PIER Working Paper

20-010

Simple Rules for a Complex World with Artificial Intelligence

JESÚS FERNÁNDEZ-VILLAVERDE
University of Pennsylvania

March 20, 2020

Simple Rules for a Complex World
with Artificial Intelligence

Jesús Fernández-Villaverde

University of Pennsylvania *

March 20, 2020

Abstract

Can artificial intelligence, in particular, machine learning algorithms, replace the idea of simple rules, such as first possession and voluntary exchange in free markets, as a foundation for public policy? This paper argues that the preponderance of the evidence sides with the interpretation that while artificial intelligence will help public policy along important aspects, simple rules will remain the fundamental guideline for the design of institutions and legal environments where markets operate. "Digital socialism" might be a hipster thing to talk about in Williamsburg or Shoreditch, but it is as much of a chimera as "analog socialism."

Keywords: Artificial intelligence, machine learning, economics, law, rule of law.

JEL codes: D85, H10, H30.

The peculiar character of the problem of a rational economic order is determined precisely by the fact that the knowledge of the circumstances of which we must make use never exists in concentrated or integrated form but solely as the dispersed bits of incomplete and frequently contradictory knowledge which all the separate individuals possess. The economic problem of society is thus not merely a problem of how to allocate "given" resources –if "given" is taken to mean given to a single mind which deliberately solves the problem set by these "data." It is rather a problem of how to secure the best use of resources known to any of the members of society, for ends whose relative importance only these individuals know. Or, to put it briefly, it is a problem of the utilization of knowledge which is not given to anyone in its totality.

F.A. Hayek, *The Use of Knowledge in Society (1945).*

1 Introduction

It is hard to open the internet or browse a bookstore without finding an article or a book discussing artificial intelligence (AI). Most of them focus on machine learning (ML), a subfield of AI that is growing in popularity in economics. We are told, with more or less accuracy, that deep learning algorithms can master the most complicated games of strategy, that robots threaten millions of jobs, and that dominance in AI will bring world supremacy.[1] Gone is the era of large armies and powerful navies. The nations with the best algorithms, as the Roman said, "will rule mankind, and make the world obey."

One thread that increasingly appears in these works is the ability of AI to address policy questions.[2] Some proposals are mild enough. For example, thanks to ML, we might be able to detect early strains in the financial markets and allow regulatory agencies to respond before damage occurs (Fouliard et al., 2019). In these proposals, AI and ML do not play a different role than traditional econometric methods used by economists since the 1940s; they just provide a more flexible statistical approach. As such, I do not have any particular problem with this application of AI and ML, and I have defended such a use.

Nevertheless, some authors go further and argue that we no longer need simple rules, such as first possession, voluntary exchange, and *pacta sunt servanda* so eloquently defended by

---

[1]Among the best of a crop of uneven quality, see Silver et al. (2016) for how deep learning algorithms mastered the ancient, but fiendishly complicated, game of Go, Baldwin (2019) on the impact of AI on the labor market, and Lee (2018) on international competition for supremacy in AI. Athey and Imbens (2019) is an outstanding survey of ML in economics.

[2]There is a collateral discussion, outside the scope of this paper, about the ethics of AI, for example, concerning its tendency to engage in statistical discrimination, and what policy should do about it (see Coeckelbergh, 2020).

Epstein (1995), to create well-behaved markets and allocate resources. Imagine the following scenario. Thanks to the development of central bank digital currencies, a central bank could, theoretically, take over financial intermediation without the logistic nightmare of setting up a national network of local branches (Fernández-Villaverde et al., 2020, describe how such a system could work). In such an environment, ML algorithms could determine the whole path of interest rates and risk premia that the central bank would pay to savers and charge to depositors. If ML were to do a good job at this task, not only would we stabilize the economy by deft control of interest rates, but financial crises and the socialization of private losses would disappear. Instead of the vagaries of investors and savers signing individual contracts, the ML algorithm would tell us the appropriate conditions for our intertemporal decisions.

Pushing the argument slightly further, AI might even realize the dream of a planned economy (although not necessarily centralized) that would deliver efficiency and equity for all. Thanks to AI, perhaps we can finally eliminate markets (see Saros, 2014, and Phillips and Rozworski, 2019, for this thesis and Medina, 2011, for a fascinating description of an early attempt in Allende's Chile called Project Cybersyn). Morozov (2019) has even coined a term for this new economic system, "digital socialism," and *The Economist* dedicated a long essay to the topic in its 2019 Christmas issue.[3] These speculations return us to classic discussions in economics from the 20th century, such as the role of experts (although, in this case, the "expert" is software code) or the socialist calculation debate.

In this paper, I will argue that while AI and ML are handy tools, there are reasons to be cautious about their broad applicability to questions of policy. More concretely, I will build a case based on three increasingly more serious barriers that ML faces in real-life policy applications.

The first barrier is that ML requires enormous data sets 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 improves upon existing methods. However, we are not (and will never be) remotely close to having the amount of data required to automatize financial intermediation.

The second barrier is that, except in a few cases, ML suffers from the Lucas critique (Lucas, 1976). Any type of estimation with non-experimental data is subject to an external validity constraint: economic agents make decisions based on expectations about policy regimes. 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 re-compute the optimal responses to the new

---

[3] "Beware the Borg: Essay." *The Economist*, vol. 433, no. 9174, 21 Dec. 2019, p. 58 (US).

policy. One can go one step further and realize that such a structural model should also incorporate the probability of policymakers changing the environment. By construction, AI and ML have very little to say about those structural models as they are reduced-form statistical representations.[4]

The third barrier, which I find the most unconquerable, comes from how societies create and use knowledge. ML cannot circumvent the problems of central planning, either on a large scale or at a much lower range. I started the paper with a quote from Hayek (1945) as a reminder that his insights are as relevant today as in 1945, well before anyone knew about AI or ML. The core of the allocation problem in a society is not how to solve an optimization problem, which AI or ML can do better in large dimensions than traditional mathematical methods. The fundamental barrier to planning is that information is dispersed and agents do not have incentives (and often not even the capabilities) to disclose such information to a central planner. The trouble with the Soviet Union was not that its computers were not powerful enough (although they were not) or that its planning algorithms were poor (they were terrible), it was that central planning is, as medieval scholastics loved to say, inefficient *in essentia sua*.[5]

Thus, I will conclude by defending that the lessons we know about what constitutes good and bad economic policies are likely to remain unchanged. "Digital socialism" might be a hipster thing to talk about in Williamsburg or Shoreditch, but it is as much of a chimera as "analog socialism." Instead, I will argue that the simple rules we have developed over centuries, including the foundations of our legal systems and market economies, came out of an evolutionary process that closely resembles another subfield of AI, reinforcement learning, except in a more suitable, decentralized fashion.

Before I move on, however, I must warn the reader that this is a discussion where I have "skin-in-the-game." I have written papers showing how ML can be applied in macroeconomics and econometrics and I teach ML to graduate students. I find research on ML fascinating, and I plan to continue working on AI and ML. But precisely because of this background, I appreciate it is crucial to frame the promises of AI realistically and avoid the disappointment that could come from overpromising. Neither AI nor public policy will be well served by hyperbolic vows.

The rest of the paper is organized as follows. Section 2 provides basic technical back-

---

[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 they can address is limited. Second, ML can help in the estimation of structural models, but that is far from what the defenders of "digital socialism" have in mind.

[5]See Kitov and Krotov (2017) for a description of the computers used by Gosplan in the Soviet Union, Heal (1973) for a survey of the state-of-the-art central planning algorithms, and Kornai (1992) for the depressing realities of socialist planning.

ground on AI. Section 3 explains the data requirements of AI. Section 4 reviews the famed Lucas critique and how it applies to AI. Section 5 discusses central planning and ML. Section 6 concludes with some remarks about the evolutionary origin of simple rules and how it compares with ML.

2 Some basic background

AI is a vast field, from expert systems to machine learning and from robotics control to natural language processing. Although modern AI started in a famed 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 the forefront of the public discourse.[6]

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.[7]

An example will help us understand the intuition. Imagine that we specify a function that maps dozens of observables from credit card transactions (e.g., the geolocation of the store where the credit card is used, the time of day and week when the transaction occurred, the 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 example, 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 previous credit card transactions and the 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 environments where the researcher might have limited knowledge. In that sense, ML is very far away from the "artificial general intelligence" of Hollywood's dystopias that was pursued, fruitlessly, during the 1960s.

---

[6] 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]. I use Goodfellow et al. (2016) to teach ML to graduate students in economics. Those in a hurry can learn much from Alpaydin (2016) and Boden (2018).

[7] 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." Many English-speaking researchers used to talk about "statistical learning," but that term has lost popularity.

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 the industry boomed. There are three reasons for such a long lag.

The first reason is that, while some key algorithms such as artificial neural networks were known for decades, the computational capability to implement them outside simple "proof-of-concept" cases was not widely available at cheap prices until two 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. comes with multiple central processing units (CPUs) and a graphics processing unit (GPU). Moreover, online services such as Amazon Web Services mean that any researcher can have access over the internet, for a few dollars an hour, to computers that were previously open only to researchers at large universities and national labs.[9] 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 data sets. 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 papers with tens of millions of observations (Chetty et al., 2014, use 47.8 million).[10] That is why AI and ML are often associated with expressions such as "big data" or "data analytics." As I will discuss in Section 3, ML algorithms are data-hungry, and there is little you 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 in 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 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 how well will ML work when applied to public policy?

---

[8]Moore's law, proposed by Gordon Moore, the co-founder of Intel, in 1965 states that the number of transistors per integrated circuit will double every 18 months (Moore, 1965). A way to think about Moore's law is that the ability of humans to perform numerical computations advanced as much between September 2018 and March 2020, when these lines are being written, as between the dawn of our species around 300,000 years ago and September 2018. See Flamm (2019) for the astonishing empirical success of Moore's prediction.

[9]Fernández-Villaverde and Valencia (2018) describe how massive parallelization has changed modern computational economics.

[10]I picked these three papers because they are considered landmark empirical works of their cohorts.

3 ML and data

The "dirty little secret" of ML is that, to work properly, it requires incredibly large data sets. 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., as I mentioned before, 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.[11]

When do we have these large data sets? 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.[12] 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. This path might seem counter-intuitive (or plainly dishonest), but it is actually easy and consistent in many environments. An 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 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 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 wining strategies (plus some changes to keep exploring new strategies). See Athey et al. (2019b) and Fernández-Villaverde et al. (2019) for how similar ideas can be applied in economics.

Unfortunately, in many problems of interest in public policy, we do not have access to such a wealth of data and, most likely, we never will. Take monetary policy, a 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)? I am skeptical.

---

[11]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.

[12]See https://www.statista.com/statistics/546894/number-of-amazon-prime-paying-members/

When the members of the FOMC set the federal funds rate, 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 quarterly data for output, consumption, and investment after World War II.[13] If we count them, from 1947:Q1 (the first "good" observation in terms of accuracy) until 2019:Q4 (the last observation as I write these lines), we have 292 data points.

But, during that time, 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). Those structural changes mean that, 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) show how these estimates change sharply depending on whether we include early observations.

Structural change matters for AI and 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 is 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) can enrich the observations, but we will still encounter severe limitations on the length and stability of micro surveys. Think, as an illustrative 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 clever 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 an insightful discussion of these topics.[14]

---

[13]There have been 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.

[14]I am not criticizing the use of ML with microdata to provide the FOMC with better information. I aspire to do that with my research. I deny the possibility that you can substitute an ML algorithm for the FOMC.

4 ML and the Lucas critique

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

Imagine that we obtain data on airline ticket prices and the occupancy rate of particular flights. Any 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 (a flight I often took a few years ago when I was invited to teach at Harvard) and low prices in the relatively empty 3:00 pm Tuesday flights from Boston to Philadelphia (my return leg the next day).

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 (all those weekly commuters!). However, in reaching such a conclusion, we have relied on basic economic theory and shown an essential limitation of ML: it is tough to use it to assert causality (although we are making 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 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 economy and business class tickets. 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 (I am paying it out of pocket, so university regulations do not bind me). My Monday lecture starts at 10:00 am, and I want to avoid the risk of suffering the discomfort of a small seat.

American Airlines prefers that I buy a business ticket 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 class 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 of the choice problem. For example, if American Airlines tightens its rules for upgrades (as it did a few years ago), my price elasticity of demand will increase. Why? The tightening of the upgrade rules hurt travelers with many segments per year of cheap tickets (i.e., the Philadelphia-Boston commuters). This change, implicitly, favored business travelers with few segments per year but of high price (my case: all those expensive business class tickets to Europe) 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. This might achieve American Airlines' goal (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 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, plenty of policies are hard to test by experimentation. The first barrier is ethical: as a society, we cannot 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. One can implement a randomized control trial (RCT) to evaluate the effect of charter schools, but one cannot change the federal funds rate to see what happens with the U.S. economy afterward. 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 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   ML and central planning

Over the last few years, a few observers have made the bold prediction that, thanks to AI, central planning is about to return (Saros, 2014, Phillips and Rozworski, 2019, and Morozov, 2019). Some of these observers are rather prominent. For example, Jack Ma, founder of Alibaba, stated in November 2016:

> 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...
>
> 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

---

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

[16]The statements in the main text require a few 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 the full model for counterfactual policy analysis. But even in that case, we still need a structural model, and ML cannot substitute for it.

smarter and make it possible to plan and predict market forces so as to allow us to finally achieve a planned economy.[17]

These proposals forget the final lesson of the socialist calculation debate, which came from Hayek (1945). The objections to central planning are not that solving the associated optimization problem is extremely complex, which it is and increasingly so in an economy with a maddening explosion of products, or that we need to gather the data and process it 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 objections to central planning are that the information one needs to undertake it 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 and innovative activity. As Steve Jobs put it: "A lot of times, people don't know what they want until you show it to them."[18]

A real-life application of central planning illustrates the point. Every year, the department of economics at the University of Pennsylvania must set up a teaching matrix for the next academic year.[19] Each member of the faculty submits his or her preferences in terms of courses to be taught, day of the week, time of day, etc. Given the teaching needs and submitted requests, the computational burden of finding the optimal allocation is manageable. We have around 32 faculty members and, once you consider that the average theorist will never request to teach econometrics and vice versa, the permutations to consider are limited. A couple of hours in front of `Excel` delivers the answer: it seems that the central teaching planner at Penn economics can do her job.

The real challenge is that, when I submit my teaching requests, I do not have an incentive to i) reveal the truth about my preferences or ii) think hard about developing a new course.[20] 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 if I do so. However, I will not be compensated for the extra effort, even if it is not high. Thus, I have an incentive to request a small section for advanced undergrads on an old-fashioned topic. This request

---

[17]See http://www.globaltimes.cn/content/1051715.shtml.

[18]Quoted in *Business Week*, 25 May 1998.

[19]I pick the example of a small organization, my department, to illustrate that the problem of planning goes well beyond the formidable task of running a national economy. Instead, it is pervasive to all forms of collective organization, especially those that, because of transaction costs, cannot rely on a price system (Coase, 1937).

[20]In my example, due to space limitations, I do not discuss two important issues. First, tacit knowledge and the difficulties in transmitting it, a point already presented by Hayek (1945), and emphasized by Polanyi (1966, p. 4) when he explained how "we can know more than we can tell." Second, the public choice problems driving the central teaching planner. While those are often considerable, I want to make the sharp point that even a benevolent central teaching planner that does not face aggregation of individual preferences problems faces daunting challenges.

is not optimal: if Penn could pay me an extra stipend, I would teach the large, innovative section, the students would be happier, and I would be wealthier.[21]

An obvious solution would be not to submit a teaching request, but a schedule of teaching supply curves, i.e., I will teach "the economics of big data" at 9.00 am on Mondays and Wednesdays at a price x or "advanced monetary theory" at 1.00 pm on Tuesdays and Wednesdays at a price $0.4x$. This scheme resembles some of the "market socialism" proposals put forward by Abba Lerner and Oskar Lange or, more recently, Roemer (1994). The central teaching planner will use the supply curves to clear the teaching market and assign a faculty member to each course "imitating the market." 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.

However, a system of teaching 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 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.

In fact, knowing that the department will assign duties using teaching supply curves, I can manipulate from the day I am hired how I behave in front of my colleagues. 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 supply curves in the future. My colleagues would know that and act accordingly, changing their supply curve to reflect that they understand that 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.[22]

There is an additional problem. Once I am assigned a course, how does Penn ensure that I teach it at the "optimal" quality level? "Optimal" cannot mean the highest possible quality. If I were to prepare every lecture as much 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.

---

[21]Sometimes deans offer small teaching grants to reward innovation in teaching, but those are rarely worth the time to fill in the application form. Consequently, we do not see much advance in teaching technologies.

[22]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.

Even forgetting about that intertemporal aspect, how do we trade off one extra minute of research (which increases Penn's academic reputation) with one extra minute of teaching preparation?[23] And how do we address heterogeneity in the comparative ability between research and teaching among faculty members when the amount of effort exerted in each activity is not observable?

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., I can always "lose" the terrible teaching evaluation I got 15 years ago and nobody will be the wiser; most recruiting committees only ask for recent evaluations). 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 below the socially optimal quality.

ML cannot address any of the concerns of the previous paragraphs, no matter how fancy the learning algorithm we apply or how much data we get from apps that the faculty and students use, because getting around information and incentive constraints is not what ML can do. Thus, ML will never fix the problem of how to determine the teaching matrix at Penn economics and to induce the "optimal" quality of the course. The problem was never about computing an optimal solution to teaching assignments given some data. The problem is, and will always be, determining the preferences, abilities, and effort of the faculty in a world where everyone has an incentive to misrepresent those preferences, abilities, and effort.

The only reliable method we have found to aggregate those preferences, abilities, and efforts is the market because it aligns, through the price system, incentives with information revelation. The method is not perfect, and the outcomes that arise from it are often unsatisfactory. Also, in some cases, such as setting up a teaching matrix at my department, there are not enough agents to set up a meaningful market and students and faculty end up unhappy. Nevertheless, as with democracy, all the other alternatives, including "digital socialism," are worse.[24]

---

[23]I once had this conversation with my deans. The answer I got was: "You need to prepare your lectures enough," which is, obviously, a meaningless platitude. I am also forgetting about the sad reality that what students think they like in a lecture when they are 18 is very different from what they would realize is essential when they are 45. Therefore, their perception of "quality" might be quite different from the socially optimal quality.

[24]Note, however, that there is the concern that ML may worsen market outcomes. Similar ML algorithms run by different companies might end up engaging in tacit collusion through the implicit correlation created in the experiments that each firm performs. See Hansen et al. (2020).

6 Concluding remarks

AI and ML are incredibly useful tools. The future of economics will be quite different because of them, and, in many policy situations, the application of AI and ML will be highly beneficial.

However, by and large, we should still rely on markets to allocate resources. Moreover, markets work when they operate under simple rules, such as first possession, voluntary exchange, and *pacta sunt servanda*. 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 (Berman, 1983). Roman law, for example, 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 to 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 comprises 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 underlying mathematical model behind the decision problem.

I 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.

There is a sharp lesson from AI: our trust in simple rules and the markets they create has deeper roots than our high positivist era of the administrative state recognizes.

References

ABDULKADIROĞLU, A., J. D. ANGRIST, S. M. DYNARSKI, T. J. KANE, AND P. A. PATHAK (2011): "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters And Pilots," *The Quarterly Journal of Economics*, 126, 699–748.

ACEMOGLU, D., S. JOHNSON, AND J. A. ROBINSON (2001): "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 91, 1369–1401.

ALPAYDIN, E. (2016): *Machine Learning: The New AI*, MIT Press Essential Knowledge, MIT Press.

ATHEY, S., M. BAYATI, G. IMBENS, AND Z. QU (2019a): "Ensemble Methods for Causal Effects in Panel Data Settings," *AEA Papers and Proceedings*, 109, 65–70.

ATHEY, S. AND G. W. IMBENS (2019): "Machine Learning Methods That Economists Should Know About," *Annual Review of Economics*, 11, 685–725.

ATHEY, S., G. W. IMBENS, J. METZGER, AND E. M. MUNRO (2019b): "Using Wasserstein Generative Adversarial Networks for the Design of Monte Carlo Simulations," Working Paper 26566, National Bureau of Economic Research.

BALDWIN, R. (2019): *The Globotics Upheaval: Globalisation, Robotics and the Future of Work*, Orion.

BERMAN, H. J. (1983): *Law and Revolution: The Formation of the Western Legal Tradition*, Harvard University Press.

BLEI, D. M., A. Y. NG, AND M. I. JORDAN (2003): "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 3, 993–1022.

BODEN, M. (2018): *Artificial Intelligence: A Very Short Introduction*, OUP Oxford.

CHETTY, R., N. HENDREN, P. KLINE, AND E. SAEZ (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.

COECKELBERGH, M. (2020): *AI Ethics*, MIT Press.

DAVIS, S. J. AND J. A. KAHN (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. AND N. CARTWRIGHT (2018): "Understanding and misunderstanding randomized controlled trials," *Social Science & Medicine*, 210, 2 – 21, randomized Controlled Trials and Evidence-based Policy: A Multidisciplinary Dialogue.

EPSTEIN, R. (1995): *Simple Rules for a Complex World*, Harvard University Press.

FERNÁNDEZ-VILLAVERDE, J., P. GUERRÓN-QUINTANA, AND J. F. RUBIO-RAMÍREZ (2015): "Estimating Dynamic Equilibrium Models with Stochastic Volatility," *Journal of Econometrics*, 185, 216 – 229.

FERNÁNDEZ-VILLAVERDE, J., S. HURTADO, AND G. NUÑO (2019): "Financial Frictions and the Wealth Distribution," Working Paper 26302, National Bureau of Economic Research.

FERNÁNDEZ-VILLAVERDE, J., D. SANCHES, L. SCHILLING, AND H. UHLIG (2020): "Central Bank Digital Currency: Central Banking For All?" Working Paper 14337, Centre for Economic Policy Research.

FERNÁNDEZ-VILLAVERDE, J. AND D. Z. VALENCIA (2018): "A Practical Guide to Parallelization in Economics," Working Paper 24561, National Bureau of Economic Research.

FLAMM, K. (2019): "Measuring Moore's Law: Evidence from Price, Cost, and Quality Indexes," in *Measuring and Accounting for Innovation in the 21st Century*, ed. by C. Corrado, J. Haskel, J. Miranda, and D. Sichel, University of Chicago Press.

FOULIARD, J., M. HOWELL, AND H. REY (2019): "Answering the Queen: Machine Learning and Financial Crises," Working Paper, LBS.

GOODFELLOW, I. J., Y. BENGIO, AND A. COURVILLE (2016): *Deep Learning*, 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.

HANSEN, K., K. MISRA, AND M. PAI (2020): "Algorithmic Collusion: Supra-competitive Prices via Independent Algorithms," Working Paper 14372, Centre for Economic Policy Research.

HAYEK, F. A. (1945): "The Use of Knowledge in Society," *American Economic Review*, 35, 519–530.

HEAL, G. (1973): *The Theory of Economic Planning*, North-Holland.

KENDRICK, J. W. (1961): *Productivity Trends in the United States*, National Bureau of Economic Research.

KITOV, V. AND N. KROTOV (2017): "The Main Computer Center of the USSR State Planning Committee (MCC of Gosplan)," in *2017 Fourth International Conference on Computer Technology in Russia and in the Former Soviet Union (SORUCOM)*, 227–232.

KORNAI, J. (1992): *The Socialist System: The Political Economy of Communism*, Clarendon Press.

LEE, K.-F. (2018): *AI Superpowers: China, Silicon Valley, and the New World Order*, HMH Books.

LUBIK, T. A. AND F. SCHORFHEIDE (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*, Basic Books.

MCCULLOCH, W. AND W. PITTS (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*, MIT Press.

MOORE, G. E. (1965): "Cramming More Components onto Integrated Circuits," *Electronics*, 38.

MOROZOV, E. (2019): "Digital Socialism? The Calculation Debate in the Age of Big Data," *New Left Review*, 116, 19–46.

MURALIDHARAN, K., P. NIEHAUS, AND S. SUKHTANKAR (2017): "General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence from India," NBER Working Papers 23838, National Bureau of Economic Research, Inc.

PHILLIPS, L. AND M. ROZWORSKI (2019): *The People's Republic of Walmart: How the World's Biggest Corporations are Laying the Foundation for Socialism*, Verso Books.

POLANYI, M. (1966): *The Tacit Dimension*, Doubleday.

ROEMER, J. (1994): *A Future for Socialism*, Verso.

RUMELHART, D. E., G. E. HINTON, AND R. J. WILLIAMS (1986): "Learning Representations by Back-propagating Errors," *Nature*, 323, 533–536.

RUSSELL, S. AND P. NORVIG (2010): *Artificial Intelligence: A Modern Approach*, Prentice Hall.

SAROS, D. (2014): *Information Technology and Socialist Construction: The End of Capital and the Transition to Socialism*, Taylor & Francis.

SILVER, D., A. HUANG, C. J. MADDISON, A. GUEZ, L. SIFRE, G. VAN DEN DRIESSCHE, J. SCHRITTWIESER, I. ANTONOGLOU, V. PANNEERSHELVAM, M. LANCTOT, S. DIELEMAN, D. GREWE, J. NHAM, N. KALCHBRENNER, I. SUTSKEVER, T. LILLICRAP, M. LEACH, K. KAVUKCUOGLU, T. GRAEPEL, AND D. HASSABIS (2016): "Mastering the Game of Go with Deep Neural Networks and Tree Search," *Nature*, 529, 484–489.

SILVER, D., T. HUBERT, J. SCHRITTWIESER, I. ANTONOGLOU, M. LAI, A. GUEZ, M. LANCTOT, L. SIFRE, D. KUMARAN, T. GRAEPEL, ET AL. (2018): "A General Reinforcement Learning Algorithm that Masters Chess, Shogi, and Go through Self-play," *Science*, 362, 1140–1144.

SUTTON, R. S. AND A. G. BARTO (2018): *Reinforcement Learning: An Introduction*, The MIT Press, second ed.