



Has machine learning rendered simple rules obsolete?

Jesús Fernández-Villaverde¹

Accepted: 6 February 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC part of Springer Nature 2021

Abstract

Epstein (Simple rules for a complex world, Harvard University Press, Cambridge, 1995) defended the superiority of simple legal rules over complex, human-designed regulations. Has Epstein's case for simple rules become obsolete with the arrival of artificial intelligence, and in particular machine learning (ML)? Can ML deliver better algorithmic rules than traditional simple legal rules? This paper argues that the answer to these question is “no” by building an argument based on three increasingly more serious barriers that ML faces to develop legal (or quasi-legal) algorithmic rules: data availability, the Lucas' critique, and incentive compatibility in eliciting information. Thus, the case for simple legal rules is still sound even in a world with ML.

Keywords Artificial intelligence · Machine learning · Economics · Simple rules

JEL codes D85 · H10 · H30

1 Introduction

Epstein (1995) made a strong case for simple legal rules. He argued that, since most people know their preferences better than policymakers do, the legal system should give them the freedom to protect and advance their interests based on their judgment. In Epstein's view, the best way to accomplish this task was through

First, I need to thank my co-authors on several projects where I have used machine learning: Pat Bajari, Sara Casella, Stephen Hansen, Samuel Hurtado, Galo Nuño, and Charlie Manzanares. Second, I thank Fernando Arteaga, Tyler Cowen, Peter Rupert, and Don Sillers for their comments. Third, I must also thank several generations of graduate students in economics at the University of Pennsylvania, Princeton University, Harvard University, Stanford University, and the University of Oxford, where I have taught courses that covered most of the material discussed here. Their questions helped to shape my thinking about this topic.

✉ Jesús Fernández-Villaverde
jesusfv@econ.upenn.edu

¹ University of Pennsylvania, Philadelphia, USA

simple legal rules that embrace autonomy, private property, and freedom of contract. Instead, complex, human-designed regulations, such as the legislation enacted during the last decades in areas such as the health and financial sectors, “place the power of decision in the hands of other people who lack the necessary information and whose own self-interest leads them to use the information that they do have in socially destructive ways” (Epstein 1995, p. XII).

The historical evidence has supported Epstein’s case for simple rules. For instance, time after time, complex financial regulations designed to “fix” the latest crisis have only been the harbinger of new problems. The collapse of the U.S. Treasuries market in March 2020, when intermediaries were overwhelmed by sales orders, was a direct consequence of the regulatory reforms passed in the aftermath of the financial crisis of 2008-2009. Higher capital requirements and increased dealer credit spreads limited the capacity of dealers to safely intermediate the market on their balance sheets after the COVID-19 shock. Only the massive intervention by the Federal Reserve System, which purchased \$1 trillion of Treasuries in late March-early April, avoided a widespread meltdown (see Duffie 2020, for details and a proposal to solve this issue through the simple rule of a broad central clearing mandate). It is easy to find similar examples of the shortcomings of complex regulations in many sectors: new problems appear as quickly as the older ones are “solved.”

Some critics of simple rules have recently contended that the problems of complex regulations highlighted by Epstein (1995) and many others apply only to human-designed rules (perhaps because the later suffer from cognitive and ideological biases). The argument is that we can take advantage of the tremendous recent advances in artificial intelligence (AI), and in particular machine learning (ML), to design complex algorithmic rules that deliver social outcomes superior to those from simple rules *and* complex human-designed rules.

For example, Jack Ma, founder of Alibaba, stated in November 2016:¹

Over the past 100 years, we have come to believe that the market economy is the best system, but in my opinion, there will be a significant change in the next three decades, and the planned economy will become increasingly big. Why? Because with access to all kinds of data, we may be able to find the invisible hand of the market.

The planned economy I am talking about is not the same as the one used by the Soviet Union or at the beginning of the founding of the People’s Republic of China. The biggest difference between the market economy and planned economy is that the former has the invisible hand of market forces. In the era of big data, the abilities of human beings in obtaining and processing data are greater than you can imagine.

¹ Other authors making similar arguments include Saros (2014), Phillips and Rozworski (2019), and Morozov (2019). The latter author has even coined a term for this alternative approach: “digital socialism.”

With the help of artificial intelligence or multiple intelligence, our perception of the world will be elevated to a new level. As such, big data will make the market smarter and make it possible to plan and predict market forces so as to allow us to finally achieve a planned economy.²

Is Jack Ma right? Has ML rendered Epstein's defense of simple rules obsolete? In other words: If a reader was convinced in 1995 by Epstein's argument, should she change her mind in 2021 because of the arrival of ML? Can, for example, a deep neural network offer better guidance to monetary policy than a simple Taylor rule? Can a recurrent neural network allocate goods better than the interaction of buyers and sellers in a free market?³

This paper argues that the answer to these questions is "no." While ML is a useful tool to solve many problems of interest, Epstein's advocacy of simple rules is still solid and ML-based mechanisms cannot replace them. More concretely, the paper builds a case based on three increasingly more serious barriers that ML faces to develop legal (or quasi-legal) algorithmic rules.

The first barrier, a rather obvious one, is that ML requires enormous datasets that are unlikely to exist in most cases of interest. To put it concisely, we probably have enough data to design an early warning system for future financial crises that can improve with respect to existing methods (Fouliard et al. 2019). However, we are not (and will never be) even remotely close to having the amount of data required to train a complex neural network that will work better than a simple Taylor rule.

The second barrier is that, except in a few cases, ML suffers from the Lucas critique (Lucas 1976). Any estimation with non-experimental data is subject to an external validity constraint: economic agents make decisions based on expectations about policy regimes. Thus, any variation in policy renders previous observations useless unless we have a structural model (i.e., an economic model that is explicit about preferences, technology, and information sets) the researcher can use to recompute the optimal responses to the new policy. We can even go one step ahead and defend the argument that such a structural model should also incorporate policymakers' probability of changing the environment. By construction, ML has little to say about structural models as they are reduced-form statistical representations. Since we need to know a structural model to design successful legal rules, ML is of limited use for this purpose.⁴

The third barrier, and the most unconquerable, comes from our understanding of how societies create and use knowledge. ML cannot circumvent the problems of

² See <http://www.globaltimes.cn/content/1051715.shtml>.

³ These questions are most relevant if one finds Epstein's defense of simple rules compelling. But even a reader who was not convinced by Epstein's arguments in 1995 might be interested in this paper. Since the next pages will argue that complex algorithmic rules are inferior to simple rules, a defender of the superiority of complex human-designed rules over simple rules can use transitivity to conclude that my argument also proves the superiority of complex human-designed rules over complex algorithmic rules.

⁴ There are two partial exceptions to this argument. First, ML is useful for policy evaluation when we have access to experimental (or quasi-experimental) data. While those situations are relevant, the range of policy questions that it can address is limited. Second, ML can also help in the estimation of structural models, but that is far from what the defenders of "digital socialism" have in mind.

truthful information elicitation. Hayek (1945) taught us, well before anyone knew much about ML, that the core of the allocation problem in a society is not how to solve an optimization problem, which ML can do better in large dimensions than traditional mathematical methods. The fundamental barrier that social organization faces is that information is dispersed and agents do not have incentives (and often not even the capabilities) to disclose such information to a centralized mechanism such as an ML algorithm. The Soviet Union's Achilles' heel was not that its computers were not powerful enough (although they were not) or that its planning algorithms were poor (they were terrible). Its problem was that it could never develop an incentive-compatible system of information elicitation.

Interestingly, the renewed interest in ML as a substitute for simple rules returns us to classic discussions in economics from the 20th century, such as the role of experts in policy formulation (although, in this case, the "expert" is an impersonal software) or the socialist calculation debate (see Levy and Peart 2016, for an introduction). In fact, there were unsuccessful attempts at setting computer-directed systems of central planning in the Soviet Union (the National Automated System for Computation and Information Processing and the System of Optimal Functioning of the Economy; see Peters 2016) and Allende's Chile (Project Cybersyn; see Medina 2011).

The previous arguments do not imply that ML cannot be applied in economics (see Fernández-Villaverde et al. 2019 as an example). But it is crucial to frame the promises of ML realistically and avoid the disappointment that could come from overpromising.

The rest of the paper is organized as follows. Section 2 provides basic technical background on AI. Section 3 explains the data requirements of ML. Section 4 reviews how the famed Lucas critique applies to ML. Section 5 discusses ML and incentive compatibility in information elicitation. Section 6 concludes with some remarks about the evolutionary origin of simple rules and how it compares with ML.

2 Some background

AI is a vast field, ranging from expert systems to ML and from robotics control to natural language processing. Although modern AI started in a famed two-month summer workshop at Dartmouth College in 1956, it has been the blossoming of ML techniques during the last two decades that has brought AI to everyone's attention.⁵

ML includes a wide variety of numerical and statistical algorithms in which a computer is programmed to learn about some properties of data or an unknown function in a relatively automatic fashion.⁶

⁵ A standard university textbook on AI is Russell and Norvig (2010), which briefly covers the history of AI, including the Dartmouth workshop [p. 17], the "AI winters" of the 1970s and 1980s [pp. 22–24], and the recent advances in the field [pp. 24–28]. Those in a hurry or less inclined to peruse technical details can learn much from Boden (2018).

⁶ While "machine learning" is an expression that can capture everyone's imagination, it is not particularly precise. Other languages, such as French ("apprentissage automatique") or Spanish ("aprendizaje automático"), use the much less catchy but certainly more accurate expression "automatic learning."

An example will help us understand the main intuition. Imagine that we can specify a very flexible function that maps dozens of observables from credit card transactions (e.g., the geolocation of the store where the credit card is employed, the time of day and week when the transaction occurred, item purchased and its price, the use of the credit card in the previous 24 hours, etc.) into a prediction of whether the use of the credit card was fraudulent. The function is sufficiently flexible (for instance, a deep neural network) such that the researcher does not need to make many choices about which observables to add (is the geolocation of the store important?) or the details of the functional form (does the item price enter in a linear or log fashion?). If the researcher has access to millions of credit card transactions and knowledge of whether the transactions were fraudulent, the functional form can be “trained” to fit the data and detect, often with fantastic success, whether a new transaction is fraudulent. At a fundamental level, there is nothing very “intelligent” here: it is an exercise in massive data fitting. What is intelligent is that the process is highly automatic and, therefore, easily scalable to multiple environments where the researcher might have limited knowledge.⁷

Although some of the key ideas in ML go back to the 1940s (see McCulloch and Pitts 1943), it was not until around 2000 that research on ML and its applications in industry boomed. Three reasons account for such a long lag.

The first reason is that, while some key algorithms such as artificial neural networks were well understood for decades, the computational capability to implement them outside “proof-of-concept” environments was not widely available at cheap prices until a couple dozen iterations of Moore’s law.⁸ Simultaneously, massive parallelization became widely available in the early 2000s. Now, nearly every laptop sold in the U.S. market comes with multiple central processing units and a graphics processing unit. Moreover, cloud services mean that any researcher can have access over the internet, for a few dollars an hour, to computers that were previously only open to researchers at large universities and national labs. Fortunately, most ML algorithms are particularly suitable for massive parallelization.

The second reason is that thanks to the internet and cheap computing, researchers and industry practitioners have been able to gather incredibly rich datasets. For example, in economics, a typical empirical paper had hundreds of observations in the 1970s (Hall 1978, uses 120 observations), thousand of observations in the late 1990s (Acemoglu et al. 2001, use 1663 observations), and today it is common to see

Footnote 6 (continued)

Many English-speaking researchers used to talk about “statistical learning,” but that term has lost popularity.

⁷ In that sense, ML is very far away from the “artificial general intelligence” pursued, rather fruitlessly, during the 1960s and that we see reflected in Hollywood’s dystopias.

⁸ Moore’s law, proposed by Gordon Moore, the co-founder of Intel, in 1965 states that the number of components (i.e., transistors) per integrated circuit (i.e., per “chip”) will double each year (Moore 1965). A simple way to think about Moore’s law is that the ability of humans to perform numerical computations advanced as much between July 2019 and December 2020, when these lines are being written, as between the dawn of our species around 300,000 years ago and July 2019. See Flamm (2019) for evidence on the astonishing empirical success of Moore’s prediction.

papers with tens of millions of observations (Chetty et al. 2014, use 47.8 million).⁹ That is why ML is often associated with expressions such as “big data” or “data analytics.” As Sect. 3 will discuss, ML algorithms are data-hungry, and there is little one can do with them if the only data at hand are a few hundred observations.

The third and final reason is that computer scientists and applied mathematicians solved some of the roadblocks to the efficient implementation of ML algorithms. Ideas such as backpropagation (Rumelhart et al. 1986) and the Latent Dirichlet allocation (Blei et al. 2003) cracked open doors that had been closed for a long time.

The combination of these three reasons has meant that ML has become ubiquitous in our lives, that more and more students are focusing on acquiring these skills, and that public policy institutions and researchers are starting to extract conclusions about how ML will affect policy. But can ML be a substitute for simple legal rules?

3 Data availability

ML requires enormous datasets to work properly. For example, how a deep neural network takes observables and outputs a prediction is indexed by dozens of parameters that must be determined from the data (i.e., the network needs to be “trained”). Although each case is slightly different, the rule-of-thumb in the industry is that one needs around 10^7 labeled observations to train a complex neural network properly, with around 10^4 observations in each relevant group.¹⁰

When do we have these large datasets? In two situations. First, when you are a firm such as Amazon or Netflix, with access to millions of observations from your customers. Every time you purchase a product from Amazon, you provide the company with one more observation of what you like, how the purchase correlates with other products you bought, how sensitive you are to price changes, etc. With around 105 million Amazon Prime customers in the U.S., a group of customers with which Amazon can expect to have repeated interactions during a calendar year, Amazon has access to all the data it can handle.¹¹ Similarly, every time you pick a movie or a show to watch on Netflix, you provide the streaming service additional information about which shows people like you enjoy.

Second, you can create your “own” data. Creating your “own” data might seem counter-intuitive (or plainly dishonest), but it is easy and consistent in many environments. A simple example of this approach is the training of AlphaZero, a computer program that learned to play Go, chess, and shogi through self-play (Silver et al. 2018). A researcher can randomly generate many initial sets of values for the parameters in the neural network that maps positions of the pieces on the board with

⁹ These three papers are considered landmark empirical contributions of their cohorts. As of December 31, 2020, Hall (1978) has 4,850 Google citations, Acemoglu et al. (2001) have 14,489, and Chetty et al. (2014) have 2,197 (an impressive number for a 7-year-old paper).

¹⁰ See Goodfellow et al. 2016, Ch. 15, for a discussion of data requirements. My experience is that one needs fewer observations, perhaps around 10^5 , if one is careful with the parameterization of the problem or is willing to impose some additional structure.

¹¹ See <https://www.statista.com/statistics/546894/number-of-amazon-prime-paying-members/>

their value and the next movement. Each set of values defines a different game strategy (i.e., should you place a bishop in this corner or move the left knight?). Then, we can pit these different strategies against each other by making them play multiple rounds of the game, select the best strategies in terms of victory percentages, and update the value of the network parameters to reflect those in the winning strategies (plus some random changes to explore new strategies). See Athey et al. (2019b) and Fernández-Villaverde et al. (2019) for recent examples of how similar ideas can be applied in economics.

Unfortunately, in most situations where we want to develop legal algorithmic rules, we do not have access to such a wealth of data and, most likely, we never will. Take the example of monetary policy, a relatively well-structured problem with much fewer instruments and targets than other policy problems. Could a neural network ever replace the Federal Market Open Committee (FOMC)? Probably no.¹²

When the members of the FOMC set the federal funds rate (the main policy instrument of the Federal Reserve System), they have access to time series of most variables of interest that are short in length. For instance, in the case of the U.S., we only have reliable data for output, consumption, and investment after World War II and, even that, only at quarterly frequency.¹³ If we count them from 1947:Q1 (the first “good” observation in terms of accuracy of our measurement) until 2020:Q3 (the last observation as when these lines are being written), we have 295 data points.

Furthermore, the U.S. economy has undergone radical structural changes. To name a few, we have moved from manufacturing into services and improved supply-chain management (Davis and Kahn 2008). Financial innovations have transformed the relationship between financial and real variables (Guerrón 2009). Monetary policy has been conducted more aptly after 1982 (Lubik and Schorfheide 2004). That is why, often, econometricians do not use observations before the early 1980s when they estimate the effects of monetary policy on output. Fernández-Villaverde et al. (2015) is an example of how these estimates change sharply depending on whether we include early observations.

Structural change matters for ML because, as time passes, we gain observations at the end of the sample, but lose informational value from the observations at the start of the sample. The net effect of more observations will be positive, but reduced. No, by 2050, we will not have radically more information about the aggregate behavior of the U.S. economy than we do today.

Going to microdata (e.g., consumption data of individual households or financial transactions) can help to enrich the observations, but we will still encounter severe limitations on the length and stability of micro surveys. Think, just as an illustrative

¹² This is not a criticism of the use of ML with microdata to provide the FOMC with better information. Instead, it is a statement questioning the possibility that you can substitute an ML algorithm for the FOMC deliberations or even a simple Taylor rule.

¹³ There have been quite valuable attempts at rebuilding series of output for the U.S. before World War II, the most famous of which is probably Kendrick (1961). However, these reconstructions incorporate enough noise that, beyond their usefulness for historical study, they are probably not robust enough to be fed into the training of a neural network or the estimation of an econometric model used for policy decisions.

case, about demographic change. How informative are the consumption patterns of married couples in the 1990s, in their early 40s with several kids at home, about the consumption patterns of single individuals in the 2020s, also in their early 40s and without kids? Comparing single individuals of the 1990s with single individuals of the 2020s will not work either because who is single in the 2020s is very different from who was single in the 1990s. The selection bias into marriage has changed dramatically. Modern microeconometrics starts from the realization of how difficult it is to control for such selection bias and to explore ways to address it. Besides, there are severe limitations on what microdata can teach us, in the absence of a structural model, about the general equilibrium effects we must know for conducting monetary policy. See Deaton and Cartwright (2018), for a discussion of some of these topics.

4 The Lucas critique

There are, however, concerns about the usefulness of ML for the development of legal rules that go beyond the limitations of data. Probably the most salient is the Lucas critique (Lucas 1976).

Imagine that we obtain data on airline ticket prices and the occupancy rate of particular flights. Any statistical or ML model will soon pick up that the ticket prices and the occupancy rates move together: we see high prices in the oversold American Airlines flights leaving Philadelphia for Boston on Monday mornings at 7:00 am and low prices in the relatively empty 3:00 pm Tuesday flights from Boston to Philadelphia.

How can American Airlines use this information to decide the profit-maximizing ticket prices (or a regulator determine the socially optimal prices)? The conclusion that high-price plane tickets cause high occupancy is senseless. Instead, we can safely conclude that American Airlines is pricing the 7:00 am flight higher because it understands the demand for that flight is strong. However, in reaching such a conclusion, we have relied on basic economic theory (which is nothing more than organized common sense) and shown an essential limitation of ML: it is tough to use it to assert causality (although we are making promising progress; see Athey et al. 2019a). For most policy questions of interest, we care about causality, not correlation. Only by understanding causality can we design better policies.

While in the airline example the direction of causality was evident, it is not in many others. Do children in charter schools do better because charter schools implement superior teaching practices or because their parents are more motivated than parents of children who stay in traditional public schools? Does having more police lower crime, or do higher crime areas attract more police officers? Does the rule of law drive economic growth, or does economic growth create a demand for the rule of law? Unfortunately, social sciences live in a world of high potential causal density (Manzi 2012). For nearly all situations, we can have a multitude of plausible causal channels between policies and outcomes.

Lucas, however, went even further than just rehearsing the old causality vs. correlation line. Let us return to the airline example. Now, consider the slightly different problem that American Airlines faces when deciding the spread between the fare

price for economy class and business class. If the price of the business ticket is too high, I will not buy it. Instead, I will bet that my frequent-flier status with the airline will allow me to upgrade. However, if the price of the business ticket is not too high, I would instead buy a business ticket to Boston as I want to ensure I am well-rested when I arrive in Boston, instead of risking not being upgraded and suffering the discomfort of a small seat. American Airlines prefers that I buy a business ticket (I pay a higher price) rather than upgrading me, but it needs to gauge my price elasticity of demand.

Can ML help here? Yes. We can find, if there is enough variation in the data, who buys business fares and who buys economy and, through information on their income, education levels, home address, etc., back up such a price elasticity of demand.

However, and this the key to Lucas' argument, the answer that comes from ML will only be valid under a constant set of circumstances. For example, if American Airlines tightens its rules for upgrades (as it did a few years ago), my price elasticity of demand will increase (note: it is the whole price elasticity of demand that changes, not just my quantity demanded). Why? The tightening of the upgrade rules hurt business travelers with many segments per year of low-price tickets (i.e., the Philadelphia-Boston commuters). This change, implicitly, favored business travelers with few segments per year but of high price (for instance, expensive business class tickets to Asia) and who now face less upgrade competition. Once I understand that upgrades are more likely, I will risk playing the upgrade lottery when the price of the business ticket goes above a lower threshold than before. One can argue that was American Airlines' goal to begin with (reward their most highly profitable transcontinental customers with easier domestic upgrades), but the point here is that no amount of ML is going to tell you by how much my price elasticity of demand will increase under the new upgrade policy. For that, you need an economic model, which will tell you about policy-invariant parameters such as risk aversion.

Some ML practitioners will reply that American Airlines can always get around the Lucas critique by experimenting with different upgrade policies. Yes and no. Yes, companies can and do experiment all the time (Manzi 2012). Nevertheless, there are limits to such experimentation. While Amazon can experiment, at meager cost, with the recommendations it displays on its homepage when you open it, American Airlines can only change the upgrade rules sporadically unless it wants to alienate its customers.

Most importantly, many policies are hard to test by experimentation. The first barrier is ethical: we do not believe, as a society, that we can play with humans to make a scientific point. Recall the 1983 comedy *Trading Places* and the Duke brothers' experiment with Dan Aykroyd and Eddie Murphy on the importance of nature vs. nurture. We celebrate the Duke brothers' ultimate ruin precisely because of our moral intuition that such experimentation is unacceptable. This ethical barrier is particularly salient in issues related to health and education.

The second barrier to experimentation is the limitation of what one can learn from such exercises. We can implement a randomized control trial (RCT) to evaluate the effect of charter schools, but we cannot change the federal funds rate to see

what happens with the U.S. economy afterward.¹⁴ Even evaluating charter schools is difficult. We can ascertain, with reasonable confidence, that sending a few thousand children with well-motivated parents to charter schools in the Boston metropolitan area has clear positive effects (Abdulkadiroğlu et al. 2011). However, it is hard to use an RCT to gauge general equilibrium effects.¹⁵ How will the program work with children of parents who did not apply to the lottery? What would happen with the location choices of parents in Boston once we generalize charter schools? And with the market for teachers? And with the composition of jobs offered by firms once the labor force is better educated?¹⁶

Interestingly, firms do not typically have to face these general equilibrium effects. If I learn, through experimentation, that placing the soda stand closer to the check-out counter increases the sale of sodas in my coffee shop, I have a minuscule effect on the national sales of sodas, their prices, and the diet of Americans. If the government mandates moving the soda stands away from the check-out counters across all shops in the country to lower the consumption of sugary drinks, we will change prices and dietary choices. Thus, the scope for experimentation that firms (or even local governments) enjoy, and the subsequent ability to employ ML is larger than the scope of national governments.

5 Incentive compatibility in eliciting information

The third and most fundamental criticism of the possibility of using ML to substitute for simple legal rules is that the ultimate reason we believe in simple legal rules is that they are better at eliciting the information required to achieve desirable societal goals.

This idea goes back to Hayek (1945), although it also appears prominently in Epstein (1995). Hayek's objection to centralized allocation mechanisms is not that solving the associated optimization problem is extremely complex –indeed, it is and increasingly so in an economy with a maddening explosion of products– or that we need to gather and process the data sufficiently fast. If that were the case, ML could perhaps solve the problem, if not now, then in a few more iterations of Moore's law. The objection to centralized allocation mechanisms is that the information one needs

¹⁴ In an RCT, we create a control and a treatment group randomly, and observe the differences in outcomes between the group that received treatment and the one that did not. A classic design experiment to gauge the effect of charter schools is to run a lottery among all the students who applied to an over-subscribed charter school. Since both the winners and the losers of the lottery applied to it, one should not expect any differences in motivation or background between the control and treatment groups (and we can test for balance on observables among the two groups).

¹⁵ Nevertheless, see Muralidharan et al. (2017) for an example of how to estimate these general equilibrium effects using a large-scale experiment.

¹⁶ The statements in the main text require a few additional caveats. For instance, one can use the results of an RCT or an ML exercise to estimate a general equilibrium model by imposing the condition that such a model replicates the RCT when we simulate it in partial equilibrium and use such a model for counterfactual policy analysis. But even in that case, we still need a structural model, and ML cannot substitute for it.

to undertake is dispersed and, in the absence of a market system, agents will never have the incentives to reveal it or even to create new information through entrepreneurial activity. As Steve Jobs put it: “A lot of times, people don’t know what they want until you show it to them.”¹⁷

A simple, real-life centralized allocation mechanism illustrates the point. Imagine a department of economics that, every year, faces the challenge of setting up a teaching matrix for the next academic year.¹⁸ Each faculty member submits her preferences in terms of courses to be taught, time slots, etc. Given the teaching needs and submitted requests, the computational burden of finding the optimal allocation is quite manageable. Given that an average department of economics at a top research university in the U.S. has around 30 faculty members and, once you consider that the average member of the theory group will never request to teach econometrics and vice versa, the permutations to consider are limited. A simple computer algorithm, such as those envisioned by the defenders of “digital socialism,” can do the job.

The real challenge is that, when I submit my teaching requests, I do not have an incentive to reveal the truth about my preferences or to think too hard about developing a new course that students might enjoy. I might not mind too much teaching a large undergraduate session on a brand-new hot topic and, if I am a good instructor, the students will be better off. However, I will not be compensated for the extra effort, even if it is not high, and I will have an incentive to request a small section for advanced undergrads on an old-fashioned topic. This request is not optimal: if the dean could, for instance, pay me an extra stipend, I would teach the large, innovative section, the students would be happier, and I would be wealthier.¹⁹

An obvious solution would be, then, not to submit a teaching request, but a schedule of teaching requests and a supply curve to do so, i.e., I will teach “the economics of big data” at 9:00 am on Mondays and Wednesdays at price x or “advanced monetary theory” at 1:00 pm on Tuesdays and Wednesdays at price $0.4x$. The software will use the supply curves to clear the teaching market and assign a faculty member to each course. This new scheme would increase the computational challenge of setting up the teaching matrix by one order of magnitude, but I can still write a short *Julia* program that will deliver an answer in minutes.

The drawback is that such a system of teaching requests and supply curves would open the door to all sorts of strategic behavior: I will consider, when I submit my supply curve, what I know about my colleagues’ tastes regarding teaching large, innovative courses. If I believe they genuinely dislike doing so, I will communicate

¹⁷ Quoted in *Business Week*, 25 May 1998.

¹⁸ The example of a small organization, a department of economics, illustrates that the problem of eliciting information is pervasive to all forms of collective organization that, because of transaction costs, cannot rely on a price system (Coase 1937).

¹⁹ Often, deans offer small teaching grants to reward innovation in teaching, but those are rarely worth even the time to fill in the application form. Consequently, we do not see much advance in teaching technologies in economics. This example, due to space limitations, does not discuss tacit knowledge and the difficulties in transmitting it, a point already present in Hayek (1945), but emphasized by Polanyi (1966, p. 4) when he explained how “we can know more than we can tell.”

a higher supply curve to teach such courses in order to clear the market at a higher price and increase my revenue. The outcome of the teaching matrix will not be efficient because I am not telling the truth, but playing strategically.

We can push the argument further. Knowing that the department will assign duties using a teaching request and a supply curve, I can manipulate from the day I am hired how I behave in front of my colleagues and the teaching requests and supply curves I submit. In such a way, I can introduce noise in their signal about my teaching preferences and exploit their incorrect inferences about my type when I submit my teaching requests and supply curve in the future. My colleagues would know that and act accordingly, changing their supply curve to reflect that they understand I tried to manipulate them. But I would also know that my colleagues know that and I will respond appropriately, and so on and so forth for one iteration after another. Those who do not believe the faculty would behave in such a way have not had experience managing academic departments.²⁰

There is an additional problem. Once I am assigned a course, how does the university ensure I teach it at the “optimal” quality level? Note that “optimal” cannot mean the highest possible quality. If I were to prepare every lecture that I give as a job market talk, the current students would love it, but I would not have time to undertake research, and my future students would get worse lectures, since my knowledge of the field would depreciate.

Even forgetting about that intertemporal aspect, how do we trade off one extra minute of research (which increases the university’s visibility and reputation) with one extra minute of teaching preparation? And how do we address heterogeneity in the comparative ability between research and teaching among faculty members when the efforts put into each activity are mostly unobservable?

Finally, we face the friction that I can carry my research with me to my next job (i.e., the publications in my C.V.) much more easily than my teaching evaluations (i.e., one can always “lose” the terrible teaching evaluation from 15 years ago and nobody will be the wiser). Also, once I get over some threshold of minimum quality in the teaching evaluations, nobody will pay much attention to an extra half point. Thus, I have an incentive to teach a course that is below the socially optimal quality.

ML will never fix the problem of determining the teaching matrix at a department of economics and inducing a course’s “optimal” quality. The problem was never about computing an optimal solution to teaching assignments given some data. The problem is, and will always be, how to truthfully elicit the faculty’s preferences, abilities, and effort in a world where everyone has an incentive to misrepresent those preferences, abilities, and effort.

Epstein (1995) argued that the only reliable method we have found to aggregate those preferences, abilities, and efforts is the market created by simple legal rules

²⁰ It is conceivable that there is an incentive-compatible teaching request mechanism that delivers an optimal allocation (this environment is akin to a multi-good reverse auction). However, once we consider the signaling and repeated behavior I described in the main text, the mechanism probably involves an inordinate degree of complexity and is unlikely to be scalable to more complex problems than allocating who will teach econometrics next fall semester.

because it aligns, through the price system, incentives with the truthful revelation of information. The method is not perfect, and the outcomes that arises from it are often unsatisfactory. Nevertheless, like democracy, all the other alternatives, including markets regulated by complex, human-made rules or “digital socialism” based on ML, are worse.

6 Concluding remarks

ML is a handy tool. The future of economics will be quite different because of them, and, in many policy situations, the application of ML will be highly beneficial.

However, by and large, Epstein’s case for simple rules is still sound. ML will never substitute first possession, voluntary exchange, and *pacta sunt servanda* as the basis of a legal system that delivers economic growth and welfare. This result is not a surprise. We did not come up with these simple rules thanks to an enlightened legislator (or, nowadays, a blue-ribbon committee of academics “with a plan”).

The simple rules were the product of an evolutionary process. Roman law, the Common law, and *Lex mercatoria* were bodies of norms that appeared over centuries thanks to the decisions of thousands and thousands of agents (see Berman 1983, and Langbein et al. 2009). For example, Roman law became predominant in Western Europe outside England in the late Middle Ages not because kings and dukes liked it (in fact, they did not), but because armies of lawyers and business people saw that it solved their problems. Good law is nothing more than applied optimal mechanism design. The forces of evolution, by trial and error, led us to the optimal solution of such a mechanism design problem, not always tidily, but inexorably.

This process is surprisingly similar to another area of AI, reinforcement learning (RL; Sutton and Barto 2018), but in a decentralized fashion. RL consists of algorithms that use training information to evaluate the actions taken by the code according to some reward function, instead of deciding whether the action was correct. RL is mighty because the programmer might not even need to be fully explicit about the decision problem’s underlying mathematical model.

One can read the history of Western law and the simple rules that emerged from it as decentralized RL. Jurists and agents, through a combination of reasoning and experience, saw what worked and what did not. Those rules that led to Pareto improvements survived and thrived. Those that did not dwindled.

After all, there is a sharp lesson to be learned from AI: our trust in simple rules has deeper roots than our high positivist era of the administrative state recognizes.

References

- Abdulkadiroğlu, A., Angrist, J. D., Dynarski, S. M., Kane, T. J., & Pathak, P. A. (2011). Accountability and flexibility in public schools: Evidence from Boston’s charters and pilots. *The Quarterly Journal of Economics*, 126, 699–748.
- Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91, 1369–1401.

- Athey, S., Bayati, M., Imbens, G., & Qu, Z. (2019a). Ensemble methods for causal effects in panel data settings. *AEA Papers and Proceedings*, 109, 65–70.
- Athey, S., Imbens, G. W., Metzger, J., & Munro, E. M. (2019). Using Wasserstein Generative Adversarial Networks for the Design of Monte Carlo Simulations, Working Paper 26566, National Bureau of Economic Research.
- Berman, H. J. (1983). *Law and Revolution: The Formation of the Western Legal Tradition*. Cambridge: Harvard University Press.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.
- Boden, M. (2018). *Artificial intelligence: A very short introduction*. Oxford: OUP Oxford.
- Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? *The Geography of Intergenerational Mobility in the United States*, *The Quarterly Journal of Economics*, 129, 1553–1623.
- Coase, R. H. (1937). The nature of the firm. *Economica*, 4, 386–405.
- Davis, S. J., & Kahn, J. A. (2008). Interpreting the great moderation: Changes in the volatility of economic activity at the macro and micro levels. *Journal of Economic Perspectives*, 22, 155–180.
- Deaton, A., & Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, 210, 2–21.
- Duffie, D. (2020). Still the World's Safe Haven? Redesigning the U.S. Treasury Market After the COVID-19 Crisis, Working Paper 62, Hutchins Center.
- Epstein, R. (1995). *Simple rules for a complex world*. Cambridge: Harvard University Press.
- Fernández-Villaverde, J., Guerrón-Quintana, P., & Rubio-Ramírez, J. F. (2015). Estimating dynamic equilibrium models with stochastic volatility. *Journal of Econometrics*, 185, 216–229.
- Fernández-Villaverde, J., Hurtado, S., Nuño, G. (2019). Financial Frictions and the Wealth Distribution, Working Paper 26302, National Bureau of Economic Research.
- Flamm, K. (2019). Measuring Moore's Law: Evidence from Price, Cost, and Quality Indexes. In C. Corrado, J. Haskel, J. Miranda, & D. Sichel (Eds.), *Measuring and accounting for innovation in the 21st century*. Chicago: University of Chicago Press.
- Fouliard, J., Howell, M., & Rey, H. (2019). Answering the Queen: Machine Learning and Financial Crises, Working Paper, LBS.
- Goodfellow, I. J., Bengio, Y., & Courville, A. (2016). *Deep Learning*. Cambridge: MIT Press.
- Guerrón, P. (2009). Money demand heterogeneity and the great moderation. *Journal of Monetary Economics*, 56, 255–266.
- Hall, R. E. (1978). Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence. *Journal of Political Economy*, 86, 971–987.
- Hayek, F. A. (1945). The use of knowledge in society. *American Economic Review*, 35, 519–530.
- Kendrick, J. W. (1961). *Productivity trends in the United States*. Cambridge: National Bureau of Economic Research.
- Langbein, J. H., Lerner, R. L., & Smith, B. P. (2009). *History of The Common Law: The development of Anglo-American legal institutions*. Philadelphia: Wolters Kluwer Law & Business.
- Levy, D. M., & Peart, S. J. (2016). *Socialist calculation debate* (pp. 1–10). London: Palgrave Macmillan UK.
- Lubik, T. A., & Schorfheide, F. (2004). Testing for indeterminacy: An application to U.S. *Monetary Policy*, *American Economic Review*, 94, 190–217.
- Lucas, R. J. (1976). Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy*, 1, 19–46.
- Manzi, J. (2012). *Uncontrolled: The surprising payoff of trial-and-error for business, politics, and society*. New York: Basic Books.
- McCulloch, W., & Pitts, W. (1943). A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 127–147.
- Medina, E. (2011). *Cybernetic Revolutionaries: Technology and Politics in Allende's Chile*. Cambridge: MIT Press.
- Moore, G. E. (1965). Cramming More Components onto Integrated Circuits, *Electronics*, 38. <https://archive.computerhistory.org/resources/access/text/2017/03/102770822-05-01-acc.pdf>.
- Morozov, E. (2019). Digital Socialism? *The Calculation Debate in the Age of Big Data*, *New Left Review*, 116, 19–46.

- Muralidharan, K., Niehaus, P., Sukhtankar, S. (2017). General equilibrium effects of (Improving) public employment programs: Experimental evidence from India, NBER Working Papers 23838, National Bureau of Economic Research, Inc.
- Peters, B. (2016). *How not to network a nation: The uneasy history of the soviet internet*. Cambridge: MIT Press.
- Phillips, L., & Rozworski, M. (2019). *The People's Republic of Walmart: How the world's biggest corporations are laying the foundation for socialism*. New York: Verso Books.
- Polanyi, M. (1966). *The tacit dimension*. New York: Doubleday.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323, 533–536.
- Russell, S., & Norvig, P. (2010). *Artificial intelligence: A modern approach*. New Jersey: Prentice Hall.
- Saros, D. (2014). *Information technology and socialist construction: The end of capital and the transition to socialism*. New York: Taylor & Francis.
- Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., et al. (2018). A General Reinforcement Learning Algorithm that Masters Chess. *Shogi, and Go through Self-play, Science*, 362, 1140–1144.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). Cambridge: The MIT Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.