Red Herrings and Revelations: Does Learning About a New Variable Worsen Forecasts?

Paul Shea*

Bates College

April 28, 2015

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.

JEL Classification: E27, E61, D81, D83. N41.

Keywords: Adaptive Learning, Underparameterization, Bounded Rationality, Forecasting.

*pshea@bates.edu. I thank the editor and two anonymous referees for helpful suggestions.

1 Introduction

Hendry and Clements (2003) state that "all econometric models are mis-specified."[1] 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.[2] 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*

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

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

herring.[3] Furthermore, this bias occasionally causes the model to move close to a singularity where forecast errors become exceptionally large.

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 analyis to whether convergence (possibly to rational expectations) occurs, or to the model's dynamics after convergence.[4] The process of convergence is thus immplicitly 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

[3]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.

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

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.[5] Estimation after the revelation is less parsimonious and may therefore

[5]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.

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.[6] Adaptive learning's econometric algorithm also adds additional noise into the system. The results show that if the exogenous and endogenous variables are sufficiently correlated, then agents are better off using adaptive learning in order to exploit the correlation between known and unknown variables. If this correlation is weak, however, then structural coefficients expectations outperform adaptive learning.

How agents form expectations is among the most controversial questions in current macroeconomics. This paper contributes to the literature that argues that rational expectations endow agents with too much information. The most popular alternative to rational expectations is adaptive learning which typically assumes that agents run a properly specified regression, free from omitted variable bias. To properly specify their econometric models, however, agents must have considerable knowledge about the true model. Like rational expectations, standard adaptive learning is therefore vulnerable to criticism for endowing its agents with excessive information. Several other papers examine this issue by modeling adaptive learning where agents use underparameterized models and then examining the resulting RPE. Cho and Kasa (2015) examine Sargent's "Conquest" (1999) model where agents use adaptive learning to estimate the endogenous data generating process. In addition to uncertainty about the coefficients, agents are also unsure of the correct specification. Agents therefore perform tests among a set of underparameterized models. Branch and Evans (2006) examine a cobweb model where agents choose among a set of underparameterized specifications. The authors demonstrate that equilibria may occur where agents heterogeneously rely on different specifications. Evans and Ramey (2006) assume that agents use underparameterized, adaptive expectations in a model of the New Keynesian Phillips Curve. They allow agents to select the optimal weights on the previous observation versus the sample mean and show that the Lucas Critique (1976) still applies to much of the parameter space. Unlike the present paper, each of these related papers assumes that agents choose among a set of underparameterized models with constant sets of regressors. This paper is the first to examine how discovering a new variable affects the model's behavior.

Most of the analysis in this paper uses a general model. Section 3, however, applies these results to the New Keynesian framework, by far the most popular modelling approach for analyzing mon-

[6]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.

etary policy.[7] Macroeconomists often suggest that policy makers created excessive inflation in U.S. during the 1960's and 1970's by mistakenly believing in a stable long-run tradeoff between inflation and output. This story may be interpreted as an example of a red herring. If we accept the current state of monetary theory, then the development of the Phillips Curve represents a genuine theoretical breakthrough; the New Keynesian framework posits a similar, though short-run, relationship. Due to endogeneity and insufficient understanding of other important variables, such as expectations, however, policy makers misunderstood the true relationship between inflation and output. As a result, the discovery of the Phillips Curve may have managed to de-stabilize the economy. In this example, all agents use adaptive learning to forecast inflation and output, but their learning algorithms omit a vector of unrevealed variables. Upon the revelation of an additional variable, the results of the general model carry through. For the calibrated New Keynesian Model, the revelation of a new variable results in a red herring between 69% and 85% of the time, and results in a catastrophic worsening of forecasts between 0.7% and 3.8% of the time.

The paper is organized as follows. Section 2 develops the model and examines the revelation of new variables under structural coefficients expectations and adaptive learning. Section 3 presents an application to a calibrated New Keynesian model. Section 4 concludes.

2 The Model and Structural Coefficients Expectations

This paper relies on a version of Sargent's (1999) "Conquest Model:"

$$U_t = U^* - \theta(y_t - x_t) \tag{1}$$

The policy maker chooses a policy, x_t, that determines the unemployment rate, U_t. The policy maker is unable to observe the partially exogenous variable, y_t, which depends on a large number of observable variables. The policy maker's objective is simply to stabilize unemployment around U^* each period. This entails setting $x_t = y_t^e$. The policy maker's problem is therefore equivalent to forming the best forecast of y_t.

[7] For details on the New Keynesian model, see Woodford (2003). Forecasters, however, typically rely on larger more complex versions.

$$y_t = Az_t + Bg_t \tag{2}$$

$$z_t = Cy_t^e + Dg_t + u_t \tag{3}$$

y_t is a scalar. g_t is a *M x 1* vector of exogenous variables that affect y_t. z_t is a *N x 1* vector of endogenous variables that also affect y_t. z_t itself depends on g_t and u_t, a *N x 1* vector of unobservable shocks. The inclusion of u_t ensures that the endogenous variables have independent explanatory power and are not simply linear combinations of the exogenous variables. Additionally, z_t depends on the policy, y_t^e. For example, it is natural to interpret y_t as inflation. In this example, z_t may include consumption and labor supply, both of which themselves likely depend on the policy instrument. I assume that each element of g_t and u_t is independently and identically distributed, $N(0, \sigma_g^2)$ and $N(0, \sigma_u^2)$.

The policy maker does not generally know all of the variables that affect y_t. At time *t*, economic theory has revealed the first *n* elements of z_t (denoted by z_t^n), and the first *m* elements of g_t (denoted by g_t^m). Note, however, that y_t depends on the entire vectors z_t and g_t, and that z_t depends on the entire vector g_t, instead of only on the revealed components. As in most macroeconomic models, I assume that there exists a continuum of measure one of micro-level agents, whose behavior then determines z_t and y_t. The unrevealed elements of z_t and g_t are therefore the result of macroeconomic theorists' inability to understand all elements of the micro-level optimization process. This limitation is then passed onto the policy maker who must forecast despite not knowing the full set of relevant variables.

It is also possible that micro-level agents' optimization depends on their forecast of the aggregate variable y_t. Throughout the paper, I assume that these agents experience the same limitations in forecasting aggregate variables as the policy maker, and that all expectations, y_t^e, are therefore identical.[8]

Finally, I assume that the instantaneous loss function equals the squared forecast error: $\ell_t = (y_t - y_t^e)^2$ and for the remainder of the paper I evaluate welfare by calculating the sum agents' discounted squared forecast errors.[9] The timing of the model works as follows: first, g_t is drawn. Second, z_t and

[8]A more general specification allows z_t to depend on both y_t, and y_t^e. The paper's major conclusions do not depend on this distinction.

[9]Granger (1969) and Granger and Newbold (1986) show that minimizing a function of the mean squared forecast error maximizes many welfare functions. Patton and Timmermann (2007), however, show that with asymmetric losses,

y_t^e are simultaneously determined. Finally, y_t is determined based on g_t and z_t.

2.1 Structural Coefficients Expectations

In this subsection, I assume that economic theory has also revealed A^n (the first *n* elements of A), and B^m (the first *m* elements of B) and that agents use these structural parameters to form their expectations. I refer to this type of expectations formation as *structural coefficients expectations*:

$$y_t^e = A^n z_t^n + B^m g_t^m \tag{4}$$

$$z_t^n = C^n y_t^e + D^n g_t + u_t^n \tag{5}$$

If agents mistakenly believe that they know all of the variables that appear in the model, then what they believe are rational expectations are, in fact, structural coefficients expectations. Structural coefficients expectations thus have three appealing properties reminiscent of rational expectations. First, both the unconditional expectational error, $E[y_t - y_t^e]$, and the expectational error conditional on the available information, $E[(y_t - y_t^e)|A^n, B^m, g_t^m, z_t^n]$, equal zero. Second, the estimate is precise; there is no estimation error associated with A^n, and B^m. Finally, because they are based on invariant structural parameters, structural coefficients expectations are not vulnerable to the Lucas Critique when economic theory reveals a new variable.

Because agents do not generally know the full model, structural coefficients expectations are not fully rational. They are therefore not necessarily optimal and they do entail a significant disadvantage. The revealed variables, z_t^n and g_t^m are typically correlated with the unrevealed variables, z_t^{-n} and g_t^{-m}. Regressing y on the revealed variables, as agents do under adaptive learning in Subsection 2.2, exploits this correlation. Structural coefficients expectations, however, ignore this correlation and therefore yield a worse in-sample fit than adaptive learning. The results of Subsection 2.2 show that if the correlation between the revealed and unrevealed variables is sufficiently strong, then structural

minimizing the mean squared error may not be optimal. I thus re-calculate the results of this section assuming either that welfare depends on the absolute value of the forecast error or that negative forecast errors are weighted between 0.1 and 10 times as much as positive forecast errors. The main results of this section are unaffected.

coefficients expectations perform worse than adaptive learning. If this correlation is weak, however, then the precision of structural coefficients expectations make them preferable.

The feasibility of structural coefficients expectations is questionable. Like rational expectations, structural coefficients expectations assume that agents possess significant knowledge of the true model. The motivation for the adaptive learning literature therefore suggests that agents are unable to form structural coefficients expectations. This section, however, assumes that structural coefficients expectations are feasible, and the remainder of the paper leaves this question open.

Substituting Equation (5) into Equation (4) yields agents' expectation of y_t.

$$y_t^e = (1 - A^n C^n)^{-1}[A^n D^n g_t + B^m g_t^m + A^n u_t^n] \tag{6}$$

Agents' loss may then be re-stated in terms of their knowledge, (*n* and *m*), and the vectors of exogenous shocks, g_t and u_t.

$$\ell_t = \{((AC - 1)(1 - A^n C^n)^{-1} A^n D^n + AD + B)g_t + Au_t ...$$

$$+[(AC - 1)(1 - A^n C^n)^{-1}](B^m g_t^m + A^n u_t^n)\}^2 \tag{7}$$

Proposition 1: If agents are fully rational ($m = M$ and $n = N$), then agents' loss, ℓ_t, equals zero.
Proof: If ($m = M$ and $n = N$), then $A^n = A$, $B^m = B$, $C^n = C$, $g_t^m = g_t$, and $u_n^t = u_t$. Equation (7) then reduces to $\ell_t = 0$.

Proposition 1 demonstrates that departures from full rationality cause a positive loss. Structural coefficients expectations and adaptive learning are two such departures. I now examine how the former affects welfare.

Suppose, however, that agents use adaptive learning where y_t is regressed on the set of revealed variables. If theory reveals a new variable, then both the way that agents form expectations and the data generating process for y_t will also change. If agents retroactively regress y_t on the expanded set of variables, then the Lucas Critique applies, and the retroactive regression coefficients will be biased relative to their optimal values for the new data generating process.[10] If C equals a null matrix, however, then the data generating process does not change and the Lucas Critique does not apply.

[10]This bias is in addition to the bias that results from omitted variables.

Under structural coefficients expectations, agents rely on invariant structural parameters to form expectations and the Lucas Critique does not apply regardless of the value of C. To simplify the analysis for the remainder of this section, I therefore set C equal to a null matrix.[11] I then re-write agents' loss under structural coefficients expectations:

$$\ell_t = [(A^{-n}D^{-n,m})g_t^m + [(A^{-n}D^{-n,-m} + B^{-m})g_t^{-m} + A^{-n}u_t^{-n}]^2 \tag{8}$$

where $D^{k,j}$ is a matrix consisting of the first *k* rows and first *j* columns of D, and $D^{k,-j}$ is matrix consisting of the first *k* rows and all but the first *j* columns of D.

Consider the $m+1^{th}$ exogenous variable, $g_{t,m+1}$. This variable contributes to agents' loss directly through the $m+1^{th}$ element of B, $b^{1,m+1}$, and indirectly though its effects on the unrevealed endogenous variables, captured by the $m+1^{th}$ element of $A^{-n}D^{-n,M}$. Under structural coefficients expectations, if economic theory reveals $g_{t,m+1}$, it also reveals its direct effect through $b^{1,m+1}$. It is likely that the direct effect of $g_{t,m+1}$ is informative about its total effect. In this case, the revelation will improve forecasting. With probability less than one-half, however, the direct effect may not be representative of the total effect, agents' forecasts will worsen, and the newly revealed variable is a red herring. Similarly, the revelation of the $n+1^{th}$ endogenous variable will likely increase welfare but will also be a red herring with probability less than one-half.

Tedious but straightforward manipulation of Equation (8) allows agents' expected loss for any $[A, B, D]$ to be re-stated as a function of the individual elements of the relevant matrices.

$$E[\ell_t] = \sigma_u^2 \sum_{i=n+1}^{N} (a^{1,i})^2 + \sum_{j=1}^{m} \sigma_g^2 [\sum_{i=n+1}^{N} \sum_{k=n+1}^{N} a^{1,i}d^{i,j}a^{1,k}d^{k,j}] + ...$$

$$\sum_{j=m+1}^{M} \sigma_g^2 [\sum_{i=n+1}^{N} \sum_{k=n+1}^{N} a^{1,i}d^{i,j}a^{1,k}d^{k,j} + 2\sum_{i=n+1}^{N} a^{1,i}d^{i,j}b^{1,j} + (b^{1,j})^2] \tag{9}$$

where $d^{k,j}$ represents the individual element in the k^{th} row and j^{th} column of D. (This notation is in contrast to $D^{k,j}$, which represents the first *k* rows and *j* columns of D.)

I now consider the effect if macroeconomic theorists reveal an additional variable to agents in the model. I treat this revelation as exogenous event that is immediately accepted by all forecasters.

[11] The major results that follow in this section are unchanged for cases where C does not equal a null matrix. These results are available upon request.

Suppose that economic theory reveals the $m+1^{th}$ exogenous variable. Equation (9) demonstrates that the change in the expected ex-post loss equals:

$$E[\Delta\ell_t|A,B,D]^{m+1,n} = -\sigma_g^2\{(b^{1,m+1})^2 + 2b^{1,m+1}\sum_{i=n+1}^{N} a^{1,i}d^{i,m+1}\} \tag{10}$$

Likewise, if economic theory reveals the $n+1^{th}$ endogenous variable, the change in the expected ex-post loss equals:

$$E[\Delta\ell_t|A,B,D]^{m,n+1} = -\sigma_g^2\{\sum_{j=1}^{M}[(a^{1,n+1}d^{n+1,j})^2 + 2a^{1,n+1}d^{n+1,j}\sum_{k=n+2}^{N} a^{1,k}d^{k,j}]...$$

$$+2\sum_{j=m+1}^{M} a^{1,n+1}d^{n+1,j}b^{1,j}\} - \sigma_u^2(a^{1,n+1})^2 \tag{11}$$

I assume that agents know the distributions of each element of $[A,B,D]$. Each element of $[A,B,D]$ is independently and identically distributed, $N(0,\sigma_q^2)$, where $q = a,b,d$. I further assume orthogonality across matrices by using the following variance-covariance matrix:

$$\begin{bmatrix} \sigma_a^2 & 0 & 0 \\ 0 & \sigma_b^2 & 0 \\ 0 & 0 & \sigma_d^2 \end{bmatrix} \tag{12}$$

Proposition 2: Ex-ante, the revelation of either the $m+1^{th}$ exogenous variable or the $n+1^{th}$ endogenous variable results in a decreased expected loss.

Proof: Under the assumption of orthogonality from Equation (12), Equation (10) reduces to $E[\Delta\ell_t]^{m+1,n} = -\sigma_g^2\sigma_b^2 < 0$, and Equation (11) reduces to $E[\Delta\ell_t]^{m,n+1} = -\sigma_g^2\sigma_a^2\sigma_d^2 - \sigma_u^2\sigma_a^2 < 0$.

Proposition 2 shows that, ex-ante, agents are always better off relying on a newly discovered variable because it is likely to result in a welfare improvement. Equations (10) and (11) demonstrate, however, that for specific draws of $[A,B,D]$, newly revealed exogenous or endogenous variables may increase in an ex-post expected loss.

Proposition 3: The probability that the $m+1^{th}$ exogenous variable is a red herring approaches zero as $\sigma_b^2 \to \infty$, $\sigma_a^2 \to 0$, or $\sigma_d^2 \to 0$. Likewise if $m+1 < M$, then this probability approaches 1/2 as $\sigma_b^2 \to 0$, $\sigma_a^2 \to \infty$, or $\sigma_d^2 \to \infty$.

Proof: By Equation (10), as $\sigma_b^2 \to \infty$, $E[\Delta\ell_t|A, B, D]^{m+1,n} \to -\infty$. Likewise, as $\sigma_a^2 \to 0$ or $\sigma_d^2 \to 0$, $E[\Delta\ell_t|A, B, D]^{m+1,n} \to -\sigma_b^2 < 0$. Also as $\sigma_b^2 \to 0$, $\sigma_a^2 \to \infty$, or $\sigma_d^2 \to \infty$, then $E[\Delta\ell_t|A, B, D]^{m+1,n} \to 0$, and because of the model's symmetry, $prob(E[\Delta\ell_t|A, B, D]^{m+1,n} > 0) \to 1/2$.

If each element of B has a large variance, then the direct effect of the $m + 1^{th}$ exogenous variable dominates its indirect effect. The revelation of the direct effect, which accompanies the revelation of the $m + 1^{th}$ exogenous variable, is therefore necessarily informative about its total effect. Likewise, if A or D equal a null matrix, then there is no indirect effect. If $\sigma_b^2 \to 0$, however, then the revelation of a new element of B is uninformative and does not affect forecasting. As $\sigma_b^2 \to 0$, it is also the case that the welfare loss associated with a red herring approaches zero. If $\sigma_a^2 \to \infty$ or $\sigma_d^2 \to \infty$, then the indirect effect of the revelation of $g_{t,m+1}$ dominates its direct effect. In these cases, the welfare loss associated with a red herring approaches infinity.

Proposition 4: The probability of the $n + 1^{th}$ endogenous variable being a red herring approaches zero as $\sigma_a^2 \to \infty$, as $\sigma_d^2 \to \infty$, or as $\sigma_u^2 \to \infty$. Likewise if $n+1 < N$, then this probability approaches one-half as $\sigma_b^2 \to \infty$.

Proof: By Equation (11), as $\sigma_a^2 \to \infty$, $\sigma_d^2 \to \infty$, or $\sigma_u^2 \to \infty$, $E[\Delta\ell_t|A, B, D]^{m,n+1} \to -\infty$. Likewise as $\sigma_b^2 \to \infty$, $E[\Delta\ell_t|A, B, D]^{m,n+1} \to 0$, and because of the model's symmetry, $prob(E[\Delta\ell_t|A, B, D]^{m,n+1} > 0) \to 1/2$.

If the importance of the endogenous variables dominates that of the exogenous variables, then the information conveyed by the revelation of an endogenous variable is necessarily informative and that variable cannot be a red herring. Proposition 3 states that the probability of an exogenous variable being a red herring approaches zero as $\sigma_b^2 \to \infty$. It is not the case, however, that the probability of an endogenous variable being a red herring approaches 1/2 as $\sigma_b^2 \to 0$. If the importance of the exogenous variables dominates that of the endogenous variables, then the revelation of an endogenous variable is not informative and the probability of that variable being a red herring approaches one-half. If $\sigma_b^2 \to \infty$, then the associated expected welfare loss approaches infinity.

I now simulate the model to quantify the probability of a red herring under structural coefficients expectations. For these simulations, I assume that $N = M = 30$. In the baseline case, I set $\sigma_a^2 = \sigma_b^2 = \sigma_d^2 = 0.25$ and, $\sigma_g^2 = \sigma_u^2 = 1$. I then take 20,000 random draws from $[A, B, D]$ and evaluate the likelihood of a red herring for all possible combinations of *m* and *n*. To access the effect of the variance terms on the model, I also simulate three low variance cases where σ_a^2, σ_b^2, or σ_d^2 equal 1/16, and three high variance cases where σ_a^2, σ_b^2, or σ_d^2 equal 1. For the revelation of an endogenous variable, I simulate two additional cases where $\sigma_u^2 = 0.5$, or $\sigma_u^2 = 2$.[12]

(FIGURE 1 HERE)

Figure 1 displays the probability of the 15^{th} exogenous variable being a red herring upon its revelation for each possible value of *n*. The middle line represents the baseline case. The probability of a red herring approaches zero as *n* approaches N. The top line represents the calibration where σ_B^2 is low, and the bottom line represents the calibration where σ_B^2 is high. Consistent with Proposition 3, the probability of a red herring decreases as σ_B^2 increases. Although not shown in either Figure 1 or Figure 2, the behavior of the high σ_A^2 and high σ_D^2 calibrations are nearly identical to the low σ_B^2 case. Likewise, the low σ_A^2 and low σ_D^2 calibrations are nearly identical to the high σ_B^2 case.

(FIGURE 2 HERE)

Figure 2 graphs the likelihood that the m^{th} exogenous variable is a red herring for $n = 15$. Once again, higher values of σ_B^2 decrease the likelihood of a red herring.

(FIGURE 3 HERE)

Figure 3 plots the likelihood that the 15^{th} endogenous variable is a red herring for different values of *m*. The range of probabilities is similar to the results for the revelation of an exogenous variable. Consistent with Proposition 4, higher values of σ_B^2 now increase the likelihood of a red herring. Higher values of σ_d^2 and lower values of σ_u^2 also increase the probability of a red herring. Although not shown in Figure 3, the low σ_a^2 calibration behaves almost identically to the high σ_b^2 calibration and vice-versa.[13]

[12]The value of σ_u^2 does not affect the likelihood of red herrings upon the revelation of an exogenous variable.

[13]The results are similar if the n^{th} endogenous variable is revealed for a fixed value of m.

2.2 Adaptive Learning

Under structural coefficients expectations, when economic theory reveals an additional variable, it also reveals the structural coefficients associated with that variable. This is reminiscent of rational expectations where agents use a model's reduced form solution to form their expectations. The adaptive learning literature often criticizes rational expectations for endowing agents with an unrealistic amount of knowledge about the economy's true model. Adaptive learning relaxes the informational requirements of rational expectations by assuming that agents use standard econometric techniques to estimate the model's reduced form. It therefore typically assumes that agents *must* use adaptive learning because they lack the information needed to form rational expectations. This justification is easily extended to the present paper by assuming that agents are unable to form structural coefficients expectations.

There is, however, an additional justification for examining adaptive learning in this paper. In the model, agents would ideally use rational expectations where they know the structural coefficients on all $N + M$ endogenous and exogenous variables.[14] The assumption that agents know only a fraction of these variables, however, implies that agents cannot use rational expectations, but may rely instead on structural coefficients expectations. Whereas rational expectations are certain to yield better forecasts than adaptive learning, adaptive learning may or may not outperform structural coefficients expectations. In this case, agents are better off *choosing* to use adaptive learning instead of structural coefficients expectations. This section addresses two questions. First, are agents better off choosing to discard structural coefficients expectations in favor of adaptive learning? Second, how likely and significant are red herrings under the assumption that agents use adaptive learning?

Under rational expectations, agents form more accurate expectations than under adaptive learning. The possibility that agents may be better off under adaptive learning arises only because agents may possess structural coefficients expectations, not rational expectations. Several papers develop different cases where agents prefer adaptive learning to rational expectations. The most common approach allows agents to form rational expectations if and only if they incur an additional cost.[15] If the cost of rational expectations is sufficiently high, then some or all agents will prefer adaptive learning

[14] Proposition 1 shows agents' loss under rational expectations always equals zero.

[15] See Evans and Ramey (1992), Brock and Hommes (1997), and Branch and McGough (2005) for prominent examples from this literature.

to rational expectations. Adam (2005) develops a model where agents learn adaptively and choose between a correctly specified model (relative to rational expectations) and an underparameterized model. If agents use the former specification, adaptive learning asymptotically converges to rational expectations. If agents use the latter specification, however, the learning process causes both models to be misspecified and it is possible that the latter model yields a better forecast on average. Learning the misspecified model is optimal, however, only if agents have already been using that model.

To model adaptive learning, I assume that agents know that y_t linearly depends on the revealed endogenous and exogenous variables, but do not know the exact coefficients.

$$y_t^e = az_t^n + bg_t^m \tag{13}$$

Under adaptive learning, I assume that agents use recursive least squares to obtain a_t and b_t. This is similar to running an OLS regression of y_t on z_t and g_t, and updating that regression each period as new data become available. Equation (13) represents agents' perceived law of motion (PLM).

By inserting agents' PLM into Equations (2) and (3), I obtain the economy's actual law of motion (ALM).

$$y_t = ACaz_t^n + (ACb + AD^m + B^m)g_t^m + (AD^{-m} + B^{-m})g_t^{-m} + ACu_t \tag{14}$$

The mapping from the PLM to the ALM may therefore be written as:

$$T\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} ACa \\ ACb + AD^m + B^m \end{bmatrix} \tag{15}$$

Under adaptive learning, it is not typically the case that *a* and *b* will converge to their structural coefficients expectations or rational expectations values. This is therefore an example of a restricted perceptions equilibrium. A restricted perceptions equilibrium is optimal in the class of PLMs that agents are considering, but may be inferior to other types of PLMs. In this model, a restricted perceptions equilibrium implies that agents are forming the best possible econometric estimate, given that they do not know all the variables included in the model.

To calculate the fixed points of the learning process, I project the ALM onto the PLM. It is convenient to write both z_t^n and y_t as functions of g_t^m, g_t^{-m}, u_t^n, and u_t^{-n}.

$$z_t^n = \alpha g_t^m + \beta g_t^{-m} + \zeta u_t^n \tag{16}$$

where $\alpha = (I - C^n a)^{-1}(C^n b + D^{n,m})$, $\beta = (I - C^n a)^{-1} D^{n,-m}$, and $\zeta = (I - C^n a)^{-1}$.

$$y_t = \chi g_t^m + \delta g_t^{-m} + \omega u_t^n + A^{-n} u_t^{-n} \tag{17}$$

where $\chi = AC(1-aC^n)^{-1}a(D^{n,m}+C^n b)+AD^{N,m}+ACb+B^m$, $\delta = AC(1-aC^n)^{-1}aD^{n,-m}+AD^{N,-m} + B^{-m}$, and $\omega = A^n + AC(1 - aC^n)^{-1}a$.

$$Var \begin{bmatrix} z_t^n \\ g_t^m \end{bmatrix} = \begin{bmatrix} \sigma_g^2(\alpha\alpha' + \beta\beta') + \sigma_u^2 \zeta\zeta' & \sigma_g^2 \alpha \\ \sigma_g^2 \alpha' & \sigma_g^2 I(m) \end{bmatrix} \tag{18}$$

$$Cov \left[\begin{bmatrix} z_t^n \\ g_t^m \end{bmatrix}, y_t \right] = \begin{bmatrix} \sigma_g^2(\alpha\chi' + \beta\delta') + \sigma_u^2 \zeta\omega' \\ \sigma_g^2 \chi' \end{bmatrix} \tag{19}$$

The fixed point of the learning process is thus given by:

$$T \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sigma_g^2(\alpha\alpha' + \beta\beta') + \sigma_u^2 \zeta\zeta' & \sigma_g^2 \alpha \\ \sigma_g^2 \alpha' & \sigma_g^2 I(m) \end{bmatrix}^{-1} \begin{bmatrix} \sigma_g^2(\alpha\chi' + \beta\delta') + \sigma_u^2 \zeta\omega' \\ \sigma_g^2 \chi' \end{bmatrix} \tag{20}$$

For simplicity, I assume that agents do not include an intercept in their regression. If agents do include an intercept, then the fixed point of the intercept estimate equals zero.

The system is stable under adaptive learning if the learning coefficients remain in the neighborhood of their restricted perceptions equilibrium values. To evaluate stability under learning, I use the related concept of E-Stability. Evans and Honkapohja (2001) demonstrate that under general conditions, a model is stable under learning if and only if it is E-Stable. E-Stability maps from the PLM to the ALM using the E-Stability differential equation.

$$d/d\tau \begin{bmatrix} a \\ b \end{bmatrix} = T \begin{bmatrix} a \\ b \end{bmatrix} - \begin{bmatrix} a \\ b \end{bmatrix} \tag{21}$$

If each eigenvalue of the Jacobian of the right hand side of Equation (21) has real parts less than zero when evaluated at its fixed point, then the model is E-Stable. Evaluating this condition using the ALM, Equation (14), and the PLM, Equation (13), shows that the necessary and sufficient condition for E-Stability is that $AC < 1$.

There are two sources of red herrings under adaptive learning. The first source is related to the added noise that an additional regressor adds to the process. It is necessarily true that including an additional regressor results in an equal or better in-sample fit than before its inclusion. Likewise the new restricted perceptions equilibrium performs better than the older equilibrium at their respective fixed points. Often, however, the inclusion of an additional regressor will introduce enough noise to the estimate so that the enlarged set of regressors provides a worse out-of-sample fit than the original set of regressors. In this case, forecasts worsen, and the newly revealed variable is a red herring.

If the model includes endogeneity (neither A nor C are null matrices), then the Lucas Critique is a second source of red herrings. I assume that once theory reveals the $n+1^{th}$ endogenous variable, agents collect past data for that variable and regress *y* on z^{n+1} and g^m. Because y_t depends on its expectation, however, the revelation of an additional endogenous variable changes the underlying data generating process. The fixed points of a_t and b_t using old data may therefore be significantly different than their fixed points in the new restricted perceptions equilibrium. While the learning process returns a_t and b_t to the neighborhood of their restricted perceptions equilibrium values, agents may make poor forecasts. If their discount factor is sufficiently small, then the revelation of the new variable may result in a larger discounted loss. The same analysis applies if theory reveals the $m+1^{th}$ exogenous variable, and agents retroactively regress *y* on z^n and g^{m+1}.

To simulate adaptive learning, I assume that agents estimate Equation (2) using recursive least squares. Agents' *(n+m) x 1* vector of regressors, $\phi = [a_t, b_t]'$, at time *t* is updated according to:

$$\phi_t = \phi_{t-1} + \gamma R_t^{-1} \phi'_{t-1}(y_{t-1} - a_{t-1} z^n_{t-1} - b_{t-1} g^m_{t-1}) \tag{22}$$

$$R_t = (1-\gamma) R_{t-1} + \gamma(\phi_{t-1}\phi'_{t-1}) \tag{23}$$

This type of recursive least squares is similar to running an OLS regression each period except that the gain (γ) represents the weight placed on the most recent observation. Under standard OLS, the gain equals the inverse of the sample size and each observation counts equally. Equations (22) and (23) are an example of constant-gain learning where agents place extra weight on more recent observations.[16] Constant-gain learning is a popular way to model learning when the model that agents

[16]For more details on constant-gain learning, see Sargent (1999), and Evans and Honkapohja (2001).

estimate is subject to structural change, such as the revelation of a new variable. In the absence of constant-gain learning, agents would be even more vulnerable to the Lucas Critique.[17] Evans and Honkapohja (2001) show that under constant-gain learning with a sufficiently small gain, E-Stability generally implies that the learning coefficients converge to normal distributions with means equal to their fixed points and variances proportional to the gain.

Asymptotically, adaptive learning dominates structural coefficients expectations. Although it yields biased regression coefficients if any variables are unrevealed, the resulting expectation of y_t is unbiased. In the absence of econometric noise, the ensuing forecasts are optimal subject to not knowing the unrevealed variables. Under structural coefficients expectations, $a = A^n$ and $b = B^m$. Structural coefficients expectations therefore provide noiseless coefficients, but at the expense of ignoring the correlation between z_t^n and g_t^{-m}, and thus being a suboptimal projection. The ensuing estimates are thus biased, conditional on the entire history of the model's variables.

Under the learning algorithm from Equations (22) and (23), there are two mechanisms whereby structural coefficients expectations may outperform adaptive learning. First, persistent learning dynamics add econometric noise to the forecasts generated by adaptive learning. Second, the occasional revelation of an additional variable adds bias to the learning algorithm through the Lucas Critique, and also prevents the learning algorithm from converging to its asymptotic properties even if agents use a decreasing gain. It is thus not obvious which type of expectations formation yields smaller errors.

Under structural coefficients expectations, the revelation of a new endogenous variable increases agents' expected loss with probability between zero and one-half. I now examine the welfare effects of discovering a new endogenous variable by simulating the model under adaptive learning. I begin the learning process by setting ϕ equal to a null vector and imposing a 1000 period burn to minimize the effects of this initial condition. Each period, I draw g_t from its distribution while using a gain equal to 0.005. I then calculate agents' loss for 500 additional periods under three different scenarios. In the first scenario, economic theory reveals the $n + 1^{th}$ endogenous variables and agents use this additional variable to forecast y_t. I assume that agents are able to obtain past data on the newly revealed variable, and that they then retroactively run the learning algorithm from Equations (22) and (23) to include the newly revealed variable. The expanded specification is then used for forecasting.

[17] A higher gain eliminates the bias from the system faster, but increases the variance of the learning parameters. Solving for and studying the optimal gain is left for future research.

In the second scenario, economic theory does not reveal the new variable and the model's behavior is unchanged from the burn period. In the third scenario, I assume that economic theory does reveal the new variable but that agents possess structural coefficients expectations as discussed in Subsection (2.1).[18] I then calculate agents' discounted loss under all three cases.

To limit computational time, I do not simulate every pair of *m* and *n*. Instead, I simulate every pair where *m* equals 2,5,8... and *n* initially equals 2,5,8,... I conduct 500 simulations for each pair of *m* and *n*. In all simulations, I set $\sigma_g^2 = 1$. My baseline calibration sets $\sigma_j^2 = 0.25$ for $j = a, b, c, d$, $\sigma_u^2 = 1$, and assumes that agents use a discount factor of 0.99. For comparison, I consider several other calibrations. I simulate five high variance cases where $\sigma_j^2 = 1$ for $j = a, b, c, d$, or where $\sigma_u^2 = 2$. I also simulate five low variance cases where $\sigma_j^2 = 1/16$ for $j = a, b, c, d$, or where $\sigma_j^2 = 1$. I only evaluate draws of $[A, C]$ that are stable under learning by discarding and replacing any draws where $|AC| > 1$.[19]

Tables 1-3 show the probability that structural coefficients expectations yield a lower discounted loss than adaptive learning upon the revelation of an additional endogenous variable.

(Table 1 HERE)

(Table 2 HERE)

(Table 3 HERE)

For the baseline and majority of calibrations, structural coefficients expectations outperform adaptive learning. In these cases, even though adaptive learning does better at its fixed point than structural coefficients expectations, the noise introduced by adaptive learning makes structural coefficients expectations preferable. In four cases, however; high σ_d^2, high σ_b^2, low σ_a^2, and low σ_u^2, adaptive learning outperforms structural coefficients expectations. These are the four cases with the strongest correlation between z_t and g_t. Adaptive learning exploits the correlation between the regressors and omitted variables and therefore performs better in these cases.

[18] The results are similar if economic theory does not reveal an additional variable and agents use structural coefficients expectations.

[19] Although draws where $AC < -1$ are stable under learning, I discard these draws to preserve the model's symmetry.

Tables 4-6 report the probability that a newly revealed endogenous variable is a red herring under adaptive learning.

(Table 4 HERE)

(Table 5 HERE)

(Table 6 HERE)

The probability of a red herring under adaptive learning is typically less than one-half. Two calibrations, however, high σ_a^2 and high σ_c^2, exhibit particularly high probabilities of a red herring. In both of these cases, the probability of a red herring exceeds one-half when agents initially know most of the relevant variables. These two cases include heightened endogeneity, which strengthens the effect of the Lucas Critique.

Tables 4-6 show that, in most cases, agents are better off using the newly revealed variable to forecast if they are concerned with the median outcome. Table 7 considers the average outcome by reporting the expected welfare loss from the revelation of the 15^{th} endogenous variable under adaptive learning.[20]

(Table 7 HERE)

Although the median loss is positive for only two calibrations, high σ_a^2 and high σ_c^2, the average loss is positive for four of the eleven calibrations. This result occurs because the Lucas Critique produces a particular type of red herring which occurs infrequently but with a potentially catastrophic welfare loss. Equation (17) shows that χ includes the term $(1 - aC^n)^{-1}$ and the model therefore contains a singularity where $aC^n = 1$. Thus as $aC^n \to 1$, z_t, y_t^e, y_t, and likely ℓ_t explode. Formally, a newly revealed variable is a catastrophic red herring if the ratio of the welfare loss after its revelation to the welfare loss had theory not revealed that variable exceeds some arbitrarily large level.

An unbiased learning algorithm tends to keep the model away from its singularity, and catastrophic red herrings occur with an exceptionally low, but positive, probability. If $\sigma_c^2 \neq 0$, however, then a new

[20]Table 7 multiplies the average loss by ten.

revelation introduces bias and catastrophic red herrings occur more often. The coefficients obtained by retroactively regressing y on z^{n+1} and g^m after the revelation will, with increased probability, cause aC^n to be arbitrarily close to one.

Table 7 shows that catastrophic red herrings occur systematically when the subset of revealed variables is large, and when the model includes high levels of endogeneity (high σ_a^2 or high σ_c^2) or when the endogenous variables exhibit low levels of noise (low σ_u^2). They can, however, occur with lower frequency for any of the other calibrations. At least one such catastrophic red herring occurs in the simulations where σ_d^2 is low, $m = 11$, and $n = 14$, causing the average loss to be positive for that calibration.[21]

This intuition also applies to the revelation of the $m + 1^{th}$ exogenous variable. Table 8 reports the probability that a newly revealed exogenous variable is a red herring, and Table 9 reports the average welfare loss.

(Table 8 HERE)

(Table 9 HERE)

The most striking difference between Table 8 and the results for the revelation of an endogenous variable is that red herrings now typically occur with probability greater than one-half. Table 9 shows that newly revealed exogenous variables may also be catastrophic red herrings. As before, catastrophic red herrings occur with particularly high frequency under the two calibrations that exhibit increased endogeneity. In all eleven calibrations, the average welfare loss is positive.

To check robustness, I also simulate four additional cases: a high gain ($\gamma = 0.15$), a low gain ($\gamma = 0.01$), a high discount factor ($\beta = 0.999$), and a low discount factor ($\beta = 0.90$).[22] Table 10 reports the results for the revelation of an additional endogenous variable and Table 11 reports those for the revelation of an additional exogenous variable.

[21] Using the mean welfare loss instead of the median welfare loss does not affect the comparison of adaptive learning and structural coefficients expectations. The former continues to outperform the latter in all but the four calibrations discussed earlier in this section.

[22] I increase the burn period to 2000 observations for the low gain case and increase the simulation length to 1000 periods for the high discount factor case.

(Table 10 HERE)

(Table 11 HERE)

Throughout this paper, I have assumed constant gain learning. A higher gain is similar to using a shorter sample period and usually causes both quicker convergence and more volatile beliefs after convergence. High gains may be especially sensible when agents are relatively concerned about structural change.[23] For the revelation of an additional exogenous variable, alternate values of the gain have little effect. A lower gain, however, significantly reduces the probability that a newly revealed endogenous variable is a red herring. The value of the optimal gain, and the process by which agents find it, is left for future research.

3 An Example: The New Keynesian Model

I now present an example that demonstrates how the general framework of this paper can easily be extended to analyze popular macroeconomic DSGE models. The New Keynesian model has emerged as the workhorse of monetary policy analysis. A simple version of the New Keynesian model consists of an Euler/IS Equation, a New Keynesian Phillips Curve, and an interest rate rule.[24]

$$\tilde{\pi}_t = \beta E_t[\tilde{\pi}_{t+1}] + \kappa(\tilde{x}_t - z_t) + \Omega^\pi \chi_t \tag{24}$$

$$\tilde{x}_t = E_t[\tilde{x}_{t+1}] + \tau(E_t[\tilde{\pi}_{t+1}] - \tilde{R}_t) + g_t + \Omega^x \chi_t \tag{25}$$

[23] See Evans and Honkapohja (2001) for more details of how the gain affects learning dynamics. Because the gain may effect welfare in other ways, finding the optimal gain is not the same as finding the gain that minimizes losses from red herrings.

[24] A few papers show that optimal policy in simple New Keynesian models depends on if agents form expectations using bounded rationality. Orphanides and Williams (2008) show that, under learning, optimal policy requires that interest rates respond strongly to inflation in order to stabilize inflationary expectations. Christev and Slobodyan (2014) consider constant gain learning and show that optimality requires a strong policy response to inflation that results in both a fast convergence speed and less volatile expectations. Salle, Yildizoğlu, and Sénégas (2013) find that in an agent based New Keynesian model with social learning, optimality requires a credible and explicit inflation target.

$$\tilde{R}_t = \lambda E_t[\tilde{\pi}_{t+1}] \tag{26}$$

where $\tilde{x}_t$, $\tilde{\pi}_t$, and $\tilde{R}_t$ represent the percentage deviation from the steady states of output, inflation, and the nominal interest rate respectively. The error term g_t is often interpreted as a white noise preference shock. The error term z_t is often interpreted as a positively serially correlated measure of potential output. Equation (26) assumes that the monetary authority can perfectly control the nominal interest rate and only targets inflation.

I have added a $Kx1$ vector of unrevealed variables, χ_t, to the model. These variables may be interpreted as those that influence micro-level agents' decisions, but that are not included in the model. I assume that agents are constrained by the theoretical model when attempting to forecast aggregate inflation and output.[25] To evaluate the model under learning, I assume that all agents initially use the following specification:

$$\tilde{\pi}_t = \alpha_{0,t}\tilde{\pi}_{t-1} + \alpha_{1,t}\tilde{x}_{t-1} + \alpha_{2,t}z_t + \alpha_{3,t}g_t + \epsilon_{\pi,t} \tag{27}$$

$$\tilde{x}_t = \alpha_{5,t}\tilde{\pi}_{t-1} + \alpha_{6,t}\tilde{x}_{t-1} + \alpha_{7,t}z_t + \alpha_{8,t}g_t + \epsilon_{x,t} \tag{28}$$

where the regression coefficients are obtained through constant-gain learning.[26] The omission of intercept terms implies that agents know the model's steady state. I assume that agents know the autoregressive coefficients for all stochastic shocks. By re-dating Equations (27) and (28) and taking expectations of g_t and z_t, the model may be restated as:

$$\begin{bmatrix} \tilde{\pi}_t \\ \tilde{x}_t \end{bmatrix} = J^{-1} \begin{bmatrix} \beta\rho_z\alpha_{2,t} - \kappa & \beta\rho_g\alpha_{3,t} \\ \rho_z(\alpha_{7,t} + \tau(1-\lambda)\alpha_{2,t}) & 1 + \rho_g(\alpha_{8,t} + \tau(1-\lambda)\alpha_{3,t}) \end{bmatrix} \begin{bmatrix} z_t \\ g_t \end{bmatrix} + J^{-1} \begin{bmatrix} \Omega^\pi \\ \Omega^x \end{bmatrix} \chi_t \tag{29}$$

[25] Although in most versions of this model, the representative household's consumption depends on the expected aggregate price levels and its own micro-level consumption (which, in equilibrium, equals aggregate consumption). It would thus be reasonable to allow $E_t[\tilde{x}_{t+1}]$ to depend on χ_t. To keep the example as simple as possible, however, I assume that all expectations are formed without observing χ_t.

[26] I also consider this example where agents rely on the lagged values of stochastic shocks instead of their contemporaneous values. The results are similar.

$$J = \begin{bmatrix} 1-\beta\alpha_{0,t} & -\kappa-\beta\alpha_{1,t} \\ -\alpha_{5,t}-\tau(1-\lambda)\alpha_{0,t} & 1-\alpha_{6,t}-\tau(1-\lambda)\alpha_{1,t} \end{bmatrix} \quad (30)$$

To simulate the model, I set $\beta = 0.99$ and $\lambda = 1.5$, both common calibrations in the related literature. There exists considerable disagreement over the proper values of τ and κ. Lubik and Schorfheide (2004) estimate a similar version of the New Keynesian model and obtain posterior gamma distributions with means of 0.58 and 1.86 for κ and τ^{-1} using U.S. data beginning in 1982, after the Volker disinflation .[27] I draw from these distributions for each simulation of the model. Following the related literature, I set the AR(1) coefficient for potential output (ρ_z) to 0.95 and that for preference shocks (ρ_g) to zero. I assume that each element of χ_t follows an AR(1) process with each AR(1) coefficient ($\rho_{\chi,i}$) drawn from a uniform distribution between 0 and 1, and that innovations to all random error terms have zeros means, standard deviations of 0.01 and are orthogonal to each other. Finally, each element of Ω is drawn from a standard normal distribution. Each simulation begins with draws from the distributions of the relevant parameters. I assume that the model begins at its steady state and I initially set all learning coefficients equal to zero. After a burn period, I assume that economic theory reveals the first element of χ_t and that agents are able to retroactively include this shock in their econometric specifications.

$$\tilde{\pi}_t = \alpha_{0,t}\tilde{\pi}_{t-1} + \alpha_{1,t}\tilde{x}_{t-1} + \alpha_{2,t}z_t + \alpha_{3,t}g_t + \alpha_{4,t}\chi_{1,t} + \epsilon_{\pi,t} \quad (31)$$

$$\tilde{x}_t = \alpha_{5,t}\tilde{\pi}_{t-1} + \alpha_{6,t}\tilde{x}_{t-1} + \alpha_{7,t}z_t + \alpha_{8,t}g_t + \alpha_{9,t}\chi_{1,t} + \epsilon_{x,t} \quad (32)$$

Inflation and output now evolve according to:

$$\begin{bmatrix} \tilde{\pi}_t \\ \tilde{x}_t \end{bmatrix} =$$

$$J^{-1} \begin{bmatrix} \beta\rho_z\alpha_{2,t}-\kappa & \beta\rho_g\alpha_{3,t} & \beta\rho_{\chi,1}\alpha_{4,t}+\Omega_1^{\pi} \\ \rho_z(\alpha_{7,t}+\tau(1-\lambda)\alpha_{2,t}) & 1+\rho_g(\alpha_{8,t}+\tau(1-\lambda)\alpha_{3,t}) & \rho_{\chi,1}(\alpha_{9,t}+\tau(1-\lambda)\alpha_{4,t})+\Omega_1^{x} \end{bmatrix} \begin{bmatrix} z_t \\ g_t \\ \chi_{1,t} \end{bmatrix} + \quad (33)$$

[27] Shea (2008), however, demonstrates that estimates of κ and τ vary widely based on the specific estimation technique and exact specification of the model.

$$J^{-1}\begin{bmatrix}\hat{\Omega}^{\pi}\\ \hat{\Omega}^{x}\end{bmatrix}\chi_t^{'} \tag{34}$$

where $\hat{\chi}_t$ is the last $K-1$ rows of χ_t and $\hat{\Omega}^i$ is the last $K-1$ columns of Ω^i. I compare the performance of the economy for the model where the new variable is revealed versus the model where it is not. Monetary policy is usually evaluated in this framework by measuring the weighted sum of squared deviations of inflation and the output gap. Because the Lucas Critique implies that forecasting is likely to improve as time passes after the revelation of a new variable, it is necessary to discount these deviations. Woodford (2003) derives the following loss function as a second-order quadratic approximation of the representative household's utility function:

$$V=\sum_{j=0}^{\infty}\beta^j[\tilde{\pi}_{t+j}^2+\frac{\kappa}{\theta}(\tilde{x}_{t+j}-z_{t+j})^2] \tag{35}$$

I calibrate $\theta = 20$ which implies a steady state markup of five percent.[28] Both sources of red herrings are present in this example. Clearly, if the determinate of J is sufficiently close to zero, then both output and inflation are likely to explode and cause a catastrophic red herring. I set the gain equal to 0.01 and simulate the model for 1000 draws for various values of K.[29] The results are reported below where ψ is the ratio of the discounted loss after the revelation to the loss had the revelation not occurred.

(Table 12 HERE)

These results are similar to those for the general model. If agents know most of the model (i.e. K is small) then the additional explanatory power of a newly revealed variable is unlikely to outweigh the added noise that it introduces. Red herrings thus occur often. As agents know less of the model, the benefits of adding another explanatory variable are greater and red herrings occur less often. As with the general model, it is possible that red herrings yield a substantial welfare loss. The third column

[28]Woodford (2003) also shows that κ is a function of the model's deeper structural parameters, including θ and β. By fixing θ and β, I am therefore implicitly assuming that the uncertainty of κ is derived from other parameters, most likely the uncertainty over the fraction of firms that may change their price in any period.

[29]Draws where the model is not stable under adaptive learning are discarded and replaced.

reports the probability that the revelation results in a welfare loss at least three times greater than had the revelation not occurred.[30] Once again, large welfare losses are more likely when agents know most of the model. It is also possible that the revelation of a new variable leads to a comparable welfare improvement, although the fourth column demonstrates that these events occur far less frequently.

4 Conclusion

This paper examines agents' forecasting problem in an environment where they know only a subset of the relevant exogenous and endogenous variables. Under structural coefficients expectations, agents use the structural coefficients associated with the revealed variables to forecast. Under adaptive learning, agents either do not know, or discard their knowledge of these structural coefficients, and instead regress the variable of interest on all revealed variables.

The paper presents two important results. First, if the exogenous and endogenous variables are significantly correlated, then agents should choose to use adaptive learning, even if structural coefficients expectations are feasible. If this correlation is weak, however, then agents are better off using structural coefficients expectations. Second, the revelation of a new variable often worsens forecasting. Under structural coefficients expectations, red herrings often occur, but never with probability greater than one-half. Under adaptive learning, however, the probability of a red herring is often greater than one-half. Furthermore, if the model includes endogeneity, then a different type of red herring will occasionally lead to a catastrophic loss in welfare.

In this paper, I assume that agents include all revealed variables in their econometric specification. As is standard in the learning literature, agents do not conduct specifications tests, or make statistical inferences as they learn. Allowing for more sophisticated agents in this framework could potentially yield some interesting results. It would be of interest, for example, to allow agents to adjust the gain in the learning algorithm based on the effects of past revelations. Additionally, the New Keynesian model analyzed in Section 3 is far simpler than those used by most Central Banks when making

[30]As with the general case, if the model is sufficiently close to its singularity, then a catastrophic red herring can produce an arbitrarily large welfare loss which significantly increases the average welfare loss across all simulations. I refrain from reporting the expected welfare loss, however, because the explosive behavior of inflation and output pushes the model beyond the region where its linearization is valid.

forecasts. Extending the framework of this paper to more complex New Keynesian models is also worthy of examination. These topics are left for future research.

References

Adam, K. 2005. "Learning to Forecast and Cyclical Behavior of Output and Inflation." *Macroeconomic Dynamics* 9: 1-27.

Branch, W., and G. Evans. 2006. "Intrinsic Heterogeneity in Expectation Formation." *Journal of Economic Theory,* 127: 264-295.

Branch, W., and B. McGough. 2008. "Replicator Dynamics in a Cobweb Model with Rationally Heterogeneous Expectations." *Journal of Economic Behavior and Organization* 65(2): 224-244.

Brock, W., and C. Hommes. 1997. "A Rational Route to Randomness." *Econometrica,* 65(5): 1059-1095.

Christev, A. and S. Solbodyan. 2014. "The Welfare Consequences of Adaptive Learning Dynamics." Presented at the 46th Annual Money, Macro, and Finance Conference, September 17-19, Durham University.

Cho, I., and K. Kasa. 2015. "Learning and Model Validation." *Review of Economic Studies,* 82(1): 45-82.

Evans, G., and G. Ramey. 2006. "Adaptive Expectations, Underparameterization and the Lucas Critique." *Journal of Monetary Economics,* 53: 249-264.

Evans, G., and G. Ramey. 1992. "Expectation Calculation and Macroeconomic Dynamics." *The American Economic Review,* 82: 207-224.

Evans, G., and S. Honkapohja. 2001. *Learning and Expectations in Macroeconomics*. Princeton, NJ: Princeton University Press.

Farmer, R., and J. Guo. 1994. "Real Business Cycles and the Animal Spirits Hypothesis." *Journal of Economic Theory*, 63(1): 42-72.

Granger, C. 1969. "Prediction with a Generalized Cost of Error Function." *Operations Research Quarterly,* 20: 199-207.

Granger, C. and P. Newbold. 1986. *Forecasting Economic Time Series,* 2nd edition, New York.: Academic Press.

Hendry, D. and M. Clements. 2003. "Economic Forecasting: Some Lessons From Recent Research." *Economic Modelling*, 20: 301-329.

Hendry, D. and M. Clements. 1998. *Forecasting Economic Time Series*. Cambridge, UK. Cambridge University Press.

Kydland, F., and E. Prescott. 1982. "Time to Build and Aggregate Fluctuations." *Econometrica,* 50(6): 1345-1370.

Lubik, T., and F. Schorfheide. 2004. "Testing for Indeterminacy: An Application to U.S. Monetary Policy." *The American Economic Review,* 94(1): 190-217.

Lucas, R. 1976. "Econometric Policy Evaluation: a Critique." *in K. Brunner and A. Meltzer, eds, The Phillips Curve and the Labor Market, Vol. 1 of the Carnegie-Rochester Conferences in Public Policy, a Supplemental Series to the Journal of Monetary Economics (North Holland-Amsterdam)*: 19-46.

Orphanides, A. and J. Williams. 2008. "Learning, Expectations Formation, and the Pitfalls of Optimal Control monetary Policy." *Journal of Monetary Economics,* 55: S80-S96.

Patton, A. and A. Timmermann. 2007. "Properties of Optimal Forecasts Under Asymmetric Loss and Nonlinearity." *Journal of Econometrics*, 140(2): 884-918.

Quinion, M. 2002. "The Lure of the Red Herring." *WorldWideWords*, url: www.worldwidewords.org/articles/herring

Roberts, J. 1995. "New Keynesian Economics and the Phillips Curve." *Journal of Money Credit and Banking,* 27(4): 975-984.

Salle, I., Yildizoğlu, M., and M. Sénégas. 2013. "Inflation targeting in a Learning Economy: An ABM Perspective." *Economic Modelling,* 34: 114-128.

Sargent, T. 1999. *The Conquest of American Inflation*. Princeton, NJ: Princeton University Press.

Shea, P. 2008. "Real-Time Rational Expectations and Indeterminacy." *Economics Letters,* 99(3): 530-533.

White, Halbert. 1994. *Estimation, Inference and Specification Analysis*. New York, NY. Cambridge University Press.

Woodford, M. 2003. *Interest and Prices*. Princeton, NJ and Oxford: Princeton University Press.

Figures and Tables

Figure 1: Effect of Discovering the Fifteenth Exogenous Variable

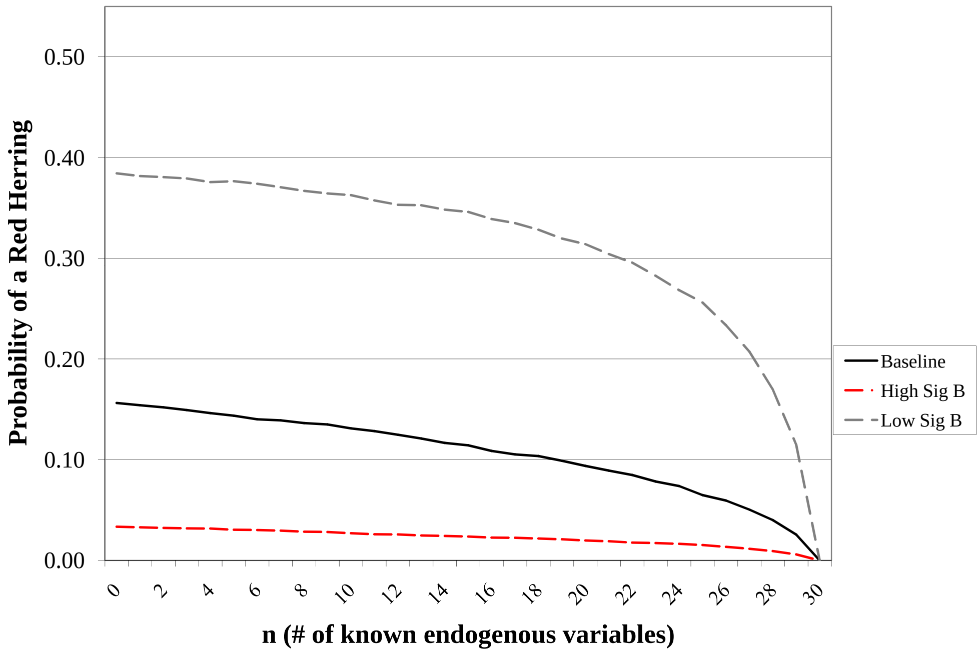

Figure 2: Effect of Discovering the m*th* Exogenous Variable (n = 15)

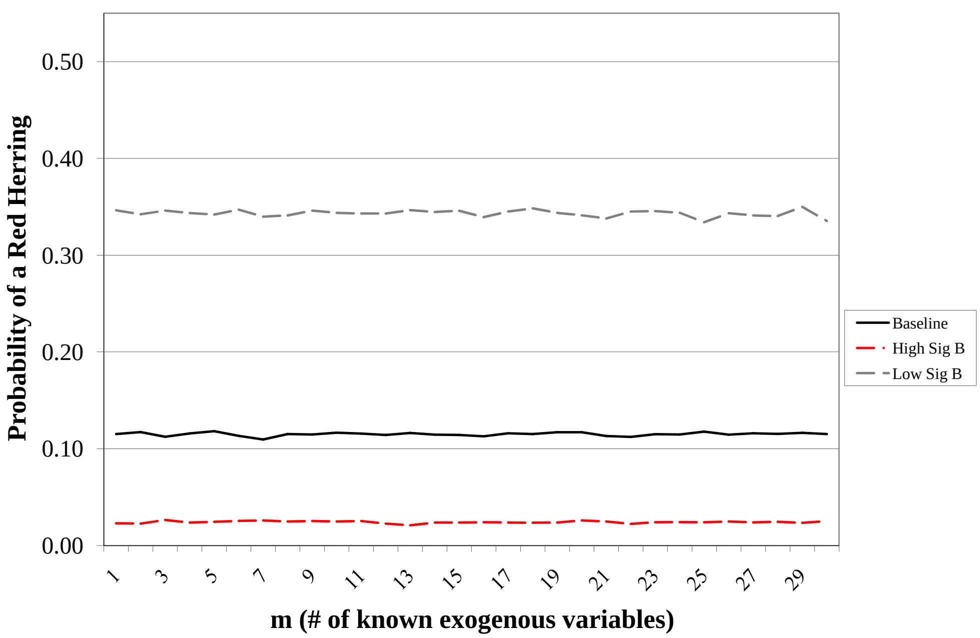

Figure 3: Effect of Discovering the Fifteenth Endogenous Variable

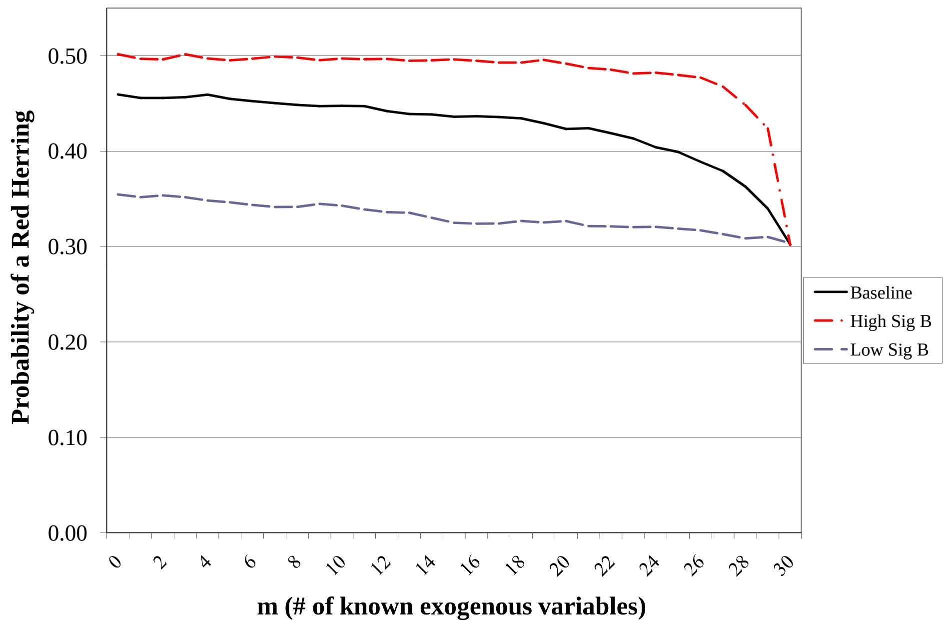

Table 1

Probability that Structural Coefficients Expectations Dominate Adaptive Learning (n = 15)

		High Variances					**Low Variances**				
m	BL	σ_a^2	σ_b^2	σ_c^2	σ_d^2	σ_u^2	σ_a^2	σ_b^2	σ_c^2	σ_d^2	σ_u^2
2	0.39	0.45	0.22	0.36	0.00	0.85	0.24	0.42	0.41	0.92	0.00
5	0.61	0.72	0.09	0.55	0.00	0.97	0.09	0.80	0.61	0.95	0.00
8	0.73	0.88	0.10	0.74	0.00	0.99	0.05	0.91	0.74	0.98	0.00
11	0.86	0.97	0.03	0.88	0.00	1.00	0.04	0.97	0.83	0.99	0.00
14	0.89	0.99	0.05	0.94	0.00	1.00	0.02	1.00	0.87	1.00	0.00
17	0.95	1.00	0.02	0.97	0.01	1.00	0.01	1.00	0.94	1.00	0.00
20	0.96	1.00	0.03	0.99	0.03	1.00	0.02	1.00	0.94	1.00	0.00
23	1.00	1.00	0.03	1.00	0.16	1.00	0.01	1.00	1.00	1.00	0.00
26	0.99	1.00	0.04	1.00	0.50	1.00	0.00	1.00	0.99	1.00	0.02
29	1.00	1.00	0.02	1.00	0.98	1.00	0.02	1.00	1.00	1.00	0.08
All	0.84	0.90	0.06	0.84	0.17	0.98	0.05	0.91	0.83	0.98	0.01

Table 2

Probability that Structural Coefficients Expectations Dominate Adaptive Learning (m = 14)

		High Variances					**Low Variances**				
n	BL	σ_a^2	σ_b^2	σ_c^2	σ_d^2	σ_u^2	σ_a^2	σ_b^2	σ_c^2	σ_d^2	σ_u^2
3	0.90	1.00	0.02	0.90	0.02	1.00	0.02	0.99	0.88	0.98	0.00
6	0.89	1.00	0.02	0.88	0.01	1.00	0.01	1.00	0.90	1.00	0.00
9	0.90	1.00	0.01	0.94	0.00	1.00	0.02	0.99	0.89	1.00	0.00
12	0.91	1.00	0.04	0.93	0.01	1.00	0.03	0.99	0.90	0.99	0.00
15	0.89	0.99	0.05	0.94	0.00	1.00	0.02	1.00	0.87	1.00	0.00
18	0.88	1.00	0.05	0.95	0.00	1.00	0.03	0.99	0.87	1.00	0.00
21	0.90	0.99	0.07	0.96	0.00	1.00	0.05	0.98	0.89	1.00	0.00
24	0.89	0.99	0.11	0.94	0.00	1.00	0.09	0.97	0.91	1.00	0.00
27	0.93	0.99	0.24	0.96	0.00	1.00	0.19	0.96	0.91	1.00	0.03
30	0.95	0.98	0.58	0.96	0.00	1.00	0.62	0.95	0.91	1.00	0.13
All	0.90	0.99	0.12	0.94	0.00	1.00	0.11	0.98	0.89	1.00	0.02

Table 3

Prob. that Structural Coefficients Expectations Dominate Adaptive Learning (m = n-1)

		High Variances					**Low Variances**				
n	BL	σ_a^2	σ_b^2	σ_c^2	σ_d^2	σ_u^2	σ_a^2	σ_b^2	σ_c^2	σ_d^2	σ_u^2
3	0.58	0.73	0.21	0.55	0.08	0.91	0.20	0.78	0.60	0.85	0.09
6	0.71	0.86	0.08	0.64	0.01	0.98	0.10	0.86	0.72	0.94	0.00
9	0.78	0.94	0.05	0.75	0.00	0.99	0.05	0.95	0.78	0.98	0.00
12	0.83	0.98	0.03	0.86	0.00	1.00	0.02	0.98	0.84	0.98	0.00
15	0.89	0.99	0.05	0.94	0.00	1.00	0.02	1.00	0.87	1.00	0.00
18	0.95	1.00	0.05	0.98	0.01	1.00	0.02	1.00	0.94	1.00	0.00
21	0.98	1.00	0.07	0.99	0.05	1.00	0.02	1.00	0.98	0.99	0.02
24	1.00	1.00	0.17	1.00	0.23	1.00	0.10	1.00	0.99	1.00	0.22
27	1.00	1.00	0.37	1.00	0.66	1.00	0.34	1.00	1.00	1.00	0.65
30	1.00	1.00	0.91	1.00	1.00	1.00	0.82	1.00	1.00	1.00	0.99
All	0.87	0.95	0.20	0.87	0.21	0.99	0.17	0.96	0.87	0.97	0.20

Table 4

Probability that the 15^{th} Endogenous Variable is a Red Herring

		High Variances					**Low Variances**				
m	BL	σ_a^2	σ_b^2	σ_c^2	σ_d^2	σ_u^2	σ_a^2	σ_b^2	σ_c^2	σ_d^2	σ_u^2
2	0.38	0.42	0.29	0.39	0.20	0.46	0.28	0.38	0.37	0.49	0.18
5	0.40	0.48	0.27	0.39	0.21	0.43	0.29	0.42	0.37	0.45	0.20
8	0.37	0.48	0.29	0.40	0.20	0.43	0.31	0.34	0.40	0.47	0.24
11	0.37	0.46	0.30	0.47	0.23	0.41	0.27	0.44	0.38	0.48	0.19
14	0.44	0.47	0.32	0.50	0.27	0.44	0.28	0.43	0.38	0.45	0.22
17	0.38	0.49	0.30	0.53	0.25	0.46	0.28	0.37	0.39	0.46	0.29
20	0.38	0.49	0.34	0.50	0.30	0.44	0.31	0.37	0.36	0.48	0.40
23	0.39	0.52	0.33	0.49	0.35	0.46	0.26	0.45	0.38	0.50	0.46
26	0.43	0.50	0.39	0.51	0.36	0.45	0.35	0.37	0.42	0.48	0.52
29	0.40	0.48	0.40	0.51	0.39	0.44	0.35	0.44	0.40	0.49	0.55
All	0.40	0.48	0.32	0.47	0.28	0.44	0.30	0.40	0.38	0.47	0.33

Table 5

Probability that the $n+1^{th}$ Endogenous Variable is a Red Herring (m = 14)

		High Variances					**Low Variances**				
n	BL	σ_a^2	σ_b^2	σ_c^2	σ_d^2	σ_u^2	σ_a^2	σ_b^2	σ_c^2	σ_d^2	σ_u^2
2	0.39	0.50	0.26	0.46	0.20	0.43	0.26	0.39	0.34	0.50	0.20
5	0.42	0.50	0.30	0.44	0.21	0.42	0.30	0.38	0.35	0.48	0.18
8	0.38	0.48	0.26	0.46	0.22	0.44	0.25	0.42	0.36	0.48	0.21
11	0.39	0.47	0.31	0.48	0.24	0.43	0.28	0.39	0.42	0.48	0.22
14	0.44	0.47	0.32	0.50	0.27	0.44	0.28	0.43	0.38	0.45	0.22
17	0.41	0.50	0.27	0.51	0.25	0.47	0.31	0.42	0.35	0.42	0.27
20	0.40	0.51	0.32	0.49	0.29	0.45	0.27	0.42	0.42	0.47	0.36
23	0.42	0.55	0.34	0.49	0.37	0.45	0.30	0.41	0.37	0.47	0.42
26	0.43	0.54	0.38	0.57	0.37	0.46	0.40	0.42	0.44	0.50	0.55
29	0.47	0.61	0.44	0.56	0.43	0.45	0.40	0.48	0.44	0.48	0.62
All	0.41	0.51	0.32	0.50	0.29	0.44	0.31	0.42	0.39	0.47	0.33

Table 6

Probability that the $n+1^{th}$ Endogenous Variable is a Red Herring (m = n)

		High Variances					**Low Variances**				
n	BL	σ_a^2	σ_b^2	σ_c^2	σ_d^2	σ_u^2	σ_a^2	σ_b^2	σ_c^2	σ_d^2	σ_u^2
2	0.31	0.38	0.30	0.33	0.21	0.40	0.30	0.40	0.35	0.46	0.28
5	0.36	0.41	0.27	0.40	0.21	0.43	0.26	0.39	0.35	0.45	0.22
8	0.39	0.43	0.27	0.40	0.19	0.47	0.27	0.40	0.37	0.48	0.24
11	0.37	0.48	0.28	0.49	0.19	0.43	0.25	0.43	0.37	0.47	0.21
14	0.44	0.47	0.32	0.50	0.27	0.44	0.28	0.43	0.38	0.45	0.22
17	0.40	0.51	0.35	0.45	0.27	0.46	0.34	0.43	0.38	0.47	0.38
20	0.46	0.52	0.37	0.51	0.36	0.48	0.34	0.44	0.43	0.48	0.55
23	0.42	0.52	0.42	0.54	0.42	0.48	0.34	0.44	0.41	0.51	0.62
26	0.45	0.59	0.46	0.51	0.46	0.50	0.38	0.46	0.47	0.51	0.63
29	0.47	0.60	0.56	0.57	0.42	0.52	0.43	0.53	0.42	0.56	0.59
All	0.41	0.49	0.36	0.47	0.30	0.46	0.32	0.43	0.39	0.48	0.39

Table 7

Expected Welfare Loss When the $n+1^{th}$ Endogenous Variable is Revealed (m = 14)

Relative to Structural Coefficients Expectations

		High Variances					**Low Variances**				
n	BL	σ_a^2	σ_b^2	σ_c^2	σ_d^2	σ_u^2	σ_a^2	σ_b^2	σ_c^2	σ_d^2	σ_u^2
2	-6.6	-6.6	-52.0	-4.3	-59.5	-8.9	-3.1	-5.2	-7.6	-5.4	-9.6
5	-5.1	9.7	-37.8	-1.4	-54.3	-7.5	-2.8	-4.3	-6.7	-35.0	-9.5
8	-4.6	10.2	-42.7	7.1	-46.6	-4.5	-2.7	-3.5	-5.5	-0.3	-7.7
11	-4.4	-32.4	-31.7	0.3	-41.5	-6.9	-2.2	-3.3	-4.5	504.3	-6.5
14	-4.7	17.0	-25.8	-2.1	-24.9	-6.6	-1.5	-2.6	-4.8	-0.3	-6.0
17	-3.5	94.7	-131.8	2.9	-16.9	-2.2	-1.5	-3.7	-4.5	-1.2	-3.0
20	-2.3	-70.0	-18.6	6.8	-10.1	-27.9	-1.1	-2.3	-3.1	-3.9	-0.9
23	-2.9	1409.3	-13.3	20.1	-4.7	-3.8	-0.7	-1.6	-2.8	0.4	-0.1
26	-2.6	2076.8	-8.6	52.5	-1.7	-1.9	-0.3	-1.7	-1.9	-3.3	89.9
29	-0.4	891.0	-1.0	91.9	-0.5	-3.3	-0.1	-0.4	-0.7	0.6	158.5
All	-3.7	440.0	-36.3	17.4	-26.1	-7.3	-1.6	-2.9	-4.2	45.6	20.5

Table 8

Probability that the $m + 1^{th}$ Exogenous Variable is a Red Herring (n = 13)

		High Variances					**Low Variances**				
m	BL	σ_a^2	σ_b^2	σ_c^2	σ_d^2	σ_u^2	σ_a^2	σ_b^2	σ_c^2	σ_d^2	σ_u^2
1	0.65	0.54	0.73	0.67	0.76	0.60	0.74	0.53	0.64	0.72	0.74
4	0.51	0.36	0.80	0.56	0.84	0.48	0.82	0.32	0.55	0.84	0.81
7	0.45	0.27	0.88	0.56	0.88	0.50	0.87	0.22	0.45	0.88	0.88
10	0.49	0.32	0.92	0.51	0.95	0.57	0.92	0.23	0.50	0.93	0.91
13	0.47	0.37	0.95	0.57	0.97	0.66	0.95	0.24	0.46	0.96	0.95
16	0.58	0.54	0.94	0.65	0.99	0.72	0.96	0.33	0.53	0.96	0.99
19	0.64	0.61	0.98	0.76	0.99	0.76	0.98	0.41	0.59	0.97	1.00
22	0.66	0.79	0.99	0.84	1.00	0.80	0.98	0.48	0.62	0.99	1.00
25	0.74	0.86	1.00	0.92	1.00	0.85	0.99	0.61	0.73	0.99	1.00
28	0.84	0.93	1.00	0.96	1.00	0.88	0.99	0.70	0.80	0.99	1.00
All	0.60	0.56	0.92	0.70	0.94	0.68	0.92	0.41	0.59	0.92	0.93

Table 9

Expected Welfare Loss When the $m + 1^{th}$ Exogenous Variable is Revealed (n = 13)

Relative to Structural Coefficients Expectations

		High Variances					**Low Variances**				
n	BL	σ_a^2	σ_b^2	σ_c^2	σ_d^2	σ_u^2	σ_a^2	σ_b^2	σ_c^2	σ_d^2	σ_u^2
1	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.0
4	0.0	-0.1	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.0
7	0.0	-0.2	0.2	0.0	0.1	0.0	0.0	0.0	0.0	102.0	0.0
10	0.0	1.5	0.3	0.1	0.1	0.2	0.0	0.0	0.0	1662.0	0.0
13	0.0	2.4	0.4	0.1	0.1	0.2	0.0	0.0	0.0	0.8	0.0
16	0.0	8.3	0.5	0.8	0.2	0.3	0.0	0.0	0.0	0.8	0.0
19	0.0	29.6	0.6	1.5	0.2	0.5	0.0	0.0	0.0	12215.9	0.0
22	0.1	91.4	0.9	3.3	0.2	1.1	0.0	0.0	0.0	2.5	0.0
25	0.1	90.3	6.6	7.6	0.2	1.7	0.0	0.1	0.0	3.8	0.0
28	0.4	195.4	3.3	108.8	0.2	4.1	0.0	0.1	0.0	9504.7	0.0
All	0.1	41.9	1.3	12.2	0.1	0.8	0.0	0.0	0.0	2349.4	0.0

Table 10

Probability that the 15^{th} Endogenous Variable is a Red Herring

m	BL	high β	low β	high γ	low γ
2	0.38	0.43	0.42	0.45	0.31
5	0.40	0.43	0.41	0.44	0.33
8	0.37	0.46	0.46	0.41	0.34
11	0.37	0.48	0.43	0.41	0.36
14	0.44	0.44	0.43	0.38	0.32
17	0.38	0.39	0.45	0.40	0.34
20	0.38	0.46	0.41	0.42	0.38
23	0.39	0.47	0.45	0.38	0.36
26	0.43	0.48	0.43	0.37	0.35
29	0.40	0.47	0.37	0.40	0.36
All	0.40	0.45	0.43	0.41	0.34

Table 11

Probability that the $m + 1^{th}$ Endogenous Variable is a Red Herring (n = 13)

m	BL	high β	low β	high γ	low γ
1	0.65	0.80	0.45	0.77	0.46
4	0.51	0.59	0.43	0.74	0.42
7	0.45	0.51	0.43	0.77	0.36
10	0.49	0.44	0.41	0.86	0.30
13	0.47	0.49	0.40	0.93	0.25
16	0.58	0.56	0.47	0.96	0.20
19	0.64	0.65	0.50	0.97	0.21
22	0.66	0.75	0.49	0.99	0.16
25	0.74	0.88	0.56	0.99	0.17
28	0.84	0.93	0.56	0.99	0.12
All	0.60	0.66	0.47	0.90	0.26

Table 12

Red Herrings in the New Keynesian Model

K	$p(RH)$	$p(\psi > 3)$	$p(\psi < \frac{1}{3})$
3	.850	.038	.000
6	.813	.016	.001
9	.778	.009	.000
12	.753	.015	.001
15	.738	.007	.000
18	.743	.012	.000
21	.718	.012	.000
24	.693	.013	.001
27	.705	.014	.001
30	.688	.018	.000