



Robust Institutions

DAVID M. LEVY*

dlevy@gmu.edu

Department of Economics, Center for Study of Public Choice, George Mason University, Fairfax, VA 22030, USA

Abstract. I propose to show how to translate the economic analysis of institutions developed in the tradition of “worst case” political economy into the *lingua franca* of robust statistics. An institution will be defined as contingent upon a design theory and the difficulty we consider is the use of the institution by the designer.

The technical bridge between institutional robustness and statistical robustness is the possibility of exploratory data analysis [EDA] with respect to the design theory. In an institutional context with EDA comes the recognition that the model is not completely specified, that we do not fully understand the structure of world *before* studying it. In a statistical context with EDA comes a non-normal error distribution.

The relationship between evolutionary institutions and robust institutions is discussed. A conjecture that “rule utilitarianism” can be thought of as “robust utilitarianism” is defended with the historical example of William Paley’s discussion of the utility of murder.

Key Words: robust, institutional analysis, public choice, exploratory data analysis

JEL classification: c4, b1, h0.

Introduction

Here is a historical-analytical puzzle. Why is it that J.M. Buchanan’s worst-case philosophy of constitutional political economy sounds so similar to J.W. Tukey worst-case philosophy of mathematical statistics?¹ The use of heavy-tailed distributions as paradigm—without worrying overly much about descriptive accuracy as well as the “unrealistic” use of Leviathan models—are defended by the desire to avoid disaster.² Underlying this common concern is their questioning of the applicability of optimization considerations.³

I should think the common factor linking Tukey and Buchanan is the distant influence of the atheoretic view of human nature taught by the Greek historians.⁴ How people behave is not, in fact, in correspondence with their own theories of how they ought to behave.⁵ Buchanan’s insight is explicitly informed by the constitutional theory of such careful students of the Greek historians as Thomas Hobbes, David Hume and John Stuart Mill.⁶ A little more oblique is the fact that behind robust statistics in general stand the contributions of John von Neumann: minimax loss theory and monte carlo method. These two contributions buy insight into real problems by sacrificing any hope of a unique analytical solution.⁷ First, von Neumann’s minimax loss approach to decision making is absolutely

*The paper was first prepared for James Buchanan’s 80th birthday celebration. That version can be found at <http://www.uni-duisburg.de/FB1/PHILO/Buchanan/files/levy.pdf>. Thanks are due to Peter Boettke, Geoffrey Brennan, Bryan Caplan, Tyler Cowen and Maria Pia Paganelli for comments on previous drafts. I have benefitted from the suggestions of the *RAE*’s readers. The errors which remain are of course my responsibility alone.

central to robust thinking. One can see this impact in A. Wald's notion of admissibility⁸ and in G. Box's use of the term "robust" to describe estimators which, as Anscombe (1960) put it, "insure" against risk by giving up efficiency at the ideal case and in the development of explicit minimax estimation by P.J. Huber.⁹ Second, without the technology provided by the Ulam-von Neumann's monte carlo method there would be no Tukey-directed Princeton Robust Study.¹⁰ Minimax-loss-thinking works when we cannot even assign probabilities to outcomes. With monte carlo methods we can find solutions where our analytical powers fails. Tools for a world where our theoretical insight fails. This is von Neumann's gift to robust statistics.¹¹

And the link to the Greek historians? Von Neumann did not, as far as I know, write on the classics. But we are told that his thoughts, as he lay dying, turned to Thucydides' account of a particularly nasty moment in the war between Athens and Sparta.¹² The Greek historians are Berlin's hedgehog. They teach one lesson, the trick which always works. The world is not as we would have it. We ought to proceed as if *our* hopes are illusions.

What's Wrong with this Picture?

We now turn to develop an analytical linkage between Tukey's and Buchanan's worries about avoiding disaster. The bridge will be the recognition that our idealized characterization of the world could be false.¹³ Of course the topics which concern Tukey and Buchanan differ. Our abstracting device is a specification of a formal language, broad enough to express some of their common concerns, but narrow enough to be tractable.

Consider the following picture which describes the performance of two artifacts—1 and 2—as a function of some theoretical information. For point of reference, we indicate one possible state of the theory which we label τ . We know that when the theory attains state τ the performance of 1 is superior to 2. It is worthy of considerably reflection that serious discussions can be conducted considering *only* the performance of 1 and 2 at τ .¹⁴

A heavily studied example is the efficiency of a least squares estimate of a regression problem when the errors are independent normal. It would be hard to find a graduate econometrics text which does not develop the proof that, under this condition, the least squares estimate attains the Cramér-Rao lower bound and thus provides the best unbiased estimate.¹⁵

What happens when the normality condition fails is somewhat less heavily discussed in the texts.¹⁶ Indeed, the student can be excused for not realizing that small deviations from normality create the situation pictured in Figure 1. If T is the set of symmetrical unimodal distributions and τ marks the special case of the normal distribution, then the efficiency properties of least squares estimation are as pictured by 1 compared with robust regression techniques pictured by 2.¹⁷ The student generally does not come away from econometrics textbooks with the sure and certain knowledge that the efficiency gains from least squares vis-à-vis robust techniques are small at normality, but the efficiency costs can be arbitrarily large as the actual distribution deviates from this ideal.¹⁸

To develop intuition about the bridge between robust statistical considerations and constitutions, consider the late central planning debate, in which L. von Mises and F.A. Hayek stood against the profession. It was established to the satisfaction of almost all the participants in the debate, that a planned economy could be more efficient than a market

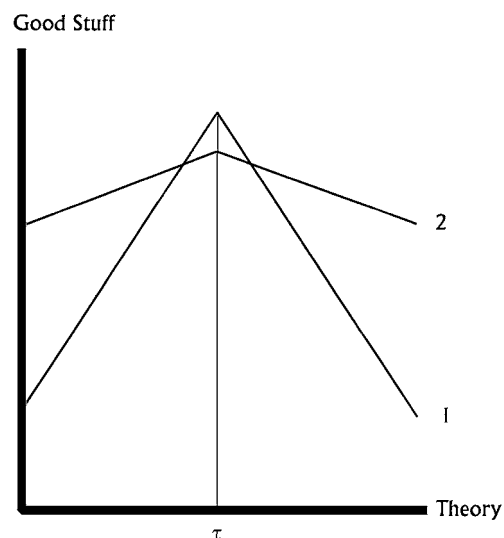


Figure 1. 2 is more robust than 1.

because the planned economy would not be troubled with the sorts of public choice considerations which plague actually existing markets.¹⁹ Perhaps the confidence with which economists pressed this case can be best judged by the widespread dissemination of the claim that for decades the planned economies' growth *exceeded* that of market economies.²⁰

If normality is the ideal condition for least squares—the τ in the picture—what serves as τ for central planning? The public spiritedness of the new socialist managers is our old friend: the benevolent despot assumption. Those who set prices would set aside any private interest to serve the public interest. And if this τ failed? As I read historical record of the central planning debate, discussion of this possibility is a little sparse.²¹ The point became obvious even to economists after the planned economies had fallen apart: the planners had used their theory to fool the public and to make themselves arbitrarily rich.

The design τ for central planning “forgot” to take into account of the interest of the theorist *cum* planner. The result was without doubt the greatest historical disaster with which economists share complicity.²²

I propose to generalize from this example with some technical machinery developed elsewhere to connect exploratory data analysis to non-normality.

Institutions and Design Theory

The building blocks of our account consist of a language T inside which we can write down a model which purports to describe an institution I . We shall suppose that each of the models of T is a linear estimating equation. The consistency condition we impose is that if τ is a model, it is identifiable.²³ Thus, for each institution we associate a design theory. This conditionality we write as $I \mid \tau$.

We define the institutional robustness with respect to the design theory τ in terms of a loss function. There exists some finite bound, m , on the difference between the institutional performance with respect to a metric U when τ holds— $U(I | \tau)$ —and when τ fails— $U(I | \sim\tau)$. Thus I is robust if and only if:

$$m < U(I | \tau) - U(I | \sim\tau).$$

It would be no surprise if a robust institution performs better when the ideal of τ holds than when it fails but that is not the point. The point is the loss from such failure is bounded. A nonrobust institution can be defined by reversing the inequality.

We must go beyond a formal language and say how the design theory τ could fail. Since we are assuming linearity and only requiring τ to be an estimating equation, and we are not all particular how it is estimated, failure of τ occurs when we leave a real variable out of consideration. The statistical casualness rests upon a supposition that there will be a competitive process at work in the estimation. If the estimation details matter, I presume this will be brought out in debate. Thus failure of τ is the same as model misspecification which is common to the language community. In the central planning example, since neither von Mises nor Hayek discussed the interests of planner *cum* economist to rig prices for their benefit—of course such a thought would never have occurred to the pro-socialist economists-planners²⁴—this consideration was given a zero weight. This is the fate of all variable dropped from the specification.

EDA is Possible iff τ Fails

Let us consider more precisely how the design theory can fail. Let us write the traditional linear model which expresses a random variable Y_t as a function of a finite number (K) of independent variables $X_{1t} \dots X_{Kt}$ and an error term ε_t . Thus for $t = 1 \dots T$

$$Y_t = \sum_{k=1}^K \beta_k X_{kt} + \varepsilon_t$$

Suppose that ε_t is normal. The Lindeberg-Feller conditions allow us to decompose a normal distributed random variable into normal and non-normal components. Those components which are normally distributed, we label η_t , and proceed to other concerns. Now, consider the infinite number (H) of non-normal components of ε_t which the Lindeberg-Feller conditions allow. If each of these random variables $X_{K+1t} \dots X_{Ht}$ is real valued then there must be infinitesimals $\delta_{K+1} \dots \delta_H$ such that we can rewrite ε_t in terms of the product of reals and infinitesimals. Thus, we obtain:

$$Y_t = \sum_{k=1}^K \beta_k X_{kt} + \sum_{k=K+1}^H \delta_k X_{kt} + \eta_t$$

While this formulation is not sufficient to establish the normality of ε_t , it is necessary. If any of the δ_k were non-infinitesimal and X_k were non-normal, then this contribution to ε_t would

have to be a real-valued, non-normal random variable thus violating the Lindeberg-Feller conditions.

The definition of EDA proposed in Levy (2000) is moving a variable from the infinitesimal δ list to the real β list. At T_{N+1} , but not at T_N , it may be possible to reject the hypothesis that β_{K+1} equals 0. We consider two types of EDA: local EDA and global EDA. Local EDA: for some $T_2 > T_1$ the model expands from K non-infinitesimal contributions to $K + 1$. Global EDA: in addition to local of EDA at T_1 , if local EDA at T_N then local EDA at T_{N+1} .

Levy (2000) establishes that, if EDA is possible, then the regression errors are non-normal. This is the bridge between robustness of institutional design and robust statistics. The ideal condition is violated if EDA is possible—something of real importance has been omitted from the model. The failure is akin to Popperian falsification: the model predicts that something is not of real important, but it is.

The robustness problem arises now. What is the consequence of this failure? A robust institution is one which puts a bound on the loss from such failure. A nonrobust institution does not have such a bound.

Against Romance and Exact Micronumerosity

One of Buchanan's persistent themes has been the importance of comparing actual institutions against actual institutions, e.g., against the possibility of market failure one must compare the possibility of government failure.²⁵

The link to robustness is obvious is we think about actually estimating an equation for a non-existing institution. We have the case that $N = 0$ —Arthur Goldberger's delightfully pungent "exact micronumerosity"—which trivially resists identification.²⁶ Presumably the equation would be identifiable when $N > K$, so we know that EDA is possible. Thus, considerations of robustness are paramount in comparing existing institutions with a nonexistent "ideal" institution. We know that any τ proposing to describe the ideal must fail for want of sufficient observations to make the identification.

Evolutionary Institutions

Suppose that an institution exists without a unique design theory, one which evolved by a process of trial and error.²⁷ It is suggestive that one institution which we have good reason to believe evolutionary—the practice of election by lot in Athens—can be shown to be statistically robust.²⁸ Indeed, since there is no design theory involved—only EDA on the part of the participants²⁹—the robustness with respect to design theory is a matter of definition.

Moreover, many of the theorists to whom evolutionary insight has been credited—Hume and Smith in particular—clearly thought in terms of a utilitarianism based on "rules" not on "actions."³⁰ Figure 1 may be one way to express the concerns of those philosophers who defend "rule utilitarianism." While there exists some state of affairs τ for which 1 is superior to 2, that is perhaps not how one wishes to bet. I believe that "rule utilitarianism" expresses robustness concerns even outside the confines of evolutionary theorizing, a case I shall make by considering the argument in William Paley in an historical appendix.

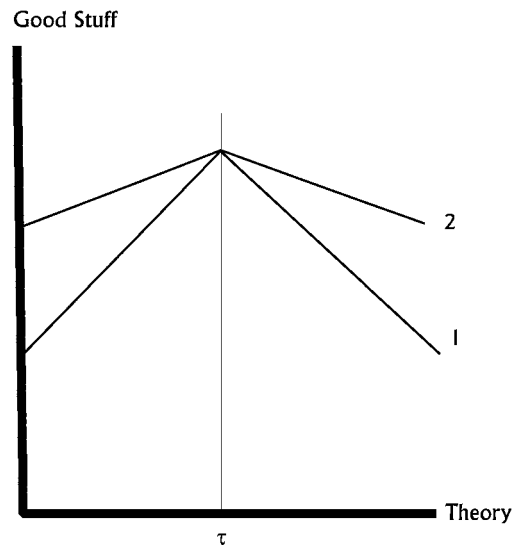


Figure 2. 2 is more efficient than 1.

Unfortunately, much modern thinking of evolutionary institutions has been expressed in terms of “efficiency,” a formulation which encourages refutation. To establish an institution efficient *without specifying* which model one is describing, one must make the case that it is superior to the alternative for *all models* of T . I suspect that this case cannot be made for any interesting institution.³¹ Or, if it were made, then the dominated institution ought to lose all its disinterested proponents! This case is drawn as Figure 2.

In a statistical context such an argument would be the demonstration that an estimator such as 1 is inadmissible; there is *no* distribution for which it is superior.³² In econometrics such proofs do indeed exist. The Gauss-Markov theorem establishes that, with respect to the class of linear unbiased estimators, only least squares is admissible.³³

As I read the record, von Mises and Hayek did try to establish that central planning was an inadmissible institution, there is no *model*—this is just a consistency condition, not a description of the world as we know it—for which planning was not dominated. Even Buchanan was not persuaded of this!³⁴

Conclusion

Inside economics the path of acceptance of both robust statistics and worst-case public choice thinking has been slow and controversial. The fact that the assumption of normality is still paradigmatic in econometrics and the fact (or so I am told) that the benevolent despot assumption is still paradigmatic in public finance are, in my way of thinking, only one fact. Best-case theorizing supposes the theorist has perfect insight into the structure of the world. EDA cannot be performed because we know the structure prior to study. This claim robust theorizing renounces. Without the assumption of perfect insight, the world is a messy

place. But Tukey said all that really needs to be said about this a long time ago when he proposed that “commandments” to empirical workers be expressed in the negative, i.e., as “badmandments”

The admonitions which follow are, indeed evil; they are not mere straw men or scape-goats. They are truly badmandments....

The great badmandment can be stated in all languages and to apply to any situation. In general allegorical language it reads:

IF IT’S MESSY, SWEEP IT UNDER THE RUG.³⁵

Appendix: William Paley, Robust Rule-Utilitarian

To come to fuller understanding of the link between robust thinking and rule-utilitarianism, we shall consider one of the set pieces in Paley’s 1785 *Moral Philosophy*.³⁶ Does a utilitarianism based on general rules differ in important ways from a utilitarianism based on acts? To address this problem Paley *counts* the number of people hurt and the number helped, he does not *add* hurts to helps to form a sum. First, he defines himself as a utilitarian:

So then actions are to be estimated by their tendency to promote happiness.—Whatever is expedient, is right.—It is the utility of any moral rule alone which constitutes the obligation of it.

Consider a policy—murdering a miser—which increases the *number of happy people*:

But to this there seems a plain objection, viz. that many actions are useful, which no one in his wits will allow to be right. There are occasions, in which the hand of the assassin would be very useful. The present possessor of some great estate employs his influence and fortune, to annoy, corrupt, or oppress all about him. His estate would devolve by his death, to a successor of an opposite character. It is useful, therefore, to dispatch such a one, as soon as possible, out of the way, as the neighborhood would exchange thereby a pernicious tyrant for a wise and generous benefactor. It may be useful to rob a miser, and give the money to the poor; as the money no doubt would produce more happiness, by being laid out in food and cloathing for a half dozen distressed families, than by continuing locked up in the miser’s chest. It may be useful to get possession of a place, a piece of preferment, or of a seat in parliament, by bribery or false swearing; as by means of them we may serve the public more effectually than in our private station... (1785:61–62).

In all cases in the paragraph quoted there are more helped than harmed. That makes the policy “expedient.” To make it “right” we must generalize the policy. (How this works will be addressed after we finish reading Paley.) Were the action generalized, the number of unhappy people would explode:

You cannot permit one action and forbid another, without shewing a difference betwixt them. Therefore the same sort of actions must be generally permitted or generally forbidden. Where, therefore, the general permission of them would be pernicious, it becomes necessary to lay down and support the rule which generally forbids them.

Thus, to return once more to the case of the assassin. The assassin knocked the rich villain on the head, because he thought him better out of the way than in it. If you allow this excuse in the present instance, you must allow it to all, who act in the same manner, and from the same motive; that is, you must allow every man to kill any one he meets, whom he thinks noxious or useless; which, in the event, would be to commit every man's life and safety to the spleen, fury, and fanaticism of his neighbour—a disposition of affairs which would presently fill the world with misery and confusion; and ere long put an end to human society, if not to the human species. (1785:64).

The “generalization” step which Paley asserts is the replacement of an action assumed motivated by benevolence—the murder of a miser to increase the number of happy people—with an action motivated by other reasons—murder. In terms of Figure 1 above, Murder | τ is not the same thing as Murder. And τ is the assumption of benevolent utilitarianism. The critical claim asserted by Paley is that Murder | τ will increase the occurrence of Murder because non-utilitarians will get in the act. Whether this is so or not, the possibility is serious and deserving of empirical attention for someone who defends Murder | τ .

Notes

1. Andrews et al. (1972:7C4): “Perhaps even more important than to say which estimators should be used, is to say which shouldn't. An important reason for not using an estimator is that it may behave catastrophically under some circumstances.” Brennan and Buchanan (1984:392): “The imputation of *homo economicus* motivation to actors in political roles may seem to violate ordinary notions about descriptive reality more than the comparable imputation to actors in the marketplace. But this difference need not provide any justification for replacing the model used for institutional comparison. It may be that judges seek to ‘uphold the law’ most of the time, that most government employees try to further their own conceptions of ‘public interest’ most of the time, and that elected politicians are genuinely concerned about promoting the ‘good society.’ But, even if this were admitted, institutional arrangements would surely be preferred which made these congruent with narrow self-interest on the part of the relevant actors. A model of human behavior in which the natural impulse toward self-interest, narrowly defined, predominates is a highly useful artifact in helping us to identify that set of arrangements that ‘economize on love.’”
2. Tukey (1960:474): “In slightly large samples, there is ground for doubt that the use of the variance (or the standard deviation) as a basis for estimating scaling type is ever truly safe. ... If contamination is a real possibility (and when is it not?), neither mean nor variance is likely to be a wisely chosen basis for making estimates from a large sample.” Mosteller and Tukey (1977:14): “... a few straggling values scattered far from the bulk of the measurements can, for example, alter a sample mean drastically, and a sample s^2 catastrophically.” Brennan and Buchanan (1984:392): “Suppose you are hiring a builder to build you a house. ... For empirical purposes, therefore, the assumption you will make about the said builder is that he is honest: you would not deal with him if you genuinely believed otherwise. But now you proceed to your lawyer's offer to draw up a contract. And in this setting, the working hypothesis you make about the builder is quite different. For the contract-drawing exercise, you make the assumption that the builder *is* going to fleece you, not because you believe this necessarily is his objective but because this is the contingency against which you wish to guard.”

3. From the 1960s both expressed grave doubts about exclusive focus on optimization accounts. As early as 1962 Tukey explicitly questioned reliance on optimization considerations, Tukey (1962). The issue was implicit as early as 1960, Tukey (1960:449–450): “Some years of close contact with the late C.P. Winsor had taught the writer to beware of extreme deviates, and in particular to beware of using them with high weights. Using second moments to assess variability means giving very high weights to extremely deviate observations. Thus the use of second moments, unquestionably optimum for normal distributions, comes into serious question.” Tukey (1977:1): **“We do not guarantee to introduce you to the ‘best’ tools, particularly since we are not sure that there can be unique bests.”** Summarizing a decade of work, Buchanan (1979:281): “... the maximizing paradigm is the fatal methodological flaw in modern economics.” Huber (1997:491): “‘Naive’ designs usually are more robust than ‘optimal’ designs.”
4. An account of common morality as false but useful theories of conduct is provided in Levy (1992).
5. One illuminating passage from Thomas Hobbes’ translation of the Melian dialogue will stand for it all. Thucydides (1975:381–382): “*Melians*: ‘We think, you well know, a hard matter for us to combat your power and fortune, unless we might do it on equal terms. Nevertheless we believe that, for fortune, we shall be nothing inferior; as having the gods on our side, because we stand innocent against men unjust; and for power, what is wanting in us will be supplied by our league with the Lacedæmonians, who are of necessity obliged, if for no other cause, yet for consanguinity’s sake and for their own honour, to defend us. So that we are confident, not altogether so much without reason as you think.’
 “*Athenians*. ‘As for the favour of the gods, we expect to have it as well as you: for we neither do, nor require anything contrary to what mankind hath decreed, either concerning the worship of the gods, or concerning themselves. For of the gods we think according to the common opinion; and of men, that for certain by necessity of nature they will every where reign over such as they be too strong. Neither did we make this law, nor are we the first that use it made: but as we found it, and shall leave it to posterity for ever, so also we use it: knowing that you likewise, and others that should have the same power which we have, would do the same.’”
 The editor notes sardonically that the Athenians sound as if they had been reading Hobbes’ *Leviathan* (p. 580).
6. Hume (1987:42): “Political writers have established it as a maxim, that, in contriving any system of government, and fixing the several checks and controuls of the constitution, every man ought to be supposed a *knave*, and to have no other end, in all his actions, than private interest.” The modern revival of this Humean point of view comes in Brennan and Buchanan (1980).
7. von Neumann and Morgenstern (1964:42): “Indeed we shall in most cases observe a multiplicity of solutions. Considering what we have said about interpreting solutions as stable ‘standards of behavior’ this has a simple and not unreasonable meaning, namely that given the same physical background different ‘established orders of society’ or ‘accepted standards of behavior’ can be built, all possessing those characteristics of inner stability which we have discussed.” Ulam (1991:199): “The one thing about Monte carlo is that it never gives an exact answer: rather its conclusions indicate that the answer is so and so, within such and such an error, with such and such probability... it provides an estimate of the value of the numbers sought in a given problem.”
8. Tukey (1986:289): “Then the work of Abraham Wald rolled back certainty still further when he showed that, insofar as procedures leading to definite actions were concerned, there could, in general, be no single optimal statistical procedure, but only a ‘complete class’ of procedures, among which selection must be guided by judgment or outside information.”
9. There is a tendency for modern economists in the Bayesian tradition to let the utility-theoretic appendix to von Neumann–Morgenstern stand for the whole book. Here is the authoritative comment of L.J. Savage on the second edition of *Theory of Games and Economic Behavior*, Savage (1972:281): “Contains, as a digression, the important new treatment of utility from which the treatment of utility in this book derives. The second edition contains more than the first on this subject, especially as a technical appendix. Also, the idea regarding multiple choices as single overall choices is discussed in great detail. Finally, the chapters on ‘zero-sum two-person’ games are mathematically intimately connected with the statistical minimax theory.”
 The close relationship between a minimax approach and an expected utility approach is doubtless responsible for Savage’s comments on the work leading to Huber (1981) at Savage (1972:291): “An important nonpersonalistic advance in the central problem of statistical robustness.” The Huber family of estimators is an explicit solution to a minimax problem, e.g., Andrews et al. (1972:2C1), Huber (1981). Thus, the estimator LJS, Savage’s contribution to Andrews et al. (1972:2C3), is a “Huber estimates are maximum likelihood estimates for a gross error model...”, i.e., a distribution which is normal in the center, double exponential in

the tails where the splicing parameter is estimated simultaneously. Its performance vis-à-vis the robustnik favorites—trimmed means, weighted quantiles and the redescending hampels—is a subject which would repay close study. The oddity of the early involvement of Bayesians in the robust movement—the term ‘robust’ was coined by a Bayesian!—vis-à-vis the current situation is noted in Huber (1997).

10. Ulam (1991:199): “The monte carlo method came into concrete form with its attendant rudiments of a theory after I proposed the possibilities of such a probabilistic schemes to Johnny in 1946 during one of our conversations. ... After this conversation we developed together the mathematics of the method. It seems to me that the name monte carlo contributed very much to the popularization of this procedure. It was named monte carlo because of the element of chance, the production of random numbers with which to play the suitable games.” Ulam (1991:200): “I felt that in a way one could invert a statement by Laplace. He asserts that theory of probability is nothing but calculus applied to common sense. Monte carlo is common sense applied to mathematical formulation of physical laws and processes.”
11. This must not be read as his *only* contribution. I remember the late Abraham Charnes saying that the Charnes-Cooper solution to the least absolute deviations regression problem depended upon devices from von Neumann.
12. Ulam (1991:102): “He was a great admirer of the concise and wonderful ways the Greek historians wrote. His knowledge of Greek enabled him to read Thucydides, Herodotus, and others in the original... The story of the Athenian expedition to the island of Melos, the atrocities and the killings that followed, and the lengthy debates between the opposing parties fascinated him for reasons which I never quite understood. He seemed to take a perverse pleasure in the brutality of a civilized people like the ancient Greeks. For him, I think it threw a certain not-too-complimentary light on human nature in general.” Ulam (1991:244): “I will never forget the scene a few days before he died. I was reading to him in Greek from his worn copy of Thucydides a story he liked especially about the Athenians’ attack on Melos, and also the speech of Pericles. He remembered enough to correct an occasional mistake or mispronunciation on my part.” ‘Hobbes’ translation is provided in note 5 above.
13. F.A. Hayek quotes Oliver Cromwell at Hayek (1960:530): “‘I beseech you, in the bowels of Christ, think it possible you may be mistaken.’” and then adds “It is significant that this should be the probably best-remembered saying of the only ‘dictator’ in British history!”
14. Tukey (1977:vi): “**The greatest value of a picture** is when it *forces* us to notice **what we never expected to see**.”
15. Chow (1983:22–24), Theil (1971:384–392).
16. Chow (1983:88–90), Theil (1971:615–622).
17. For the one-dimension case of a least squares regression (the sample mean) Tukey (1960), Andrews et al. (1972). For the multiple dimension case Mosteller and Tukey (1977) and Koeneker and Bassett (1978). Gib Bassett drew Figure 1 at a 1998 Public Choice Wednesday Seminar when he presented the results in Bassett and Persky (1999).
18. Remarkably enough students often emerge from an econometrics course under the impression that the efficiency properties of least squares are something which can be proven from the Gauss-Markov theorem! Perhaps the textbooks do not make clear that the “linearity” restriction in the Gauss-Markov theorem is can be expressed in terms of influence curves. From a linear influence curve it is immediate that the influence curve is unbounded; hence, the fragility of the estimate. Influence curves play a critical role in Andrews et al. (1972) and Mosteller and Tukey (1977). There is a certain delightful irony in *this* confusion because, as I shall argue below, it reappears when those who “Austrians” who work in the evolutionary tradition in opposition to the conventional neoclassical econometric approach try to make an “efficiency” case ought to be trying to make a “robustness” case instead.
19. Documentation for this reading of the debate is found in Levy (1990).
20. The embarrassing numbers are pictured in Levy (1993) as well as a characteristic claim made about them on highest economic authority.
21. As far as can see only J. Schumpeter made the point, but he made it in such a way that it would not be easily understood, Levy (1990).
22. The professional reaction to the work of Warren Nutter ought to be studied as an hideous example of the reaction awaiting someone whose results do not fit with the reigning preferences of how the world works.
23. In formal logic a model is defined as a collection of sentences from which all other sentences in the language cannot be deduced, Chang and Keisler (1978:9). This consistency condition in econometrics fails when

- singular matrices occur, the multiple-dimension equivalent of division by zero. For my tastes Theil (1971) remains unsurpassed on the various aspects of the identification question, e.g., the link between colinearity and identification (1971:152) and the problem of non-existing moments in popular simultaneous equation procedures (1971:532). A “consistent estimator” is a well-defined term with a meaning unrelated to model-theory “consistency.” “Identifiable” is as close to model-theory consistency as I think we can get.
24. Try writing this sentence with a straight face. Lord Shaftesbury’s doctrine—the truth cannot be ridiculed—often proves its worth.
 25. Buchanan (1984:11): “I have often said that public choice offers a ‘theory of government failure’ that is fully comparable to the ‘theory of market failure’ that emerged from the theoretical welfare economics the 1930s and 1940s. In that earlier effort, the system of private markets was shown to ‘fail’ in certain respects when tested against the idealized criterion for efficiency in resource allocation and distribution. In the later effort, in public choice, government or political organization is shown to ‘fail’ in certain respects when tested for the satisfaction of idealized criteria for efficiency and equity.”
 26. Goldberger (1991:248–249): “Econometrics texts devote many pages to the problem of multicollinearity in multiple regression, but they say little about the closely analogous problem of a small sample size in estimating a univariate mean. Perhaps that imbalance is attributable to the lack of an exotic polysyllabic name for ‘small sample size.’”
 27. Hayek is uniquely responsible emphasizing the importance of evolutionary institutions in the thinking of David Hume, Adam Smith. Hayek (1960) is perhaps his most systematic attempt. I am persuaded that most, if not all, of Hayek’s insights can be translated into robust terms even when they do not fit so nicely in standard efficiency terms. Part of my confidence is that there is evidence, which is independent of Hayek’s account, of robust considerations. Hayek points to Adam Smith as a formative theorist in the evolutionary tradition. Adam Smith’s utilitarianism—which Hayek does not consider—seems to have been made in terms of the well-being of majority, i.e., robust utilitarianism, Levy (1995). The same attention to median well-being is found in Malthus, cf. Hollander (1997). Malthus is explicit in pointing out this in Smith’s work. Tyler Cowen suggests that the discussion of ‘worst off’ in Rawls (1971) might be a generalization of the Smithian approach as it allows *any* quantile to be selected.
 28. Levy (1989) discusses how Athenian practice used both voting and random representation. The Athenian understanding—from what we know about this at our distance—suggested that random representation ought to be employed in a bimodal context in which majority rule voting is nonrobust. Bassett and Pesky (1999) establish the formal equivalence of voting and estimation.
 29. Levy (1992) gives an explanation of evolutionary institutions in classical antiquity, e.g., the Athenian lot, the Homeric epics and the cult of the hero.
 30. Levy (1999) gives an account of Smith’s katalactic approach vis-à-vis the neo-classical “Robinson Crusoe” approach in terms of robustness.
 31. Levy (1999) argues that such evolutionary strategies as “tit-for-tat” are robust even if the “proof” of efficiency is simple hand-waving by supposing a zero rate of time preference.
 32. Wald (1950) formalized the concept of admissibility using von Neumann-Morgenstern minimax devices.
 33. The reader who thinks I neglect the condition that the Gauss-Markov theorem requires that the error distribution have the first two moments might reflect upon Tukey’s point, Tukey (1986:411), that there is no way to distinguish between a Cauchy with the tails trimmed at $\pm 10^{-40}$ and the theoretical Cauchy. In the former case all moments exist, in the latter none exist.
 34. This remarkable bit of doctrinal history is provided in Levy (1990).
 35. Tukey (1986:199). This is an unpublished article from 1960 resulting from Tukey’s year associating with behavioral scientists.
 36. This section exists to answer Geoffrey Brennan’s query.

References

- Andrews, D. F., Bickel, P. J., Hampel, F. R., Huber, P. J., Rogers, W. H., and Tukey, J. W. (1972) *Robust Estimates of Location*. Princeton: Princeton University Press.
- Anscombe, F. J. (1960) “Rejection of Outliers.” *Technometrics*, 2: 123–147.

- Bassett, G. Jr. and Persky, J. (1999) "Robust Voting." *Public Choice*, 99: 299–310.
- Box, G. E. P. (1953) "Non-Normality and Tests on Variances." *Biometrika*, 40: 318–335.
- Brennan, G. and Buchanan, J. M. (1980) *The Power to Tax: Analytical Foundations of a Fiscal Constitution*. Cambridge: Cambridge University Press.
- Brennan, G. and Buchanan, J. M. (1984) "The Normative Purpose of Economic 'Science.'" In: Buchanan, J. M. and Tollison, R. D. (Eds.) *The Theory of Public Choice—II*. Ann Arbor: University Press.
- Buchanan, J. M. (1979) In: Brennan, H. G. and Tollison, R. D. (Eds.) *What Should Economists Do?* Indianapolis: Liberty Press.
- Buchanan, J. M. (1984) "Politics without Romance." In: Buchanan, J. M. and Tollison, R. D. (Eds.) *The Theory of Public Choice—II*. Ann Arbor: University Press.
- Chang, C. C. and Keisler, H. J. (1978) *Model Theory*. 2nd edn. Amsterdam: North-Holland.
- Chow, G. C. (1983) *Econometrics*. New York: McGraw-Hill.
- Goldberger, A. S. (1991) *A Course in Econometrics*. Cambridge, Mass.: Harvard University Press.
- Hayek, F. A. (1960) *The Constitution of Liberty*. Chicago: University of Chicago Press.
- Hollander, S. (1997) *The Economics of Thomas Robert Malthus*. Toronto: University of Toronto Press.
- Huber, P. J. (1997) "Robustness: Where are We Now?" In: Dodge, Y. (Ed.) *L₁-Statistical Procedures and Related Topics*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, Vol. 31.
- Huber, P. J. (1981) *Robust Statistics*. New York: John Wiley.
- Hume, D. (1987) In: Miller, E. F. (Ed.) *Essays Moral Political, and Literary*. Revised edition. Indianapolis: Liberty Classics.
- Koeneker, R. and Bassett, G. Jr. (1978) "Regression Quantiles." *Econometrica*, 46: 33–50.
- Levy, D. M. (1989) "The Statistical Basis of Athenian-American Constitutional Theory." *Journal of Legal Studies*, 18: 79–103.
- Levy, D. M. (1990) "The Bias in Centrally Planned Prices." *Public Choice*, 67: 213–226.
- Levy, D. M. (1992) *Economic Ideas of Ordinary People: From Preferences to Trade*. London, 1992.
- Levy, D. M. (1993) "The Public Choice of Data Provision." *Accountability in Research*, 3: 157–163.
- Levy, D. M. (1995) "The Partial Spectator in the Wealth of Nations: A Robust Utilitarianism." *European Journal of the History of Economic Thought*, 2: 299–326.
- Levy, D. M. (1999) "Katallactic Rationality: Language, Approbation & Exchange." *American Journal of Economics and Sociology*, 58: 729–747.
- Levy, D. M. (1999/2000) "Problem and Solution: Non-Normality and Exploratory Data Analysis." *Econometric Theory*, 16: 99.3.3 [Solution]; 15: 99.3.3 [Problem].
- Mosteller, F. and Tukey, J. W. (1977) *Data Analysis and Regression*. Reading, Mass.: Addison-Wesley.
- Paley, W. (1785) *The Principles of Moral and Political Philosophy*. London.
- Rawls, J. (1971) *A Theory of Justice*. Cambridge, Mass.: Harvard University Press.
- Savage, L. J. (1972) *The Foundations of Statistics*. New revised edition. New York: Dover.
- Theil, H. (1971) *Principles of Econometrics*. New York: John Wiley.
- Thucydides. (1975) In: Schalatter, R. (Ed.) *The History of the Peloponnesian War*. Translated by Thomas Hobbes. [=Hobbes's Thucydides.] New Brunswick: Rutgers University Press.
- Tukey, J. W. (1960) "A Survey of Sampling from Contaminated Distributions." In: Olkin, I., Ghurye, S. G., Hoefding, W., Madow, W. G., and Mann, H. B. (Eds.) *Contributions to Probabilities and Statistics*. Stanford: Stanford University Press.
- Tukey, J. W. (1962) "The Future of Data Analysis." *Annals of Mathematical Statistics*, 33: 1–67.
- Tukey, J. W. (1977) *Exploratory Data Analysis*. Reading, Mass.: Addison-Wesley.
- Tukey, J. W. (1986) In: Jones, L. V. (Ed.) *Philosophy and Principles of Data Analysis*. Vol. 3 of *The Collected Works of John W. Tukey*. Monterey: Wadsworth & Brooks.
- Ulam, S. M. (1991) *Adventures of a Mathematician*. Berkeley: University of California Press.
- Von Neumann, J. and Morgenstern, O. (1964) *Theory of Games and Economic Behavior*. 3rd edn. New York: John Wiley.
- Wald, A. (1950) *Statistical Decision Functions*. New York: John Wiley.