Reading List for Macroeconomic Analysis with Machine Learning and Big Data

Weinan E and Yucheng Yang
Summer 2019

Last Update: July 20, 2019

# 1 Overview

## 1.1 Introduction

Einav, L. and Levin, J., 2014. Economics in the age of big data. *Science*, 346(6210), p.1243089.

Lecture Notes on Using Big Data to Solve Economic and Social Problems by Raj Chetty and Greg Bruich.

Abraham, K., Jarmin, R., Moyer, B., and Shapiro, M., Forthcoming. Big Data for 21st Century Economic Statistics. University of Chicago Press. `https://www.nber.org/books/abra-7`

Kolanovic, M. and Krishnamachari, R.T., 2017. Big data and AI strategies: Machine learning and alternative data approach to investing. JP Morgan Global Quantitative & Derivatives Strategy Report.

## 1.2 Basics of Machine Learning for Macroeconomics

Lecture Notes on Mathematical Introduction of Machine Learning by Weinan E.

Friedman, J., Hastie, T. and Tibshirani, R., 2001. The elements of statistical learning. New York: Springer series in statistics.

Goodfellow, I., Bengio, Y. and Courville, A., 2016. Deep learning. MIT press.

Shai Shalev-Shwartz and Shai Ben-David. 2014. Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press.

Athey, S. and Imbens, G., 2019. Machine Learning Methods Economists Should Know About. arXiv preprint arXiv:1903.10075.

Athey, S., 2018. The impact of machine learning on economics. In *The Economics of Artificial Intelligence: An Agenda*. University of Chicago Press.

Mullainathan, S. and Spiess, J., 2017. Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2), pp.87-106.

Nakamura, E. and Steinsson, J., 2018. Identification in macroeconomics. *Journal of Economic Perspectives*, 32(3), pp.59-86.

# 2 Statistical Model in Macroeconomics and Machine Learning

## 2.1 Vector Autoregressive Model and Structural VAR

Lecture Notes on Time Series Econometrics by James Stock and Mark Watson.
Webpage: `https://www.aeaweb.org/conference/cont-ed/2019-webcasts`

Kilian, L. and Lütkepohl, H., 2017. Structural vector autoregressive analysis. Cambridge University Press.
Webpage: `https://sites.google.com/site/lkilian2019/textbook/preliminary-chapters`

Del Negro, M. and Schorfheide, F., 2010. Bayesian macroeconometrics.`https://cpb-us-w2.wpmucdn.com/web.sas.upenn.edu/dist/e/242/files/2017/04/bayesian_macro-pjuc2n.pdf`

## 2.2 State Space Model, Filtering Problem and EM Algorithm

Lecture Notes on Time Series Econometrics by James Stock and Mark Watson.
Webpage: `https://www.aeaweb.org/conference/cont-ed/2019-webcasts`

Stock, J.H. and Watson, M.W., 2016. Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. In Handbook of macroeconomics (Vol. 2, pp. 415-525). Elsevier.

Lecture Notes on Machine Learning and Dynamical System by Weinan E.

## 2.3 Recurrent Neural Network and LSTM Network

Karpathy, A., 2015. The unreasonable effectiveness of recurrent neural networks. `http://karpathy.github.io/2015/05/21/rnn-effectiveness/`

Christopher Olah. 2015. Understanding LSTM Networks. `https://colah.github.io/posts/2015-08-Understanding-LST`

Rangapuram, S.S., Seeger, M.W., Gasthaus, J., Stella, L., Wang, Y. and Januschowski, T., 2018. Deep state space models for time series forecasting. NIPS.

Doerr, A., Daniel, C., Schiegg, M., Duy, N.T., Schaal, S., Toussaint, M. and Sebastian, T., 2018, July. Probabilistic Recurrent State-Space Models. ICML.

## 2.4 VAR and State Space Model with Non-standard Data

Chang, M., Chen, X. and Schorfheide, F., 2018. Heterogeneity and Aggregate Fluctuations. Manuscript, University of Pennsylvania.

Liu, L. and Plagborg-Møller, M., 2019. Full-Information Estimation of Heterogeneous Agent Models Using Macro and Micro Data.

Mongey, S. and Williams, J., 2017. Firm dispersion and business cycles: Estimating aggregate shocks using panel data. Manuscript, New York University.

Foroni, C. and Marcellino, M.G., 2013. A survey of econometric methods for mixed-frequency data.

Schorfheide, F. and Song, D., 2015. Real-time forecasting with a mixed-frequency VAR. Journal of Business & Economic Statistics, 33(3), pp.366-380.

Bok, B., Caratelli, D., Giannone, D., Sbordone, A.M. and Tambalotti, A., 2018. Macroeconomic nowcasting and forecasting with big data. Annual Review of Economics, 10, pp.615-643.

# 3 Structural Model in Macroeconomics and Machine Learning

## 3.1 Representative Agent Model and DSGE

Lecture Notes on Graduate Macro by Eric Sims `https://www3.nd.edu/~esims1/grad_macro_17.html`

Galí, J., 2015. Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications. Princeton University Press.

Herbst, E.P. and Schorfheide, F., 2015. Bayesian estimation of DSGE models. Princeton University Press.
Website: `https://web.sas.upenn.edu/schorf/companion-web-site-bayesian-estimation-of-dsge-models/`

## 3.2 Heterogeneous Agent Model: Aiyagari-Bewley-Huggett to Krusell-Smith

Krueger, D., 2012. Macroeconomic theory. Lecture Notes Chapter 10. `https://www.ssc.wisc.edu/~aseshadr/econ714/MacroTheory.pdf`

Krusell, P. and Smith, Jr, A.A., 1998. Income and wealth heterogeneity in the macroeconomy. Journal of political Economy, 106(5), pp.867-896.

Den Haan, W.J., 2010. Comparison of solutions to the incomplete markets model with aggregate uncertainty. Journal of Economic Dynamics and Control, 34(1), pp.4-27.

Winberry, T., 2018. A method for solving and estimating heterogeneous agent macro models. Quantitative Economics, 9(3), pp.1123-1151.

## 3.3 Heterogeneous Agent Model in Continuous Time: HACT and HANK

Lecture Notes on Distributional Macroeconomics by Benjamin Moll.
Webpage: `http://www.princeton.edu/~moll/notes.htm`

Lasry, J.M. and Lions, P.L., 2007. Mean field games. Japanese Journal of Mathematics, 2(1), pp.229-260.

Guéant, O., Lasry, J.M. and Lions, P.L., 2011. Mean field games and applications. In Paris-Princeton lectures on mathematical finance 2010 (pp. 205-266). Springer, Berlin, Heidelberg.

Achdou, Y., Han, J., Lasry, J.M., Lions, P.L. and Moll, B., 2017. Income and wealth distribution in macroeconomics: A continuous-time approach (No. w23732). National Bureau of Economic Research.
Webpage: `http://www.princeton.edu/~moll/HACTproject.htm`

Kaplan, G., Violante, G.L. and Weidner, J., 2014. The Wealthy Hand-to-Mouth. *Brookings Papers on Economic Activity*, 2014(1), pp.77-138.

Kaplan, G., Moll, B. and Violante, G.L., 2018. Monetary policy according to HANK. *American Economic Review*, 108(3), pp.697-743.

Kaplan, G. and Violante, G.L., 2018. Microeconomic heterogeneity and macroeconomic shocks. *Journal of Economic Perspectives*, 32(3), pp.167-94.

Ahn, S., Kaplan, G., Moll, B., Winberry, T. and Wolf, C., 2018. When inequality matters for macro and macro matters for inequality. NBER Macroeconomics Annual, 32(1), pp.1-75.

Auclert, A., 2019. Monetary policy and the redistribution channel. *American Economic Review*, Forthcoming.

Auclert, A., Bardóczy, B., Rognlie, M. and Straub, L., 2019. Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models.
Webpage: `https://github.com/shade-econ/sequence-jacobian/#sequence-space-jacobian`

### 3.4 Solving High-dimensional Stochastic Control and PDEs with Deep Learning

Han, Jiequn and E, Weinan, 2016. Deep learning approximation for stochastic control problems. NIPS Workshop.

Han, Jiequn, Jentzen, A. and E, Weinan, 2018. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences*, 115(34), pp.8505-8510.

E, Weinan, Han, J. and Jentzen, A., 2017. Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *Communications in Mathematics and Statistics*, 5(4), pp.349-380.

Lecture Notes on Reinforcement Learning by David Silver. `http://www0.cs.ucl.ac.uk/staff/d.silver/web/Teaching.html`

### 3.5 Solving Structural Model using Deep Neural Networks

Duarte, V., 2018. Machine Learning for Continuous-Time Economics.

Fernández-Villaverde, J., Hurtado, S. and Nuno, G., 2019. Financial Frictions and the Wealth Distribution. Manuscript, University of Pennsylvania.

## 4 Empirical Macroeconomic Analysis with Big Data

### 4.1 Credit & Consumption Data and the Great Recession

Mian, A., Rao, K. and Sufi, A., 2013. Household balance sheets, consumption, and the economic slump. *The Quarterly Journal of Economics*, 128(4), pp.1687-1726.

Mian, A. and Sufi, A., 2009. The consequences of mortgage credit expansion: Evidence from the US mortgage default crisis. *The Quarterly Journal of Economics*, 124(4), pp.1449-1496.

Mian, A. and Sufi, A., 2015. House of debt: How they (and you) caused the Great Recession, and how we can prevent it from happening again. University of Chicago Press.

Mian, A. and Sufi, A., 2018. Finance and business cycles: the credit-driven household demand channel. *Journal of Economic Perspectives*, 32(3), pp.31-58.

Mian, A.R. and Sufi, A., 2019. Credit Supply and Housing Speculation. `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3209564`

Baker, S.R., 2018. Debt and the response to household income shocks: Validation and application of linked financial account data. *Journal of Political Economy*, 126(4), pp.1504-1557.

Kueng, L., 2018. Excess sensitivity of high-income consumers. *The Quarterly Journal of Economics*, 133(4), pp.1693-1751.

Ganong, Peter, and Pascal Noel. 2019. Consumer Spending during Unemployment: Positive and Normative Implications. *American Economic Review*, 109 (7): 2383-2424.

## 4.2 Tax Data, Inequality and Economic Opportunity

Chetty, R., Hendren, N., Kline, P. and Saez, E., 2014. Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics*, 129(4), pp.1553-1623.
Data Page: `https://opportunityinsights.org/paper/land-of-opportunity/`

Chetty, R. and Hendren, N., 2018. The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects. *The Quarterly Journal of Economics*, 133(3), pp.1107-1162.
Data Page: `https://opportunityinsights.org/paper/neighborhoodsi/`

Chetty, R. and Hendren, N., 2018. The impacts of neighborhoods on intergenerational mobility II: County-level estimates. *The Quarterly Journal of Economics*, 133(3), pp.1163-1228.
Data Page: `https://opportunityinsights.org/paper/neighborhoodsii/`

Chetty, R., Friedman, J.N., Hendren, N., Jones, M.R. and Porter, S.R., 2018. The opportunity atlas: Mapping the childhood roots of social mobility.
Data Page: `https://opportunityinsights.org/paper/the-opportunity-atlas/`

Saez, E. and Zucman, G., 2016. Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *The Quarterly Journal of Economics*, 131(2), pp.519-578.
Data Page: `http://gabriel-zucman.eu/uswealth/`

Piketty, T., Yang, L. and Zucman, G., 2019. Capital Accumulation, Private Property and Rising Inequality in China, 1978-2015. *American Economic Review*, 109 (7): 2469-96.
Data Page: `http://gabriel-zucman.eu/china/`

Fagereng, A., Guiso, L., Malacrino, D. and Pistaferri, L., 2019. Heterogeneity and persistence in returns to wealth.

Fagereng, A., Holm, M.B. and Natvik, G.J.J., 2019. MPC heterogeneity and household balance sheets.

Fagereng, A., Holm, M.B., Moll, B. and Natvik, G.J.J., 2019. Saving Behavior Across the Wealth Distribution: The Importance of Capital Gains.

## 4.3 Scanner Data, Prices and Monetary Policy


Nakamura, E. and Steinsson, J., 2008. Five facts about prices: A reevaluation of menu cost models. *The Quarterly Journal of Economics*, 123(4), pp.1415-1464.

Stroebel, J. and Vavra, J., 2019. House prices, local demand, and retail prices. *Journal of Political Economy*, Forthcoming.

Wong, Arlene, 2019. Refinancing and the Transmission of Monetary Policy to Consumption. Manuscript, Princeton University.

Cavallo, A., 2018. Scraped data and sticky prices. Review of Economics and Statistics, 100(1), pp.105-119.

Hua, Amy, 2017. Learning the Structure of Price Stickiness in Scanner Data. Manuscript, Princeton University.

DellaVigna, S. and Gentzkow, M., 2017. Uniform pricing in US retail chains. *The Quarterly Journal of Economics*, Forthcoming.

Kaplan, G. and Schulhofer-Wohl, S., 2017. Inflation at the household level. *Journal of Monetary Economics*, 91, pp.19-38.


## 4.4 Social Network Data and Macroeconomic Implications


Bailey, M., Cao, R., Kuchler, T., Stroebel, J. and Wong, A., 2018. Social connectedness: Measurement, determinants, and effects. *Journal of Economic Perspectives*, 32(3), pp.259-80.
Visualization: `https://www.nytimes.com/interactive/2018/09/19/upshot/facebook-county-friendships.html`

Bailey, M., Cao, R., Kuchler, T. and Stroebel, J., 2018. The economic effects of social networks: Evidence from the housing market. *Journal of Political Economy*, 126(6), pp.2224-2276.

Bailey, M., Dávila, E., Kuchler, T. and Stroebel, J., 2019. House price beliefs and mortgage leverage choice. *The Review of Economic Studies*, Forthcoming.


A website for Twitter data research: `https://blogs.lse.ac.uk/impactofsocialsciences/2019/06/18/using-twitter-as-a-data-source-an-overview-of-social-media-research-tools-2019/`

## 4.5 Firm Data and Macroeconomic Implications


Dai, Ruochen, Mookherjee, D., Munshi, K. and Zhang, Xiaobo, 2019. The Community Origins of Private Enterprise in China.

Shi, Yu, Townsend, R. and Zhu, Wu, 2019. Internal Capital Markets in Business Groups and the Propagation of Credit Supply Shocks.

Cong, L. W., Gao, H., Ponticelli, J. and Yang, X., 2019. Credit allocation under economic stimulus: Evidence from China. *The Review of Financial Studies*, Forthcoming.

Autor, D.H., Dorn, D. and Hanson, G.H., 2013. The China syndrome: Local labor market effects of import competition in the United States. *American Economic Review*, 103(6), pp.2121-68.

Resource Page: http://chinashock.info/

Autor, D.H., Dorn, D. and Hanson, G.H., 2016. The China shock: Learning from labor-market adjustment to large changes in trade. *Annual Review of Economics*, 8, pp.205-240.

Bernard, A.B., Dhyne, E., Magerman, G., Manova, K. and Moxnes, A., 2019. The origins of firm heterogeneity: a production network approach.

### 4.6 Employer-Employee Data, Job Posting Data and Firm Dynamics

Song, J., Price, D.J., Guvenen, F., Bloom, N. and Von Wachter, T., 2018. Firming up inequality. *The Quarterly Journal of Economics*, 134(1), pp.1-50.

Card, D., Heining, J. and Kline, P., 2013. Workplace heterogeneity and the rise of West German wage inequality. *The Quarterly Journal of Economics*, 128(3), pp.967-1015.

Davis, S.J., Faberman, R.J. and Haltiwanger, J.C., 2013. The establishment-level behavior of vacancies and hiring. *The Quarterly Journal of Economics*, 128(2), pp.581-622.

Webb, Michael, 2019. The Impact of Artificial Intelligence on the Labor Market. http://papers.nber.org/conf_papers/f123534.pdf

Liu, Yukun, 2019. Labor-based Asset Pricing.

### 4.7 Alternative Data and New Measures of Economic Indicators

Jean, N., Burke, M., Xie, M., et al, 2016. Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301), 790-794.
Resource Page: http://sustain.stanford.edu/predicting-poverty

Blumenstock, J., Cadamuro, G., & On, R. 2015. Predicting poverty and wealth from mobile phone metadata. *Science*, 350(6264), 1073-1076.

Kreindler, G.E., Miyauchi, Y., 2019. Measuring Commuting and Economic Activity inside Cities with Cell Phone Records.

Lei Dong, Carlo Ratti, and Siqi Zheng, 2019. Predicting neighborhoods' socioeconomic attributes using restaurant data. *Proceedings of the National Academy of Sciences*, Forthcoming.

Glaeser, E.L., Kim, H. and Luca, M., 2017. Nowcasting the local economy: Using Yelp data to measure economic activity (No. w24010). National Bureau of Economic Research.

Yin, M., Sheehan, M., Feygin, S., Paiement, J.F. and Pozdnoukhov, A., 2017. A generative model of urban activities from cellular data. IEEE Transactions on Intelligent Transportation Systems, 19(6), pp.1682-1696.

Gebru, T., Krause, J., Wang, Y., Chen, D., Deng, J., Aiden, E.L. and Fei-Fei, L., 2017. Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States. *Proceedings of the National Academy of Sciences*, 114(50), pp.13108-13113.

Chi, G., Liu, Y., Wu, Z. and Wu, H., 2015. Ghost cities analysis based on positioning data in China. arXiv preprint arXiv:1510.08505.

Clark, H., Pinkovskiy, M. and Sala-i-Martin, X., 2017. China's GDP growth may be understated (No. w23323). National Bureau of Economic Research.

### 4.8 Textual Data, Uncertainty and Sentiments

Gentzkow, M., Kelly, B.T. and Taddy, M., Text as Data. *Journal of Economic Literature*, Forthcoming.

Baker, S.R., Bloom, N. and Davis, S.J., 2016. Measuring economic policy uncertainty. *The quarterly journal of economics*, 131(4), pp.1593-1636.

Soo, C.K., 2018. Quantifying sentiment with news media across local housing markets. *The Review of Financial Studies*, 31(10), pp.3689-3719.

Atalay, E., Phongthiengtham, P., Sotelo, S. and Tannenbaum, D., 2019. The Evolution of Work in the United States. *American Economic Journal: Applied Economics*, Forthcoming.
Data Page: `https://occupationdata.github.io/`

Jegadeesh, N. and Wu, D.A., 2017. Deciphering fedspeak: The information content of fomc meetings.

Turrell, A., Speigner, B., Djumalieva, J., Copple, D. and Thurgood, J., 2018. Using job vacancies to understand the effects of labour market mismatch on UK output and productivity.

Hassan, T.A., Hollander, S., van Lent, L. and Tahoun, A., 2019. Firm-level political risk: Measurement and effects. *Quarterly Journal of Economics*, Forthcoming.

Kelly, B., Papanikolaou, D., Seru, A. and Taddy, M., 2018. Measuring technological innovation over the long run (No. w25266). National Bureau of Economic Research.

Cong, L.W., Liang, T. and Zhang, X., 2019. Textual Factors: A Scalable, Interpretable, and Data-driven Approach to Analyzing Unstructured Information.