# *Computational Learning in Macroeconomics, 1970-2022*

## Farewell Lecture at Gabelli Finance

Paul D. McNelis, S.J.[1]

January 25, 2023

[1]Professor Emeritus, Gabelli School of Business, Fordham University

# Wastelands, Blind Sightings, Sunny Uplands

- 1970's: Macroeconometric Models vs. Simple Calibrated Models: Multipliers and Parameter Stability (Niehans, Heffernan, and Neftci)
- 1980's: Hyperinflations in Latin America and Israel: Indexation (Barbosa, Bigman)
- 1990's: Irish Experience and Discovering Asia: Liberalization and Capital Controls (Browne, Corbo)
- 2000's: Discovering Neural Nets for Forecasting and Solving Macro Models (McAdam, Lim, Duffy)
- 2010's: Becoming a Bayesian with Counterfactual Simulations: Gold Standard and Singapore Exchange-Rate Management (Fagan, Lothian, Chow, Lim)
- 2020's: Big Data and Machine Learning: Banking and Contagion (Lai)

# Wastelands: Big Mods with Little Theory

- The age of big econometric models in the 1960's: FRB-MIT-Penn, Michigan Quarterly Econometric Model, Brookings, Hickman-Coen, DRI, Klein LINK
- When I applied to grad school, wanted to go to Penn and be a student of Klein
- Large grant from Ford Foundation to link country models, called Project Link. One student goes to Paris, another Rome, with meetings in exotic places every six months
- Simulate the linked country-specific large models, compare multipliers, sell forecasts to industry or government agencies
- Intellectual wasteland: specification of the equations had little or nothing to do with what we were learning in theory

# Zeitgeist: Analytical Dynamics

- Interest was on dynamics, age of comparative statics was over
- We used two dimensional dynamic processes, looked for analytic conditions of stability
- For a matrix with two state variables, $k_t$, $\pi_t$, conditions for stability were a negative trace and positive determinant
- $\frac{\partial \dot{k}}{\partial k} + \frac{\partial \dot{\pi}}{\partial \pi} < 0$
- $\frac{\partial \dot{k}}{\partial k}\frac{\partial \dot{\pi}}{\partial \pi} - \frac{\partial \dot{k}}{\partial \pi}\frac{\partial \dot{\pi}}{\partial k} > 0$
- Again, a few good papers, eg. Miguel Sidrauski on money and growth, but not much was happening with two-dimensional dynamic models
- As usual folks were quibbling about ranges of parameter values for stability. Not much insight. An intellectual wasteland.
- So how to work with a dynamic model with more than two dynamic equations: numerical simulation

# Beginnings of Computation

- A friend suggested numerical simulation of dynamic models with three for four dynamic equations.
- Use Fortran IV for solving dynamic adjustment paths from one steady state to another, on a DEC10 Mainframe at Johns Hopkins
- Too small to estimate with econometric methods but too large for analytical solutions.
- But I could "parameterize" the functions and then simulate the models for different policy changes. We now call this a "calibration" exercise.
- Alice Orcutt, daughter of Guy Orcutt of Yale, a fellow grad student, tried to talk me out of this approach.
- She was a true-blue econometrician. Now in Canada, she is the mother of Emi Nakamura of Berkeley, recent Clark medalist, who does numerical simulation

# How to simulate?

- Issue was how to simulate? I took courses in Numerical Analysis at Johns Hopkins Applied Physics Lab the summer after my second year and passing comps. Really crazy: should have gone on a long vacation. I did take some tennis lessons, to be sure.
- At that time we had to use punch cards, and submit punch cards to a computer center and get turnaround a few hours later. I have to call on Fortran 4 subroutines based on the Runge-Kutta method for numerical integration.
- One advantage, when air-conditioning was not as widespread, was that the computer centers were well air conditioned.
- Use of the Runge-Kutta method was widely used by NASA for the Apollo program.
- Idea of round-off error: for simulation, is $1/3 = .33$ or $.3333333$? When does truncation because a serious problem?

# Taylor Expansion

- Alternative to Second Order Taylor Expansion
- Based on first and second-order derivatives, which may not exist!

$$y_t = y_0 + f'(x_0, y_0)\Delta x_0 + .5 * \Delta x_0' f''(x_0, y_0)\Delta x_0 \tag{1}$$

# Runge-Kutta

- This was an alternative to Taylor expansion
- No need to take derivatives

$$y_1 = y_0 + f(y_0, x_0)$$

$$\begin{aligned} y_1 &= y_0 + \tfrac{1}{6}(k1 + 2k2 + 2k3 + k4) \\ k1 &= hf(x_0, y_0) \\ k2 &= hf(x_0 + .5h, y_0 + .5k1) \\ k3 &= hf(x_0 + .5h, y0 + .5k2) \\ k4 &= hf(x_0 + h, y_0 + k3)(2) \end{aligned}$$

# Theological Interlude

- I had an interlude of Theology studies for three years but worked summers at Johns Hopkins with Jurg Niehans. I followed in the footsteps of Adam Smith and Thomas Bayes as well as the Salamanca Jesuits.
- At that time, most of us were blind-sided by Rational Expectations. Initial papers were rejected because I was using adaptive expectations instead of rational expectations.
- One key insight of my thesis was that once-over money-stock multipliers are much larger that once-over expenditure changes, if we do a proper stock/flow comparison.
- To finance a deficit due to an expenditure change, there has to be continued bond expansion over many periods.
- This is not a proper comparison with a once-over money supply change. So the comparison should be once-over bond-finance expenditure changes, up and down, with once-over money stock changes.

# Being in the Right Place

- I could not use my simulation model to show this (it was neither econometric nor rational expectations). So I went to the FED and got a tape of the FRB-MIT-Penn model and simulated it.
- I used this to show that big Keynesian models produced strong monetarist results
- I presented this at a conference in Europe, which Karl Brunner attended. He liked the paper, warmly applauded. Only later did I know he was editor of the Journal of Monetary Economics. So I sent the paper there. It certainly helps to be in the right place at the right time.
- I was disowned and excommunicated by several professors who taught me at Boston College as an undergraduate.
- How dare I point out that the FED-MIT-Penn model had monetarist conclusions? One former undergraduate professor admitted that he even simulated the model to try to overturn my results.

# Kalman Filter

$$y_t = x_t \beta_t + e_t$$

$$\beta_t = \beta_{t-1} + u_t$$

$$\beta_t = \beta_{t-1} + \mathbf{K}_t[y_{t-1} - \beta_{t-1} x_{t-1}] (3)$$

- First equation is the measurement equation and second is the state equation
- In the third equation, $\mathbf{K}_t$ is the Kalman gain
- It is a function of the time-varying variance-covariance matrix hyperparameters, Q,R, of the Linear Quadratic Regulator problem $e_t' Q e_t + u_t' R u_t$
- Objective function: minimize measurement equation errors with small changes in the controls $u_t$
- Basic idea of the Kalman filter is to find an optimal weighting scheme of past data to give the best one-period predictions

# Sargent and Fair on Supply Side Specification

- Sargent published a classical macro-econometric model of the US with rational expectations. Ray Fair put out a Keynesian econometric model with adaptive expectations.
- Salih was at GW at the time. I started to use TROLL for teaching and research, a processor for main frames dedicated to model estimation and simulation, even with foreward-looking expectations.
- Salih discovered that it had a Kalman filtering package, to estimate time-varying parameters (with a bit of micro-surgery on some algorithms). One idea: compare the parameter stability of the Sargent supply side equation with that of the Fair equation. After all, the Lucas criticism is about proper specification of models. Models with rational expectations should have more stable parameters.
- We found that the Sargent supply-side equation had less parameter variation than the Fair supply-side specification

# Blind Sightings and Picking the Wrong Battle

- Close call with JPE: Lucas rejected it with "great regret". RESTAT took it. The one negative report for RESTAT was just a xerox of the negative report from JPE.
- Sargent later told me that he had abandoned his research program of econometric estimation of RE in favor of calibrated RBC analysis and did not want to get into a fight with Ray Fair over econometric models.
- One again, blind-sighted. Profession was moving to the Kydland-Prescott age, even with Kalman Filtering
- Mantra of Minnesota Mafia: Progress not Regress (as in regression).
- Neil Ericson cited this paper many times over as the first empirical test of the Lucas critique.

# Application to Hyperinflations in Brazil and Israel

- Much time in Brazil, Chile, Argentina and Israel in the 1980's.
- I was fascinated with inflation, especially hyperinflation.
- Issue of inertial inflation: $\pi_t = \gamma_{t-1}\pi_{t-1} + \beta X_t$, with $X_t$ a collection of macro policy variables
- The higher the degree of inertia, with $\gamma_t \to 1$, the more resistant inflation stabilization was to changes in monetary/fiscal austerity policies.
- Indexation was the underlying cause of the inertial inflation.

- Two types of indexation: wage indexation and asset-market indexation.
- Asset market indexation was of two types: Crawling pegs for the exchange rate and for bank deposits.
- In Israel, the accounts were called PATAM's and in Brazil, Overnights.
- To make inflation more response to fiscal/monetary stabilization, there was need for *de-indexation*.
- Question was, which type of indexation was the key driver of the coefficient of inertia, $\gamma_t$?
- For Israel, David Bigman and I found that asset indexation mattered more. For Brazil, wage indexation was more important.

# Trade Liberalization in Chile and Korea

- One of the key questions in Latin America was that the quick trade liberalization engineered by the "Chicago boys" did rapidly reduce or moderate the pricing of domestic manufactured goods, even though they were almost identical to imported manufactured good.
- Latin America was isolated for several decades as a result of the *Singer-Prebish* hypothesis, which forecasted a declining terms of trade for commodity expots. Solution: protective tariffs for domestic manufacturing.
- In the 60's in Brazil and later in Argentina, Chile and Uruguay we saw the Southern Cone experiments with rapid trade liberalization. But prices remained high, much higher than international prices.

# Estimating Equation

$$p_{m,t} = \gamma_t S_t p^*_{m,t} + (1 - \gamma_t) X_t + \epsilon_t \tag{4}$$

- $S_t, p_{m,g}, p^*_{m,t}, X_t$ represent the exchange rate, the price of domestic manufactured goods, the foreign price of such goods and a matrix of domestic demand variables.
- As trade liberalization progresses, $\gamma_t \to 1$.
- This did not happen in Chile and Korea, when Vittorio Corbo and I compared many different indices of trade liberalization with the time-varying $\gamma_t$.
- Reason: despite trade liberalization there is a *network* of distribution, marketing and repair services which has to be put in place.
- This takes time.

# Capital Controls in Ireland-I

- In 1979 the Republic of Ireland severed its linking to the UK Pound Stering and entered the European Monetary System, linking to the DM.
- This was a major decision.
- Ireland and the UK are not BFF's, but there was no question of severing ties with UK Sterling through the initial years of independence, the trade wars of the 30's, WWII, the mass emigration of 50's, IRA troubles of the 60's.
- Yet they broke off for the EMS and eventually to the Euro. But the UK is still their major trading partner.
- After the oil shocks of the 70's, the Central Bank of Ireland attempted to impose *exchange controls* in order to drive down interest rates below that of the EMS.
- In this setting, domestic interest rates have the following process: $i_t = \gamma_t i_t^* + (1 - \gamma_t) X_t + \epsilon_t$

# Capital Controls in Ireland - II

- $i_t, i_t^*, X_t$ represent the domestic and foreign rate and domestic liquidity factors.
- Will full financial integration, of course, $\gamma \to 1$. With tight controls, $\gamma \to 0$.
- With Frank Browne of the Central Bank of Ireland, I examined the relationship between three exchange-control episodes in the 80's and the behavior of $\gamma_t$.
- Initially the drop in $\gamma_t$ lasted for about 10 months and then returned to a value close to unit.
- Thenm another round of capital controls and the interest rates dropped for about four months and bounced back.
- After a third round, it dropped for about a month and then popped back to one.
- Message: capital controls have only transitory effects at best once a financial system is fully integrated with the rest of the world.

# Neural Networks

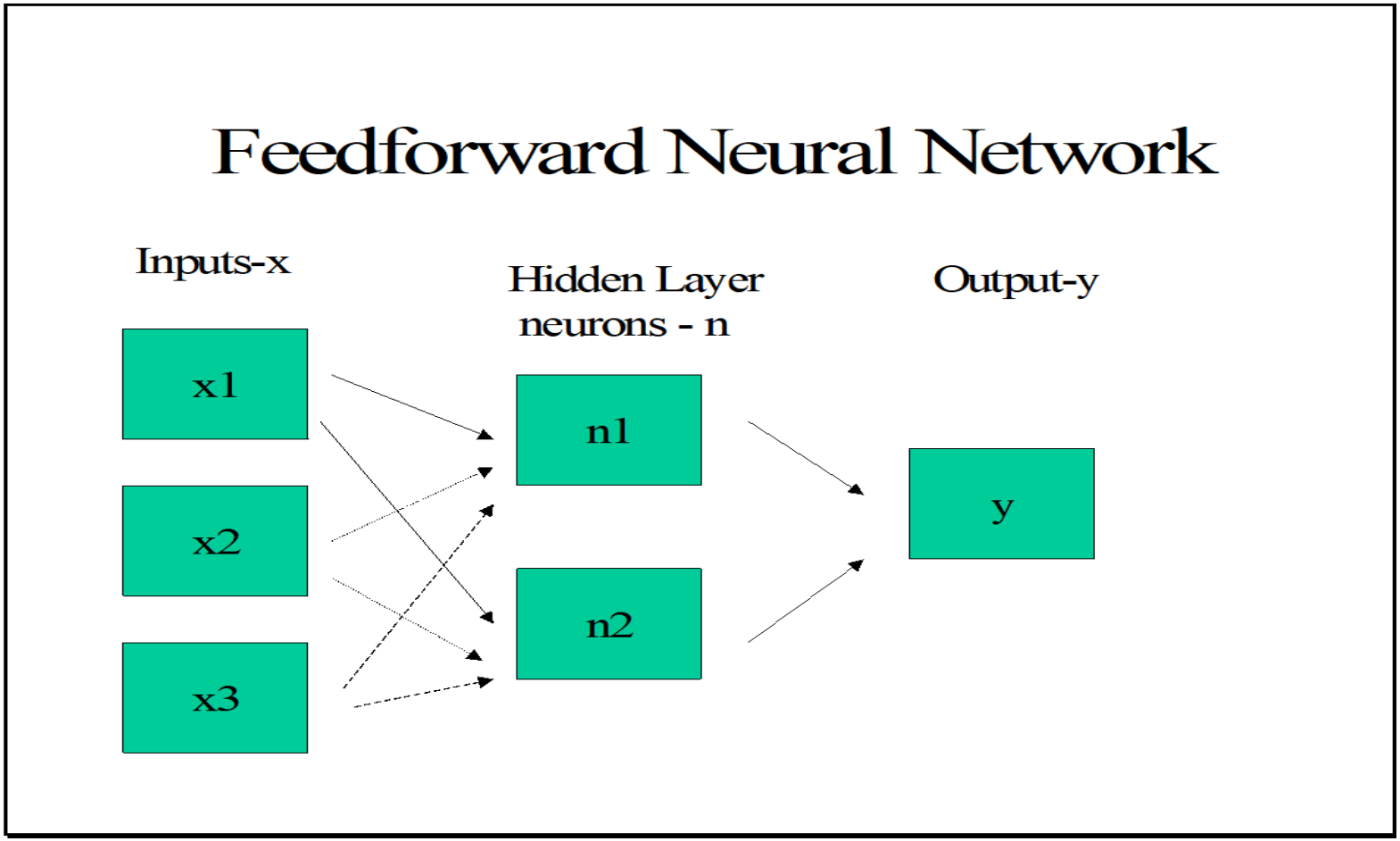
- Nagging sense the real world is nonlinear and there are threshold effects as variables reach very low or very high values
- Yet most new ways of doing rational expectations were based on the Blanchard-Kahn method for forward-looking linear models.
- In the BK method, simply find the roots of a linearized model. The number of forward-looking variables must equal the number of roots outside the unit circle
- Simulation involved the perturbation method, which is based on first-order Taylor approximation with "small shocks".
- Were the Asian Financial Crisis, the Global Financial Crisis, the Pandemic "small" shocks?
- Yet that is the working assumption for solving these models these models.

# Taking Nonlinearity Seriously

- Before the nonlinearity of the Zero Lower Bound, we observed threshold patterns in financial markets.
- Besides forecasting the solution and simulation of dynamic modeling was linear.
- The prevailing way to go beyond linearity was with polynomial approximations.
- However there is the case of the famous *curse of dimensionality*.

# Visualization of a Feedforward Network

Figure: Feedforward Network: 3 Inputs, 2 Hidden Layers, One Output



# Properties of a Network

- This single-layer feedforward or multiperceptron network with one hidden layer is the most basic and commonly used neural network in economic and financial applications.
- More generally, the network represents the way the human brain processes input sensory data, received as input neurons, into recognitionas as an output neuron.
- As the brain develops, more and more neurons are interconnected by more synapses, and the signals of the different neurons, working in parallel fashion, in more and more hidden layers, are combined by the synapses to produce more nuanced insight and response.
- Activation function is usually a logistic or logsigmoid function

# Representation

$$
\begin{aligned}
\mathsf{n}_{k,t} &= \omega_{k,0} + \textstyle\sum_{i=1}^{i*} \omega_{k,i} x_{i,t} \\
\mathsf{L}(\mathsf{n}_{k,t}) &= \frac{1}{1+exp(-n_{k,t})} \\
\mathsf{N}_{k,t} &= \mathbb{L}(n_{k,t}) \\
\mathsf{y}_t &= \gamma_0 + \textstyle\sum_{k=1}^{k*} \gamma_k N_{k,t}
\end{aligned}
\tag{5}
$$

# Thick Models

- The choice of the number of hidden layers and the choice of activation function (such as hyperbolic tangent instead of logistic functions) refers to *hyperparameters* of the network.
- Obviously there is a lot of freedom and as Lucas cautioned us, *beware of economists bearing free parameters* what do we do?
- With Peter McAdam of the ECB, we used a set of neural net specifications and created a *thick model* for forecasting inflation in the Euro Area and USA.
- For out-of-sample accuracy, this neural-net thick model outperformed linear models with many lags and polynomial expansions

# Solving a Nonlinear DSGE Model

- Problem is to solve the forward-looking difference equation

$$U'(c_t) = \beta \mathbb{E}([U'(c_{t+1})])R_t \tag{6}$$

- This is the famous Keynes-Ramsey rule for optimal saving, and the asset-pricing Kernel
- This equation has a unit root so we cannot solve $C_{t+1}$ as a function of $C_t$
- Note that the expectations operator $\mathbb{E}$ itself is a function, so that the K-R rule is a *functional* equation

# PEA

- Idea is to guess an approximation to the Expectation function
- Early stage is to take a first order Taylor expansion and solve the model with the Blanchard-Kahn method
- This is only useful for small shocks
- Marcet develop the Parameterized Expectations Algorithm (PEA) for solving the model. Guess a solution for the forward expectation based on current state variables and iterate forward. Keep adjusting the parameters of the approximating function to minimize the expectation errors.
- Marcet used Second order polynomial expansion
- With John Duffy, in one paper and Guay Lim, in a book and many papers, over many decades, I used neural network approximations and showed that they gave more accurate approximations

# Bayesian Estimation and Counterfactual Simulation

- One of the drawbacks of the DSGE modeling was calibration: how to match these models to real world data?
- One way was to compare the impulse response paths of calibrated DSGE models generating artificial data with VAR models estimated from actual data
- Civil war within the Minnesota Mafia, Chari-Kehoe vs. Chrisitano and Eichenbaum.
- Sargent stepping in to make peace over this with the "ABC's and D's of VARs
- Sims and his students (Del Negro, Pesenti, and Schorfheide in particular) at Yale pioneered the use of Bayesian estimation for these models to tie them down to data.

# Bayes Rule

$$Pr(\Omega|y,x) \propto Pr(y|\Omega,x)Pr(\Omega) \tag{7}$$

- The left-hand term is the Posterior Probability
- The first term on the right hand side is the Likelihood
- The second term is the Prior Probability for the parameter set $\Omega$
- Method is to first optimize the likelihood function for the mode of the parameter set $\Omega$
- The use a Monte Carlo Markov Chain for random draws from the prior distributions
- Then calculate the median, mean and 95 percent confidence intervals for the parameters

# Applications-No Cross of Gold

- With James Lothian and Gabriel Fagan, we estimated a small dynamic model of the USA with data for the Gold Standard, 1870-1914.
- We fitted the model under the gold standard and the did a counterfactual simulation with a Taylor rule in place
- Contrary to the rhetoric William Jennings Bryan, we did not find any "cross of gold".
- Even if we could have put clones of Greenspan and Volker into a time machine and have the FED in place, things would not have been very different.
- Rationale: most of the driving forces of GDP growth and inflation were productivity shocks.

# Fear of Floating in Singapore

- With Hwee Kwan Chow and Guay Lim, we estimated a SOE model for Singapore with an exchange-rate rule, which is the instrument for monetary policy
- We then put in place a counterfactual Taylor rule and let the exchange rate adjust to interest rate differentials.
- We found that things would not be better with a Taylor rule.
- Rationale: Singapore is a highly open economy, with terms of trade more important than productivity forces driving the economy.
- Argument: an exchange-rate rule works better when terms of trade are the driving forces while a Taylor rule is preferable when productivity is the driving force.

# Big Data and Banking Contagion: Using Elastic Net and Cross Validation

- In collaboration with Jennifer Lai, a study of contagion effect of Chinese banks as well as the offshore and onshore RMB markets and external measures of global economic policy uncertainty (EPU) indices developed by Nick Bloom.
- We used daily data for share prices, and obtained range volatility based on the opening,closing, high and low values reported for each data.
- Lots of banks, generous lag structures: too many parameters.
- We use Elastic Net (EN) with Cross Validation (CV).

$$\beta_{Enet} = \overset{Min}{\beta} \sum_{t=1}^{T}(y_t - \sum_{i} \beta_i x_{it})^2 + \lambda \sum_{i=1}^{k}[(\alpha|\beta_i|) + (1-\alpha)\beta_i^2] \quad (8)$$

# More on EN-CV

- The parameter $\alpha$ is a hyperparameter. With $\alpha = 1$ we have LASSO, with $\alpha = 0$ we have Ridge Regression.
- Most use $\alpha = .5$.
- We find the optimal
- The optimal $\lambda^*$ is the one which delivers the lowest out-of-sample mean squared error for a given number of folds.
- With Jennifer we found with EN-CV for Chinese banks, underctainty indices, that the offshore-onshore RMB spreads play a key role for share-price volatility among Chinese banks.

# Policy Positions from 50 Years Comptational Learning

- Money matters. But it is not all that matters (Niehans)
- The FED should be hawkish on inflation (Bigman, Barbosa)
- The effects of rapid trade liberalization take time (Corbo)
- Capital controls have only temporary effectivenss, as the experience of Ireland showed (Browne)
- There was no "cross of gold" (Fagan, Lothian).
- An exchange-rate rule can be more effective as an interest-rate rule(Lim and Chow)
- Having a dual exchange rate system, as China does, puts puressure on the stability of the banking system (Lai)
- Being in the Euro Area mitigates the downsides of financial crises, but increases the frequency (Fagan)
- Marginally raising taxes on capital to finance infrastructure has positive efects on both growth and the labor share (Lim).

# Reflection on Computational Methods

- What would we have learned from Fisher (Ronald and Irving), Kolmogorov, Neyman, Pearson, Zellner if they had these tools?
- Great respect for results of Cagan's econometrics done on hand calculator.
- It is clear we have to evolve into computational polyglots or at least bilingual coders if we want to keep up with developments in computational Macro and Finance, including Machine Learning and AI applications.
- We should also realize that our "toys" are not that interesting in themselves. These methods are old hat at ORSA-TIMS and IEEE meetings

# Agenda for Research

- Our contributions will come from putting these tools to work with new and interesting data sets.
- Bottom line: there is no *free lunch* in empirical work, now and then.
- We have to think long and hard about the issues we investigate, the data we use and the models we apply
- We should look for *robustness* of results across a variety of ML algorithms.
- Even with our computational tools, it is important to spend time reading articles, taking notes with pen or pencil on paper, reading classical writings as well as economic history, talking with colleagues.
- To paraphrase Norman Maclean, "all good things,like catching trout, or for myself, having a consistently good golf swing, and doing significant research, come by grace, and grace comes by art, and art does not come easy".