggle is to separate legitimate authority from the corruption concealed behind it later investigates a real industry whose promoters made expansive claims about freedom and prosperity, and the amusement is recorded here without being offered as evidence. The analytical relevance, however, lies in the evidentiary distinction and not in the casting history. Public confidence is an observable social outcome, whereas the justification for that confidence requires independent examination. The authors' account is itself an argument that must be assessed on its evidence, and the present essay takes no position on how that assessment should conclude. It is noted only that the structure of the problem, in which the observable is confidence and the unobservable is soundness, is the structure of the problem treated formally below.

*Statement of Non-Support.* This section does not support any claim about the soundness of any financial product, the sincerity of any promoter, or the accuracy of any book.

## 6 Legitimacy as an Evidentiary Problem

The preceding cases suggest that legitimacy should not be inferred directly from the success of its performance. Three questions must be distinguished, a separation that follows the broader tradition of treating legitimacy as a matter of rules, justifiability by shared beliefs, and expressed consent, each analytically separable from the others [1]. Suchman's survey of strategic and institutional approaches likewise distinguishes legitimacy as a resource that organizations pursue from legitimacy as a condition that audiences confer, and observes that the first can be cultivated without the second being deserved [14].

First, how does an actor obtain recognition as an authority? This is a descriptive question concerning persuasion, reputation, institutional position, and social coordination. Second, what justifies the actor's exercise of authority? This is a normative question concerning authorization, rights, obligations, procedural constraints, and the interests of affected parties. Third, what consequences follow from that exercise? This is an empirical question concerning outcomes, causal attribution, and the persistence or elimination of the conditions the authority was supposed to address.

These questions may yield different answers. A leader can be popular without being justified, formally authorized without being effective, or effective in a limited intervention without acquiring a general entitlement to govern. The central error occurs when evidence answering one question is treated as sufficient to answer the others. The error is rarely committed deliberately by audiences. It is more often induced by performers, for whom the conversion of an answer to the first question into an apparent answer to the second and third is the principal object of the performance.

The table below sets out several recurring performances, the inference each invites, and the evidence the inference would actually require. The final row is included for completeness.

*Statement of Non-Support.* The table does not support the claim that the listed inferences are always unwarranted. In many cases the evidence in the third column exists and the inference is correct.

### An Older Formulation

The problem is not recent. In the second book of the *Republic*, Glaucon asks what distinguishes justice from the successful performance of justice, and constructs the case to make the question unanswerable by appearances: the perfectly unjust man who holds the reputation of perfect justice is set beside the perfectly just man who holds the reputation of injustice, so that conduct and reputation have been entirely separated [11]. The challenge to Socrates is to show that justice is worth having when it is accompanied by neither recognition nor reward, and that the unjust man with the best reputation is nonetheless worse off. The formal result of the next section is a statement of the first half of Glaucon's setup in modern terms. If the reputations are the same, nothing an observer can do with the reputations will

| Performance | Inference invited | Evidence the inference requires |
|---|---|---|
| Disciplinary severity | Reform | Removal of the conditions that enabled the original failure, observed after the disciplinarian departs |
| Philanthropic spectacle | Benevolence | Independent measurement of outcomes for the stated beneficiaries |
| Technological sophistication | Soundness | Verifiable accountability for losses, assignable to named parties |
| Crisis resolution | Competence | Evidence that the intervention was effective relative to alternatives; separately, evidence on the actor's causal role in the crisis, which bears on good faith |
| Formal apparatus, theorem, and bibliography | Correctness | Not supplied by the present essay |

Table 1: Performances, the inferences they invite, and the evidence those inferences require.

separate the two men. Glaucon's remedy was an argument about the soul, and the present essay offers no argument about anyone's soul.

A related observation concerns Batman, the figure toward whom *Gotham* is, in a long deferral, directed. The series spends years on the social conditions that make such a figure imaginable, namely the loss of moral credibility by the institutions charged with administering justice. Batman is one possible response to that loss, and the narrative leaves open the uncomfortable question whether replacing accountable institutions with an exceptionally powerful individual addresses the underlying problem or reproduces it in a more theatrical form. Superhero narratives are commonly read as morality plays in which villains embody exaggerated dispositions, such as intellectual vanity, social ambition, or a divided conception of justice. *Gotham* complicates the reading by locating the dispositions in a city whose supposedly legitimate institutions are themselves compromised, and the same complication is the reason the three performances of Section 2 cannot be sorted by their declared objectives.

## 7 A Formal Statement of Non-Identifiability

The informal claim that a reputation does not identify its own origin can be given a precise form. Let $\mathcal{O}$ denote a measurable space of public records, interpreted as everything an audience can observe about an actor, including statements, endorsements, visible interventions, and recorded approval. Let $\mathcal{M}$ denote a set of mechanisms, where each mechanism $m \in \mathcal{M}$ is a process that generates a public record and therefore induces a probability distribution $P_m$ on $\mathcal{O}$. Let $h : \mathcal{M} \to \{0, 1\}$ be a hidden indicator, with $h(m) = 1$ when the mechanism consists of conduct that actually justifies the authority claimed and $h(m) = 0$ when it

consists of conduct arranged to produce the appearance of such justification. The audience observes a record $o \in \mathcal{O}$ and applies a test $T : \mathcal{O} \to \{0, 1\}$, where $T(o) = 1$ means that the audience concludes the authority is legitimate.

**Definition 1.** *Fix a genuine mechanism $G$ with $h(G) = 1$ and a manufactured mechanism $F$ with $h(F) = 0$. The* true positive rate *of a test $T$ is* $\mathrm{TPR}(T) = P_G(T = 1)$*, and its* true negative rate *is* $\mathrm{TNR}(T) = P_F(T = 0)$*. The* balanced accuracy *of $T$ is* $\frac{1}{2}(\mathrm{TPR}(T) + \mathrm{TNR}(T))$*.*

**Proposition 1** (Non-identifiability of legitimacy from its performance)**.** *If $P_G = P_F$, then every test $T$ satisfies* $\mathrm{TPR}(T) + \mathrm{TNR}(T) = 1$*, so that every test has balanced accuracy exactly* $\frac{1}{2}$*.*

*Proof.* For any test $T$, $\mathrm{TNR}(T) = P_F(T = 0) = 1 - P_F(T = 1)$. Since $P_F = P_G$, this equals $1 - P_G(T = 1) = 1 - \mathrm{TPR}(T)$. Adding $\mathrm{TPR}(T)$ to both sides gives the claim. □

The proposition says that when a genuine and a manufactured mechanism generate the same public record, the record carries no information about which one produced it, and no amount of analytical ingenuity applied to the record alone can change this. The condition $P_G = P_F$ is strong, and a weaker form is more realistic.

**Proposition 2** (Optimal discrimination is fixed by distance between performances)**.** *Let $\nu$ be a measure dominating both $P_G$ and $P_F$ (for instance $\nu = P_G + P_F$), with densities $p_G$ and $p_F$, and let* $\mathrm{TV}(P_G, P_F) = \frac{1}{2}\int |p_G - p_F|\, d\nu$ *denote total variation distance. Then every test $T$ satisfies* $\mathrm{TPR}(T) + \mathrm{TNR}(T) \le 1 + \mathrm{TV}(P_G, P_F)$*, with equality for the test $T^*(o) = 1$ if and only if $p_G(o) > p_F(o)$, where the decision on the set $\{p_G = p_F\}$ is arbitrary. Consequently the supremum of balanced accuracy over all tests equals* $\frac{1}{2}(1 + \mathrm{TV}(P_G, P_F))$*.*

*Proof.* Write $\mathrm{TPR}(T) + \mathrm{TNR}(T) = 1 + P_G(T = 1) - P_F(T = 1) = 1 + \int_{\{T=1\}}(p_G - p_F)\, d\nu$. The integral is at most $\int (p_G - p_F)^+ d\nu$, since discarding the negative part of the integrand can only increase it. Because $\int (p_G - p_F)\, d\nu = 0$, the positive and negative parts have equal integrals, so $\int (p_G - p_F)^+ d\nu = \frac{1}{2}\int |p_G - p_F|\, d\nu = \mathrm{TV}(P_G, P_F)$. Equality holds when $\{T = 1\} = \{p_G > p_F\}$, which is the event on which the integrand is positive. Dividing the sum by two gives the statement about balanced accuracy. □

The second proposition makes explicit what a sophisticated manufacturer of appearances is doing. The manufacturer's objective is to reduce $\mathrm{TV}(P_G, P_F)$, that is, to make the manufactured record resemble the genuine one. Every increment of successful imitation lowers the ceiling on the best achievable discrimination, and the ceiling is not raised by the audience becoming more attentive. The statement concerns the ceiling and not the performance of any particular observer. An audience using a poor test may well do better or worse as the two records converge, and the proposition says only that no audience can do better than $\frac{1}{2}(1 + \mathrm{TV})$. The proposition also has a respectable ancestor in the theory

of signaling, where a signal separates types only when producing it is more costly for the type it is meant to exclude, and where cheap or equally affordable signals produce a pooling outcome in which the audience learns nothing from the signal [13]. Public displays of virtue are inexpensive in many institutional settings, and the relevant question for any particular display is whether it is costly for the manufactured type to produce.

The underlying logic is the logic of identification in econometrics, in which two structures that imply the same distribution of observables cannot be distinguished by any amount of data drawn from that distribution [5]. The causal formulation is equally familiar: the observational distribution does not determine the interventional one without assumptions that the data do not themselves supply [10]. Nothing in the argument is new. Its application to the reputations of leaders is offered only because the audience for reputations so frequently behaves as though the result were false.

*Remark* 1. The propositions are conditional. They assert nothing about whether $P_G$ and $P_F$ are close for any actual pair of mechanisms. Where the manufactured performance is crude, $\mathrm{TV}(P_G, P_F)$ is large and discrimination is easy, and the historical record contains many examples of exactly this. The result is relevant to the cases in which imitation is good, and the essay does not claim to know in which cases that is.

*Statement of Non-Support.* Propositions 1 and 2 do not support the claim that any observed reputation was manufactured. They are compatible with every reputation in the world having been earned.

## 8 The Replacement Test

One way to distinguish personal performances from institutional improvements is to ask what happens when the celebrated leader is replaced. If corruption returns immediately after a forceful reformer departs, the reform may have depended upon personal surveillance or intimidation rather than durable institutional constraints. If a philanthropic organization can continue violating its beneficiaries' interests after its executives are replaced, the problem cannot be adequately explained by the character of those executives alone. If public confidence depends upon the continued repetition of a founder's heroic narrative, it may be necessary to ask whether the institution has developed independent means of demonstrating its value. The test is developed at greater length in a companion essay on institutional incentives and the recurrence of harm [2].

The test can be stated within the framework above, provided that it is stated causally. Let an institution consist of a set of constraints $C$ and an occupant $A$ drawn from a population $\mathcal{A}$ of possible occupants, and let $Y$ denote an outcome of interest such as the incidence of harm to beneficiaries. Write $P_{a,c}(Y) = P(Y \mid \mathrm{do}(A = a), \mathrm{do}(C = c))$ for the distribution of $Y$ when the occupant and the constraints are each set by intervention. The replacement test compares $P_{a,c}$ across occupants $a$ while holding $c$ fixed.

**Proposition 3** (The replacement test locates attribution and not valence)**.** *Fix constraints $c$ and suppose $P_{a,c}(Y) = P_{a',c}(Y)$ for all $a, a' \in \mathcal{A}$. Then the occupant has no distributional causal effect on $Y$ under $c$. This conclusion holds equally when the common distribution is favorable and when it is unfavorable.*

*Proof.* A distributional causal effect of the occupant is a dependence of $P_{a,c}(Y)$ on $a$. Invariance in $a$ means there is no such dependence. The value of the common distribution does not enter the definition of invariance, so the conclusion is unaffected by whether that value is desirable. □

*Remark* 2. The interventional form is essential. Observed outcomes may be invariant across occupants because leaders are systematically assigned to institutions of a particular kind, or because an institution adjusts its constraints in anticipation of a particular arrival, and in either case observational invariance would not establish that replacement is inconsequential. Distributional invariance also leaves open that two occupants produce different outcomes for particular beneficiaries while the aggregate distribution is unchanged, so that an institution may be invariant under replacement at one level of observation and conceal substantial change at another.

The proposition has an uncomfortable corollary, which is that a reliably harmful institution passes the replacement test in the sense of reproducibility. An institution that harms its beneficiaries regardless of who administers it has, in a technical sense, achieved perfect independence from the character of its leadership, and this is a property that reformers frequently say they wish to obtain. The same reasoning suggests why reforms that merely replace personnel may be indistinguishable, at the distributional level, from no reform at all. The test is therefore diagnostic of where responsibility resides and silent on whether the responsibility is being discharged well. A passing score should be read as an invitation to examine the constraints $C$, and not as a verdict on them.

*Statement of Non-Support.* The replacement test does not establish legitimacy. It does not support the claim that any institution has passed or failed it, since no replacement is performed in this essay.

## 9 The Present Essay as a Performance

An argument that public displays of competence do not establish competence must, if it is consistent, accept the same limitation for itself. The present essay has a title with a colon, an abstract, numbered sections, three propositions with proofs, a table, a bibliography in headline case, and a prevailing tone of caution. Each of these is a recognized instrument for producing the impression of rigor. By the reasoning of the preceding sections, the impression is a record $o$, the soundness of the argument is a hidden indicator $h$, and a reader who concludes from the apparatus that the argument is sound is applying a test $T$ whose accuracy

depends on how closely careful essays and careless essays resemble one another in their apparatus. The resemblance is, on inspection of the genre, considerable.

The proofs do not escape the difficulty, though the nature of the difficulty should be stated carefully. The propositions are correct, and their correctness can be checked independently of the author. The second in particular is a genuine characterization of optimal binary discrimination and not an artifact of convenient premises. What the proofs do not supply is any reason to think that the quantity they bound is small for a real pair of mechanisms, or that the replacement test has been run on a real institution. The essay has proved something, but it has not proved that what it proved is what its readers wanted to know. A theorem can be correct, well presented, and almost entirely irrelevant to the empirical proposition for which it is rhetorically recruited, and a reader who finds the formal results persuasive about institutions has been persuaded by the form, which a manufactured argument would also display.

The repeated Statements of Non-Support introduce a further complication. Their presence may indicate a genuine effort to delimit the argument, or they may function as a system of rhetorical indemnification, allowing the author to suggest conclusions through the surrounding prose while disclaiming responsibility for their adoption. The reader cannot determine which function they serve by counting them. A series of disclaimers is also a performance in its own right, namely the performance of epistemic restraint, and restraint is among the most reliable ways of acquiring a reputation for responsibility without incurring the costs of the responsible acts. The essay therefore does not support the claim that its Statements of Non-Support constitute evidence of intellectual honesty. This observation is itself presented in the customary language of intellectual honesty.

It follows that the essay does not support its own central thesis in application to any real case, and the following consequences are stated plainly so that they need not be inferred. The essay does not support the claim that any named leader, corporation, promoter, or institution is illegitimate. It does not support the claim that any named leader, corporation, promoter, or institution is legitimate. It does not support the claim that the mechanisms described in Sections 2 through 5 are common, rare, or present anywhere outside the works of fiction from which the examples were drawn. It does not support the claim that the replacement test is a good test, only that it answers a particular question. Finally, it does not support the claim that this statement of limitations is itself sincere and not a further performance of intellectual modesty, a possibility that the framework in Section 7 is unable to exclude and that the author is content to leave open.

The reader is accordingly in the position the essay describes. A document has presented a record, and the reader must decide, without access to the hidden indicator, how much weight the record deserves. The only remedy the preceding analysis offers is the one it offers in every other case, which is to seek evidence that is expensive for a manufactured argument to produce. In the case of an essay, such evidence would include independent verification of its sources, attempts at refutation by hostile readers, and a record of predictions that could

have failed. None of these is supplied here.

*Statement of Non-Support.* This section does not support the claim that the essay is unsound. It supports only the claim that the essay's apparatus cannot be cited in favor of its soundness.

## 10 Conclusion: Beyond the Heroic Administrator

The enduring appeal of the heroic administrator lies in the promise that institutional complexity can be overcome through exceptional personal character. The strong leader will impose discipline, the benefactor will provide resources, the entrepreneur will innovate, and the reformer will eliminate corruption. These promises are narratively attractive because they transform diffuse institutional problems into conflicts between recognizable individuals. Yet an institution's legitimacy cannot be reduced to the reputation, intentions, charisma, or disciplinary temperament of its leadership. The decisive questions concern what authority permits, what constrains it, what evidence supports its claims, and whether its beneficiaries retain meaningful opportunities to challenge its exercise.

The moral lesson of *Gotham* is therefore not simply that apparently respectable leaders may conceal criminal ambitions. It is that the social mechanisms through which people recognize authority can be manipulated independently of the conditions that make authority legitimate. The corresponding lesson of *Fallout* is that an institution can preserve the language of public benefit long after its operating practices have ceased to serve that benefit. In both cases the performance of legitimacy becomes most dangerous when it is accepted as a substitute for its demonstration.

The formal content of the essay can be summarized in a sentence. Observational success in producing public confidence does not identify the legitimacy of the process that produced it, because genuinely beneficial conduct and strategically manufactured appearances can generate the same public reputation, and the better the imitation, the less any observer can learn. The argument is about non-identifiability and not about fraud. It does not treat every forceful leader as suspect or every philanthropic institution as exploitative, and it asks only that public displays of virtue, competence, or benevolence be recognized as insufficient, by themselves, to determine whether the underlying institution deserves trust.

The essay has not shown that any such institution does or does not deserve trust, has not shown that its own propositions apply to any real case, and has not shown that it is correct. It has shown that it is possible to write at some length, with a theorem, about the difference between appearing to have shown something and having shown it.

## References

[1] Beetham, David. *The Legitimation of Power*. London: Macmillan, 1991.

[2] Flyxion. "Institutional Incentives and the Recurrence of Harm." https://standardgalactic.github.io/alignment/institutional-incentives.pdf, November 2026.

[3] Goffman, Erving. *The Presentation of Self in Everyday Life*. New York: Doubleday, 1959.

[4] Klein, Naomi. *The Shock Doctrine: The Rise of Disaster Capitalism*. New York: Metropolitan Books, 2007.

[5] Koopmans, Tjalling C. "Identification Problems in Economic Model Construction." *Econometrica* 17, no. 2 (1949): 125–144.

[6] McKenzie, Ben, and Jacob Silverman. *Easy Money: Cryptocurrency, Casino Capitalism, and the Golden Age of Fraud*. New York: Abrams Press, 2023.

[7] McKenzie, Ben, dir. *Everyone Is Lying to You for Money*. Documentary film, 90 min. The Forge, 2026.

[8] Merton, Robert K. "Bureaucratic Structure and Personality." *Social Forces* 18, no. 4 (1940): 560–568.

[9] Meyer, John W., and Brian Rowan. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." *American Journal of Sociology* 83, no. 2 (1977): 340–363.

[10] Pearl, Judea. *Causality: Models, Reasoning, and Inference*. 2nd ed. Cambridge: Cambridge University Press, 2009.

[11] Plato. *Republic*. Translated by G. M. A. Grube, revised by C. D. C. Reeve. Indianapolis: Hackett, 1992.

[12] Scott, James C. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven: Yale University Press, 1998.

[13] Spence, Michael. "Job Market Signaling." *Quarterly Journal of Economics* 87, no. 3 (1973): 355–374.

[14] Suchman, Mark C. "Managing Legitimacy: Strategic and Institutional Approaches." *Academy of Management Review* 20, no. 3 (1995): 571–610.

[15] Weber, Max. *Economy and Society: An Outline of Interpretive Sociology*. Edited by Guenther Roth and Claus Wittich. Berkeley: University of California Press, 1978.