Economic Modelling 49 (2015) 395–406

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.elsevier.com/locate/ecmod

Red herrings and revelations: does learning about a new variable worsen forecasts?

Paul Shea [1]

Bates College, Lewiston, ME 04240, United States

ARTICLE INFO

Article history:
Accepted 27 May 2015
Available online 18 June 2015

Keywords:
Adaptive Learning
Underparameterization
Bounded rationality
Forecasting

ABSTRACT

I develop a framework where agents forecast despite knowing only a subset of the variables in the true economic model. I then examine whether the discovery of an additional variable improves forecasting. Because agents do not know all of the variables in the model, they form expectations using bounded rationality. Under *adaptive learning*, agents form expectations by regressing a variable of interest on the revealed variables. Surprisingly, I find that the revelation of an additional variable often worsens forecasts, an event deemed a *red herring*, with probability greater than one-half. If the model includes endogenous variables that depend on agents' expectations, then revealing a new variable will occasionally lead to a catastrophic worsening of forecast accuracy. Under *structural coefficients expectations*, agents know how each revealed variable appears in the true model and they use this information to forecast. Now, the revelation of a new variable improves forecasting more often than not. I then apply the framework to a calibrated New Keynesian model and find that the revelation of a new variable usually worsens forecasting. Collectively, these results show that learning about a new variable may actually make forecasts less accurate.

1. Introduction

Hendry and Clements (2003) state that "all econometric models are mis-specified."[2] Indeed, most econometricians would agree that any interesting econometric estimate surely omits some variables, and that the goal of econometricians should be to minimize this bias or convincingly argue that it works in their favor. Likewise, even good macroeconomic theory models, by design, omit some major aspects of the economy. This paper does not argue that models should be made larger to reduce this misspecification. Instead, it examines the effect of this type of misspecification in a theoretical self-referential model where agents form forecasts despite only knowing some of the variables that appear in the true model. It finds that learning about an additional variable may actually worsen forecasting.

The baseline approach in macroeconomic theory is to examine a model after it has converged to its rational expectations equilibrium (REE). Rational expectations assume that agents know the model's reduced form solution, and use this solution to form optimal forecasts. An alternative approach is to instead assume bounded rationality where agents must forecast in the presence of some informational limitation. An example of bounded rationality that has attracted considerable interest recently is the *restricted perceptions equilibrium* (RPE), where agents forecast based on only a subset of the variables that appear in the model's solution.[3] A RPE is a good way to model the omissions that exist in actual empirical and theoretical macroeconomic work.

This paper addresses a novel question related to restricted perceptions equilibria. If economic theory reveals a new variable, does this new information make forecasts better or worse? Given enough data, we might expect the answer to be yes if agents are simply forecasting an exogenous process. In this paper, however, agents choose the model's endogenous variable based on their expectation of its future value. Because the model is self-referential in this way, the revelation of a new variable changes the data generating process so that forecasts based on the old data generating process may be biased. As a result, revelations often cause worse forecasts, in some cases with probability greater than one-half. I refer to cases where the revelation of a new variable worsens forecasts as a *red herring*.[4] Furthermore, this bias occasionally causes the model to move close to a singularity where forecast errors become exceptionally large.

E-mail address: pshea@bates.edu.

[1] I thank the editor and two anonymous referees for helpful suggestions.

[2] This sentiment is also the theme of several econometric textbooks including Hendry and Clements (1998), and White (1994).

[3] For a more detailed discussion of restricted perceptions equilibria, see Branch and Evans (2006).

[4] A red herring is a metaphor used to describe an object that distracts an investigation, diverts attention to a side issue, or provides useless but confusing information. Its origins date to seventeenth century England where a herring, reddened by salting and smoking, was used to confuse hounds pursuing a fox or other prey. See Quinion (2002) for more details.

http://dx.doi.org/10.1016/j.econmod.2015.05.012
0264-9993/

Most macroeconomic work analyzes models after convergence to rational expectations has occurred. Even the bounded rationality literature, including work on restricted perceptions equilibria, typically limits its analysis to whether convergence (possibly to rational expectations) occurs, or to the model's dynamics after convergence.[5] The process of convergence is thus implicitly assumed to be either inheriently uninteresting, or to occur fast enough so that it may safely be ignored.

This paper instead focuses on the process of convergence after the revelation of a new variable. By showing that agents' discounted sum of squared forecasts errors are often larger after a revelation, it shows that focusing only on a model post-convergence may obscure important results. Although the model does converge to rational expectations as all variables are revealed, and forecast errors go to zero as this convergence occurs, forecast errors do not steadily decline. Rather, as new variables are revealed, forecast errors often initially increase. Sufficiently clever agents might even stop relying on newly revealed variables altogether, preventing convergence from ever occurring.

Macroeconomic theory may be viewed as a process of "revealing" relevant variables to forecasters. Consider a few examples, accepting their contributions for the sake of argument. Keynes's General Theory (1936) may be interpreted as revealing the importance of nominal rigidities. Most models of business cycles now include some type of productivity shock. The work of Kydland and Prescott (1982), which demonstrated that exogenous variation in the Solow Residual could help explain aggregate fluctuations, may thus be interpreted as theory revealing the importance of productivity shocks to agents in the true model. Likewise, the animal spirits hypothesis predicts that agents' self-fulfilling beliefs may have important aggregate effects. The work of Farmer and Guo (1994) may thus be interpreted as revealing the importance of indeterminacy to forecasters. Although the set of major and genuine theoretical contributions is debatable, many applied macroeconomic theorists hope that their work ultimately enlightens agents attempting to forecast aggregate variables. This paper models this process and shows that even genuine theoretical advances may yield adverse outcomes more often than not.

This paper examines the effect of theoretical revelations, like those from the preceding paragraph, in a self-referential model where agents must forecast despite only knowing a subset of the variables that appear in the true model. In this environment, expectations are formed in two, boundedly rational, ways. Under *structural coefficients expectations*, economic theory reveals not only a subset of relevant variables, but also the associated structural coefficients in the underlying theoretical model. Agents then form expectations using these coefficients and without relying on past data. Expectations are not fully rational, however, because agents do not know any of the structural coefficients associated with the unrevealed variables.

The second type of expectations formation is *adaptive learning* where agents use all revealed variables as regressors to estimate the variable of interest. Adaptive learning typically relaxes rational expectations' strong informational requirements by assuming that agents form expectations using standard econometric techniques. The adaptive learning hypothesis is supported by the observation that professional forecasters almost uniformly rely on empirical estimation, are uncertain of the underlying theoretical model, and rarely issue predictions without the aid of extensive data.

I develop a simple linear model where the variable that agents forecast depends on large sets of exogenous and endogenous variables. The vector of endogenous variables depends on both the set of exogenous variables and agents' expectations. At any time, economic theory has revealed only subsets of both the exogenous and endogenous variables. I then address two questions related to the revelation of a new variable. One, does the revelation of an additional variable improve agents' forecasts? Under structural coefficients expectations, the revelation of an additional variable improves forecasting with probability between one-half and one. If agents seek to minimize their discounted squared forecast errors, they are therefore generally better off using newly revealed variables to forecast. There exists, however, a significant (though less than one-half) probability that a revelation may be a red herring which worsens forecasts.

Under adaptive learning, there are two sources of red herrings. Although an additional variable necessarily improves forecasting at the fixed point of the learning process, it also introduces additional noise into the system.[6] Estimation after the revelation is less parsimonious and may therefore yield a larger mean squared error out-of-sample. This source frequently causes red herrings, but the accompanying worsening of forecasts is relatively small. The second source of red herrings is the Lucas Critique (1976). The revelation of an additional variable changes both the way that agents form expectations and the data generating process. If agents retroactively regress the variable of interest on an expanded set of regressors, then the resulting estimate will be biased. The Lucas Critique results in a small but positive probability that this bias will move the model arbitrarily close to a singularity where the endogenous variables and the mean squared forecast error explode. This type of red herring may therefore result in a catastrophic worsening of forecasts.

Under adaptive learning, a newly revealed endogenous variable is a red herring with a significant, but usually less than one-half, probability. The possibility of catastrophic red herrings, however, causes the accompanying average welfare loss to be positive in four of the eleven calibrations, including two with heightened levels of endogeneity. A newly revealed exogenous variable is always a red herring with probability usually greater than one-half. Catastrophic red herrings may also occur upon the revelation of an exogenous variable, and, on average, forecasts worsen for all eleven calibrations. Collectively, these results suggest that genuine theoretical breakthroughs often destabilize the economy.

Structural coefficients expectations assume that theory reveals not only the existence of a new variable, but also exactly how this new variable appears in the true model. This type of expectations formation is thus vulnerable to the same critique that has been controversially levied against rational expectations; that it endows agents with implausibly high amounts of knowledge about the true model. In contrast, adaptive learning assumes that agents only learn about the existence of a new variable and must estimate how it affects the model through econometric methods. Agents are thus unsure both about which variables appear in the model and the correct coefficients for the known variables whereas under structural coefficients expectations agenst are only unaware of the former. Thus, a reader might conclude that only adaptive learning is feasible and that a comparison between the two is of minimal concern.

Because the critique that rational expectations endows agents with too much knowledge is not entirely accepted, the other question that this paper explores is whether agents are better off choosing to use structural coefficients expectations or adaptive learning, assuming that both are feasible. Because structural coefficients expectations fail to exploit the correlations between the revealed and unrevealed variables, they deliver conditionally biased forecasts. Adaptive learning is initially biased, but converges towards delivering unbiased forecasts after the new variable is revealed.[7] Adaptive learning's econometric algorithm also adds additional noise into the system. The results show that if the

[5] Under constant gain learning, which this paper employs, the learning algorithm often converges to a distribution around rational expectations rather than rational expectations themselves.

[6] Here, the fixed point of the learning process refers to value of the regression coefficients as the sample size under the new data generating process goes to infinity.

[7] The coefficient estimates, however, suffer from omitted variable bias. In the model, agents only care about the squared error of their forecasts and not their parameter estimates.