Institute for New Economic Thinking

Working Papers
ineteconomics.org/workingpapers

Stock-Market Expectations: Econometric Evidence that *both* REH and Behavioral Insights Matter

Roman Frydman[1] and Joshua R. Stillwagon[2]

Working Paper No. 44

May 19, 2016

Abstract

Behavioral finance views stock-market investors' expectations as largely unrelated to fundamental factors. Relying on survey data, this paper presents econometric evidence that fundamentals are a major driver of investors' expectations. Although expectations are also in part extrapolative, this effect is transient. The paper's approach underscores the central importance of opening models to structural change and imposing discipline on econometric analysis through specification testing. Our findings support the novel hypothesis that rational market participants, faced with unforeseeable change, base their forecasts on both fundamentals - the focus of the REH approach - *and* the psychological and technical considerations underlying behavioral finance.

[1] Department of Economics, New York University, and Institute for New Economic Thinking (INET), Program on Imperfect Knowledge Economics (IKE), rf3@nyu.edu

[2] Department of Economics, Trinity College, and INET Program on IKE, joshua.stillwagon@trincoll.edu

Keywords: Behavioral finance, REH, Knightian uncertainty, survey expectations, structural change, model specification, automated model selection.

JEL Codes: G12, G14, G02, C22

The authors are grateful to the Institute for New Economic Thinking and Trinity College for support of this research. Earlier versions of this paper were presented at the University of Copenhagen and the 17th Oxmetrics Conference at George Washington University. The authors thank Neil Ericsson, David Hendry, Soren Johansen, Felix Pretis, Anders Rahbek, Peter Sullivan, and Morten Nyboe Tabor for insightful comments and suggestions that led to substantial improvements of the paper.

1 Introduction

The Rational Expectations Hypothesis (REH) relates market participants' expectations to fundamental factors (such as company earnings and macroeconomic variables). In a pathbreaking paper, Shiller (1981) presented evidence that the REH-based present-value model is grossly inconsistent with persistent swings in stock prices. He interpreted his findings as evidence that market participants' expectations, which drive these movements, are largely unrelated to fundamental factors. This interpretation provided the *raison d'être* of the behavioral-finance approach, which hypothesized that participants' expectations, and thus stock-price swings, are driven by psychological and technical considerations. The commonly invoked examples of such considerations are market sentiment (optimism or pessimism) and bandwagon effects (participants' mechanical extrapolation of past returns into the future).[1]

Frydman and Goldberg (2011, 2013a,b) advanced an alternative interpretation of Shiller's findings: REH does *not* represent how rational, profit-seeking participants forecast prices in asset markets. The reason is simple: By design, REH models are completely closed to unforeseeable change in the process underpinning outcomes.[2] Frydman and Goldberg hypothesized that faced with such change, *rational* participants would base their forecasts on fundamentals – the focus of the REH approach – as well as draw on psychological and technical considerations. This hypothesis, particularly the central role that it accords to structural change in modeling investors' expectations, guides an econometric analysis in this paper.

The main contribution of this paper is to present econometric evidence that trends in fundamentals are a major driver of investors' expectations. Investors' expectations are also in part extrapolative. However, the effect of extrapolation is short-lived, largely reversing itself after one month. The paper also finds that the effects of both fundamentals and extrapolation vary

[1] For extensive surveys of the behavioral-finance approach see Shleifer (2000), Barberis and Thaler (2003), and references therein.

[2] REH models assume away unforeseeable change by presuming that all changes in the process undepinning outcomes can be represented *ex ante* with a probabilistic rule.

in magnitude over time.

The paper's approach underscores the key importance of opening models to structural change and imposing discipline on econometric analysis by requiring that empirical models be well specified, in the sense of passing a battery of standard specification error tests.

The paper joins a growing literature relying on survey data of investors' expectations to understand prices and risk premiums in asset markets.[3] Prior to the use of these data, researchers relied on the indirect implications of alternative theoretical representations of expectations for asset-price movements. For REH, this typically involved imposing consistency within a specific model and testing its predictions for the quantitative co-movements between prices and fundamental factors, rather than investigating directly whether these factors drive investors' expectations.

As is well known, the "joint hypothesis" problem makes it very difficult to ascertain whether the failure of such studies to detect the role of fundamentals in asset-price movements arises from the invalid specification of the market's expectation or the wrong model of equilibrium returns. This observation has enabled proponents of REH to maintain that a better risk-premium model could overturn failures of the hypothesis in asset markets.[4]

The availability of survey data has made it possible to investigate the empirical relevance of REH and behavioral approaches directly. However, given sensitivity to question framing and interpretation, the evidence from survey data has been considered unreliable. It is argued that the surveys are either too noisy or unrepresentative to be useful, or that respondents are misinterpreting the question (Cochrane, 2011).

Greenwood and Shleifer (2014) convincingly argue against this dismissal of the stock market survey data. Their paper shows that various measures of expected returns (seven different sources of survey data in total) are highly correlated with one another, suggesting that they are not merely uninforma-

[3]For recent studies using stock market survey data and references to earlier literature, see Williams (2013) and Greenwood and Shleifer (2014).

[4]Although the search for such a risk-premium model has not been successful, many (notably Cochrane, 2011) continue the quest.

tive noise. Furthermore, they are highly correlated with mutual fund flows, demonstrating that they are representative of expectations that are relevant for market participants' decisions. These findings vindicate the use of survey expectations.

Greenwood and Shleifer (GS) estimate a relationship between investors' expectations, fundamentals and extrapolation. For each of their seven survey measures, they present regression results supporting the behavioral-finance hypothesis that these expectations are almost purely extrapolative and largely unrelated to fundamentals.

Our analysis uses the longest available sample of GS's survey measures, spanning the period from 1963 to 2015. In contrast to GS, however, our econometric analysis yields a relationship that accords a major role to fundamentals and a transient role to extrapolation in driving investors' expectations.

We reach this very different conclusion by adhering to a key methodological principle: for an estimated relationship to serve as the basis for assessing empirical relevance of alternative theoretical approaches it should be well specified, in the sense of passing standard tests of specification error.[5]

We show that GS's econometric model is strongly rejected by each of the standard specification error tests. We trace this misspecification to two main sources: non-stationarity of regressors and the model's time-invariant structure, which presumes that the same set of variables, with unchanging parameter values, can represent how investors form expectations at every point in time.

Both of these shortcomings have detrimental effects for the validity of inference in the GS model. For example, the test for no autocorrelation is rejected with a $p-$value of 0.0000. Bauer and Hamilton (2015) have shown that serially correlated errors have led to erroneous deductions in the expectations hypothesis literature, where well-known findings of the predictability of bond returns from factors outside of the yield curve have proved not to be robust.

[5] For arguments concerning the key importance in econometric analysis of achieving well-specified models, see Campos, Ericsson, and Hendry (2005), Johansen (1995), and Juselius (2006).

They trace this problem to the inclusion of highly persistent variables in the estimated model. Using the Augmented Dickey-Fuller test, we show that a number of variables used by GS are highly persistent.

Whereas estimation and inference problems stemming from non-stationarity are well recognized, detrimental effects of assuming away structural change have been largely overlooked. Indeed, constraining an unstable model to be time-invariant renders its error term autocorrelated, heteroskedastic, and correlated with the regressors.[6]

There are good reasons to surmise that more than one structure would be required to represent investors' expectations during any sufficiently long sample period.[7] Yet a vast majority of studies of expectations estimate a regression model that presumes that the same structure can explain how investors form expectations at each point in time.

Our econometric approach addresses both of the foregoing shortcomings. In order to avoid misspecification and unreliable inference arising from non-stationarity of variables, we use first differences for all such variables. Moreover, we place structural change at the center of our analysis.

The first stage of our investigation uses the Autometrics tree-search algorithm. Automated model selection has advanced dramatically over the last decade and a half, owing to much-improved algorithms, beginning, for example, with the multi-path search of Hoover and Perez (1999). The properties of the Autometrics procedure used here, building off of Hendry and Krolzig (2001, 2005), have been demonstrated by Doornik (2009) to overcome the previously documented biases of step-wise regression. These procedures rely on the general-to-specific methodology, whereby all potential variables are in-

[6]See Tabor (2013) for an econometric analysis of autocorrelation and the ARCH effects arising from assuming away structural change.

[7]The importance of structural change in an REH context has been emphasized by Lucas (1976) and Hamilton (1988, 1994). Frydman and Goldberg (2007, 2011, 2013a,b) analyze the theoretical and empirical implications of unforeseeable structural change in the foreign exchange and equity markets. For an overview of various approaches to modeling changes in expectations, see Frydman and Phelps (2013).

cluded from the outset.[8]

Autometrics is well suited to econometric modeling of investors' expectations for a number of reasons. First, it provides a disciplined way to select an empirical model involving a subset of variables chosen from a large potential set of regressors. This is crucial in modeling investors' expectations, because existing theories provide no guidance concerning specific factors that might drive these expectations. Predictions yielded by the REH and behavioral approaches concern broad sets of factors – fundamental and behavioral, respectively – that might be empirically relevant.

Thus, in order to examine whether fundamental and/or behavioral considerations drive expectations, an investigator must examine a variety of potential specifications involving different subsets of some large set of candidate regressors. Autometrics does so in a disciplined way by requiring that the selected model pass a battery of specification error tests.

Second, Autometrics adjusts parameter-estimates and test statistics for the model selection bias arising from repeated reestimation using the same set of variables.[9]

Third, the procedure provides a way to diagnose the importance of allowing for structural change in achieving a well-specified model. To this end, Autometrics uses step indicator dummies to test for potential shifts in the constant term.(Castle *et al.* 2015).

As expected, we find that the specification selected by Autometrics undergoes such shifts. However, we would also expect that the estimated shifts in the constant term reflect, at least in part, shifts in the parameters attached to the model's variables.

Consequently, the second step of our approach to model selection allows for structural change in the constant term and the regressors' coefficients. This step involves testing for structural change in the specification that relates investor expectations to the set of variables selected by Autometrics. We then

[8]For a comprehensive overview of arguments in favor of this methodology and automatic model selection, see Hendry and Doornik (2014).

[9]Although GS present and reference a number of alternative specifications, they do not discuss whether they adjusted their estimates for the model-selection bias.

estimate separately a model for expectations involving these variables within each subperiod of statistical constancy, as judged by the structural-change test. The resulting piece-wise linear model is considered well specified if each of its linear segments passes a battery of standard specification error tests.

Both Autometrics and the structural-change tests show that in order to achieve a well-specified model, we must allow its structure to change over time. This finding is inconsistent with the vast majority of existing models, which, regardless of whether they are based on REH or behavioral considerations, attempt to approximate investors' expectations with time-invariant structures.

Although any single structure eventually fails specification error tests, there may be protracted periods of time during which investors' expectations can be approximated with linear segments. Our empirical model for expectations involves three linear relationships, *each* of which is well specified.

The empirical relevance of structural change implies that the model's quantitative predictions vary across the linear segments. However, the model generates qualitative predictions that enable us to assess the empirical relevance of alternative theoretical approaches to modeling investors' expectations.[10]

The estimates of the model that passes our rather stringent selection process indicate that the trend of at least one fundamental variable – the rate of interest and/or unemployment – is a major driver of investors' expectations in *every* subperiod of approximate parameter constancy. In contrast, extrapolation plays a transient role in every linear segment.

We also find that the composition of the variables accounting for investors' expectations differs across subperiods.[11] Whereas both the interest rate and the unemployment rate drive expectations in one of the subperiods, only one

[10]Generating quantitative predictions that span more than one segment requires further restrictions on change, for example, that change between segments is governed by a Markov switching rule. For a demonstration, see Hamilton (1994) and Frydman and Goldberg (2007). However, Stillwagon and Sullivan (2016) and Frydman *et al* (2016) show that, although a Markov switching rule might provide an *ex post* approximation of the process during a sample period, this empirical characterization eventually fails to represent structural change in future periods.

[11]This result of our econometric analysis corroborates descriptive evidence provided by Frydman *et al.* (2015), who find that the number and composition of fundamental variables driving stock-market participants' expectations vary over time.

of them matters in the other two.

Moreover, the estimated qualitative effects (signs of parameters) of fundamental variables appear to provide a sensible explanation of these movements during the subperiods approximated by each of the linear segments. According to the model, from 1963 to 1980, investors' expectations appear to have been driven by both the interest rate and business-cycle effects. The bull market from the 1980s to 1999, however, appears to have been driven primarily by falling interest rates, while post-1999 expectations again focused on the macroeconomic outlook, as proxied by changes in unemployment.

The paper is organized as follows: Section 2 shows that the model that served as the basis for GS's conclusion concerning the unimportance of fundamentals is grossly misspecified, as judged by the battery of standard tests. This section also uses the Augmented Dickey-Fuller test to show that a number of key regressors used by GS are non-stationary. Section 3 sketches our approach to model selection, which relies on Autometrics, structural change, and specification error tests in searching for a well-specified model. Sections 4 and 5 formulate an unrestricted model, apply Autometrics, and correct the estimated t-ratios and estimates of the parameters for the selection bias. Section 6 carries out the structural-change step of our approach to model selection. Section 7 discusses the results of our econometric investigation, based on a well-specified, piece-wise linear model selected by our approach.

Finally, section 8 places the paper's findings in a broader context. It sketches how opening models to Knightian uncertainty is the key to incorporating both REH and behavioral insights into representations of rational forecasting. Although opening models to unforeseeable change poses considerable challenges for both model-building and econometric methodology, the paper concludes that overcoming these challenges is one of the important objectives of macroeconomics and finance research.

2 Survey Data

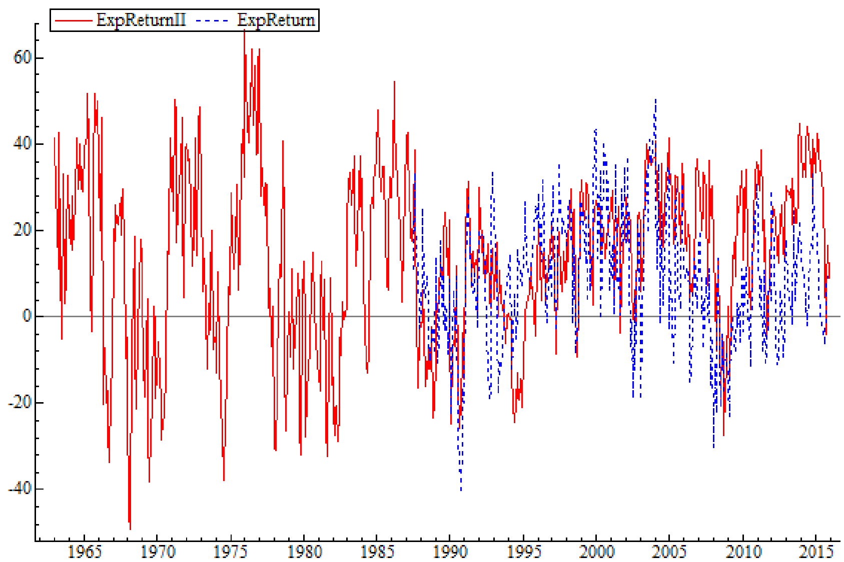
Greenwood and Shleifer use seven measures summarizing surveys of investors' expectations. An important contribution of their paper is to show that these measures are highly correlated with market participants' decisions to invest their capital in mutual funds. Furthermore, GS show that the measures co-move strongly and positively, even though the surveys that underpin them rely on very different methodologies. This evidence buttresses their argument that survey measures are not just "meaningless noise" (p. 715).

We illustrate the correlation between alternative survey proxies with the two longest available measures: one summarizing the survey by the Investors Intelligence Newsletter (II), and the other based on the survey by the American Association of Individual Investors (AA). These surveys record the percentage of their participants' bullish, neutral, and bearish forecasts on a weekly basis. Given that most of the other variables are measured at monthly intervals, we use a monthly average. Moreover, following GS, we proxy investors' time-t expectation of "raw" stock returns (stock-price change) over the succeeding 12-month period, $t+12$, with the difference between the proportion of investors who are bullish and bearish at t concerning stock prices at $t+12$:

$$Exp_{t|t+12} = [\%\ bullish_t - \%\ bearish_t] \tag{1}$$

Measures computed according to (1) are not numerical observations of price changes expected by survey participants. However, GS show that these proxies are highly correlated with the shorter available sample from Gallup surveys which provide numerical forecasts of stock returns from September 1998 through May 2003. Figure 1 shows a close co-movement of the II and AA measures.

Figure 1: II and AA Survey Measures



Caption: The solid (broken) series depict the II and AA measures, respectively.

3 GS's Model-Selection Approach: A Critical Assessment

Greenwood and Shleifer estimate a number of alternative regression models that relate each of their seven survey measures to a proxy for extrapolation and a set of fundamental variables. They pick their preferred specification by selecting the subset of the variables that are consistently estimated irrespective of the survey measure used to proxy expectations.

Fundamentals are often statistically insignificant in the regressions estimated by GS. Moreover, whenever they are significant, variables that seem to matter have the wrong sign or differ across estimated specifications. By contrast, the proxies for extrapolative expectations are statistically significant and have the correct sign in specifications using different survey measures.

Based on these results, GS pick as their preferred model a specification that accords no role to fundamentals and represents stock-market expectations as purely speculative. They conclude that these expectations "are well explained by two variables. First, when recent past returns are high, investors expect higher returns going forward. Second,...investor expectations are positively correlated with the price dividend ratio" (p. 729).

However, reliance on standard t-ratios and "correct" signs of the effects of the extrapolative variables across all specifications are far from sufficient to support GS's conclusion that investor expectations "are well explained" by these two variables. The reason stems from three major shortcomings of GS's approach: model selection bias arising from repeatedly searching among alternative specifications involving the same set of variables, non-stationarity of the regressors, and misspecification arising from assuming away structural change.

Although GS display some alternative specifications that they estimated, and reference many others, they do not elaborate on precisely how they arrived at the final model or whether they adjusted their estimates for the model selection bias. As our analysis in section 5 shows, this bias can be quite large and correcting for it alters both the estimates and test statistics substantially.

As we discussed in the introduction, non-stationarity of regressors and constraining an unstable model to be time-invariant may result in serious misspecification that would render both estimates and test statistics unreliable. In order to illustrate how misspecification affects the error term of GS's empirical models for investors' expectations, we replicate GS's regressions for the II and AA measures and subject them to five standard tests for specification error.

3.1 GS's Regressions

Greenwood and Shleifer estimate a model of investor expectations by relating each survey measure to variables that proxy extrapolation and a set of fundamental variables. Their set of extrapolative variables consists of the percentage change in the S&P 500 over the last year (the measure of past returns R_{t-12}) and $\ln(P_t/D_t)$ – the log of the price dividend ratio.[12] Fundamental variables used in the results presented by GS include: unemployment u_t, the one-year Treasury rate i_t, and the growth rate of earnings $\Delta\ln(E_t)$.

Table 1 displays results for the four specifications estimated by GS. In GS Models II1 and II2, the dependent variable is the proxy for the expected rate of return based on the II surveys. In GS models AA1 and AA2 the dependent variable is the AA proxy.

Table 1

Estimates of GS specifications

	GS Model II1	GS Model II2	GS Model AA1	GS Model AA2
C	-37.891 [-3.35]	2.106 [0.10]	-63.9351 [-4.58]	-77.785 [-2.58]
R_{t-12}	0.510 [7.00]	0.545 [8.30]	0.290 [5.38]	0.267 [4.44]
$\ln(P_t/D_t)$	13.242 [4.39]	2.958 [0.66]	18.076 [5.05]	20.517 [3.96]
i_t		-2.109 [-4.69]		0.655 [0.84]
u_t		1.328 [1.44]		0.357 [0.24]
$\Delta\ln(E_t)$		1.516 [0.10]		0.796 [0.07]

Caption: The dependent variable is [%bulls-%bears]. Sample from 1963:01-2015:06 for II and 1987:07-2015:06 for AA. Newey-West t-values in brackets.

We use somewhat different data sources for the regressors and our sample ends about 3.5 years later. Nevertheless, the estimates in Table 1 are quite similar to those reported by GS in their Table 3 (p. 730).[13]

As GS observe, extrapolation is quite evident. The effect of the past return is highly significant in all models with and without fundamentals, but

[12] GS sketch why they consider the price-dividend ratio an extrapolative variable. This ratio is "a measure of the price-level" (p.729). which is "essentially the sum of past returns" (p.731). However, as we discuss in section 5 the price-dividend ratio involves the role of both extrapolation and fundamentals in stock-market expectations.

[13] We use the Shiller data for the S&P 500, earnings, and dividends. Industrial production, the one-year Treasury rate, the U-3 unemployment rate, and personal consumption expenditure are from the FRED database.

the $\ln(P/D)$ becomes insignificant in the longer sample when fundamentals are included. By contrast, fundamentals do not seem to matter, let alone consistently, across models II2 and AA2. Earnings growth and unemployment are both insignificant, and the latter has the wrong sign. The interest rate is significant and has the correct (negative) sign in model II2. However, once the AA measure is used as a proxy for the expected return in model AA2, the interest rate switches sign to positive and loses significance completely. Based on such results across all seven measures, GS conclude that investor expectations are purely extrapolative.

3.2 Specification Error Tests

GS's conclusion that empirical evidence is inconsistent with fundamentals-based account of investors' expectations rests on the adequacy of their estimated regression relationships and their test statistics. In order to assess this adequacy we subject the regressions in Table 1 to a battery of standard specification tests.

The diagnostics in Table 2 include the Lagrange multiplier test of serial correlation, labeled as AR (Godfrey 1978), autoregressive heteroskedasticity or ARCH (Engle 1982), normality (Doornik and Hansen 1994), heteroskedasticity (White 1980) and the RESET test of model misspecification (Ramsey 1969).

Table 2:
Specification Tests

	GS Model II1	GS Model II2	GS Model AA1	GS Model AA2
AR	0.0000	0.0000	0.0000	0.0000
ARCH	0.0000	0.0000	0.0000	0.0000
Normality	0.6604	0.7263	0.2094	0.3442
Hetero	0.0000	0.0000	0.0098	0.0215
RESET	0.0000	0.0000	0.8502	0.8548

Caption: The figures represent the p-values for the respective tests and models.

These results indicate that regressions II and AA in Table 1 are grossly misspecified. Their errors are strongly autocorrelated, are heteroskedastic and suffer from ARCH effects.

GS rely on the Newey and West (1987) approach to correct the standard t-ratios for autocorrelation and heteroskedasticity with Heteroskedasticity and Autocorrelation Consistent (HAC) standard errors. However, Bauer and Hamilton (2015) show that reliance on HAC-corrected t-ratios does not adequately address the bias of standard errors in models that include highly persistent or non-stationary variables. Bauer and Hamilton argue that this bias is substantial and has led to erroneous conclusions concerning factors driving bond premia. For example, they show that "..the tests employed by Ludvigson and Ng (2009), which are intended to have a normal size of five percent, can have a true size of up to 54%" (pg. 3).

Spanos and Reade (2016), using simulations, similarly conclude that HAC standard errors suffer from significant size and power distortions even under best-case scenarios.

We examine stationarity of all variables in GS regressions with the standard Augmented Dickey-Fuller (ADF) test (Said and Dickey 1984). As reported in Table A1 in the appendix, the hypothesis of a unit root was not rejected for one of the extrapolative variables, $\ln(P_t/D_t)$, and two of the fundamental variables, unemployment u_t, and the one-year Treasury rate i_t.

These results enable us to make general statements concerning the adequacy of all specifications estimated by GS – that is, regardless of the particular proxy used for investor expectations. Given that all of these models include non-stationary variables, the Bauer and Hamilton analysis implies that HAC-adjusted t-ratios are likely to lead to unreliable inference concerning the determinants of investor expectations.

Beyond the detrimental effect of including non-stationary regressors on inference, GS's regressions suffer from another key shortcoming. Like a vast majority of existing models in macroeconomics and finance, they constrain the specification of investor expectations to remain unchanging over time.

On theoretical and empirical grounds, time-invariant specifications are likely to provide a grossly inadequate representation of participants' expectations in real-world markets. As we show in section 6.2, once we replace non-stationary regressors with their stationary first differences and allow for structural change, the specification of investor expectations improves markedly. The results of misspecification tests, such as those in Table 2, turn from strongly significant to statistically insignificant.

4 Our Approach to Model Selection

Our approach to modeling the role of extrapolation and fundamentals in investor expectations attempts to remedy three of the main econometric shortcomings in GS's analysis: non-stationarity of the regressors, selection bias arising from repeatedly searching among alternative specifications involving the same set of variables, and misspecification as a result of assuming away structural change.

We avoid the first shortcoming by using only first differences of all of the variables that have been found non-stationary with the ADF test. Second, we rely on Autometrics to select from myriad potential alternative specifications and correct for the selection and other biases inherent in the search for well-specified models.

Automentrics is also helpful in mitigating the detrimental effect on model specification of ignoring structural change. Autometrics allows for shifts in the constant term of the estimated model, though it constrains other parameters to remain unchanging over time. We show that by controlling for such shifts, Autometrics improves the specification of the model relative to its time-invariant counterpart. However, attempts to force Autometrics to select the model that passes standard tests (such as those in Table 2) at higher significance levels (5%) results in the proliferation of shifts in the constant term. This suggests that allowing for structural change in all of the parameters, rather than just in the constant term, may result in a model that is well specified at 5% significance *and* undergoes relatively few structural changes. Achieving well-specified models is particularly important if we rely on Autometrics to help us test the empirical relevance of alternative theoretical explanations. As this is our objective here, we set the significance level at the customary 1%. This leaves some misspecification in the model. However, we rely on Autometrics to suggest the set of variables that might be relevant in modeling investor expectations.[14]

In view of these considerations, the second step of our approach allows for structural change in both the constant term and other model parameters, and we consider the model well specified if it passes the standard battery of specification tests at the 5% level. This step involves testing for structural change in the specification that relates investor expectations to the set of variables selected by Autometrics. We then estimate separately a model for expectations involving these variables within each subperiod of statistical constancy, as judged by the structural-change test.

4.1 Addressing Selection Bias: Autometrics

Autometrics relies on the general-to-specific methodology, whereby all potential variables are included from the outset. The properties of the procedure used here have been demonstrated by Campos, Ericsson, and Hendry (2005) and Doornik (2009) to avoid omitted-variable and other biases of step-wise regression.

In order to allow for structural change and control for outliers, Autometrics uses an impulse indicator saturation (IIS) and a step indicator saturation (SIS) procedure (see Hendry, Johansen, and Santos 2008 for IIS, and Castle *et al.* 2015 for SIS). This allows for an impulse and step indicator for each observation, conducted in block searches or sub-samples, to permit feasible estimation (given that the number of included variables automatically exceeds the number of observations). The indicators are included simultaneously with all potential regressors to mitigate the possibility of incorrect inclusion or exclusion of either.

[14] See Ericsson (2012) for the use of Autometrics and indicator saturation as a diagnostic tool.

4.2 General Unrestricted Model

We start our model selection with a general unrestricted model that includes all candidate regressors used by Greenwood and Shleifer. However, in order to avoid misspecification and inference problems stemming from non-stationarity, we include only first differences among the candidate regressors.[15] Furthermore, two lags are included for each regressor and the dependent variable to address potential issues of serial correlation. Our initial unrestricted specification can be written as follows

$$Exp_{t|t+12} = c + \Sigma_{j=1}^{2} Exp_{t-j|t+12-j} + \Sigma_{i=1}^{8} \Sigma_{j=0}^{2} \beta_{i,j} X_{i,t-j} + \varepsilon_t \qquad (2)$$

where ($c, \beta_{i,j}$ for $i = 1...8$, and $j = 0, 1, 2$) is a vector of parameters, $Exp_{t|t+12}$ denotes the proxy for the expected rate of return, and $X_i = [R_{t|t-12}, \Delta \ln(P_t/D_t), \Delta u_t, \Delta i_t, \Delta \ln(E_t), \Delta \ln(C_t), \Delta \ln(D_t), \Delta \ln(Y_t)]$ is a vector of variables including, respectively, the return over the past year to the S&P 500, the differences in the log price dividend ratio, unemployment, the one-year Treasury bill rate, earnings, consumption, dividends, and industrial production.

5 Results from Autometrics

In order to facilitate the comparison with GS's results, we have applied Autometrics to the model including $\Delta \ln(P_t/D_t)$. The second and third columns in Table 3 below (labeled Model A1) present the results for this specification.

[15] We already mentioned that the ADF test rejects stationarity of the interest rate and unemployment. This test also finds non-stationarity of other variables used by GS: consumption and industrial production have a unit root, and dividends and earnings have a linear trend.

Table 3

Autometrics Results for the II Survey Measure.

	Model A1		Model A2	
	Coeff.	Bias-adj.	Coeff.	Bias-adj.
$Exp_{t-1\|t+11}$	0.729 [32.66]	0.729 [32.66]	0.810 [37.84]	0.810 [37.84]
$R_{t\|t-12}$	0.021 [0.79]	0.000 [0.00]	0.946 [13.67]	0.946 [13.67]
$R_{t-1\|t-13}$			−0.888 [−12.74]	−0.888 [−12.74]
$\Delta\ln(P_t/D_t)$	198.444 [21.73]	198.444 [21.73]		
$\Delta\ln(i_t)$	−2.414 [−3.41]	−2.279 [−3.22]	−4.259 [−5.11]	−4.257 [−5.11]
$\Delta\ln(u_{t-2})$	−2.779 [−1.40]	0.000 [0.00]	−3.809 [−1.75]	0.000 [0.00]
$\Delta\ln(D_{t-2})$	146.281 [2.53]	103.450 [1.79]		
$\Delta\ln(P_{t-2}/D_{t-2})$	9.640 [0.99]	0.000 [0.00]		
$\Delta\ln(Y_{t-2})$	−107.072 [−2.21]	−55.379 [−1.14]		
$\Delta\ln(E_{t-2})$	8.933 [1.27]	0.000 [0.00]		
$\Delta\ln(E_t)$	−14.959 [−2.08]	−6.230 [0.87]		
c	4.791 [7.11]	4.791 [7.11]		

Caption: The Coeff. columns show the coefficient, with $t-$values beneath in brackets, while the bias adjusted column shows them after the bias-correction procedure from Hendry and Krolzig (2005).

The inclusion of the price-dividend ratio in the unrestricted model may make it difficult to ascertain whether fundamentals play a role in driving investors' expectations. This variable captures both the effect of extrapolation and fundamentals through their effect on both the change in the stock price and dividends. Thus, we have also applied Autometrics to the model that initially excludes $\Delta\ln(P_t/D_t)$. These results are presented in the fourth and fifth columns of Table 3 (labeled Model A2).

5.1 Bias Correction

The estimates reported in the "Coeff" and "bias-adjusted" columns in the table make clear the importance of the correction for the selection bias. Although some of the variables remain significant after the correction, the correction reduces the value of others to zero.

These results illustrate the argument in Lovell (1982) that, conditional on being retained, the estimates resulting from step-wise regression procedures are biased away from zero. The Hendry and Krolzig (2001, 2005) procedure adjusts the estimates toward zero. The degree of the correction depends on the estimated t-value and the significance level set for the selection procedure. Greater correction is applied to the less significant variables.

5.2 Specification Tests

By design, Autometrics improves model specification. As reported in Table 4, the autocorrelation and heteroskedasticity are no longer significant in Model 1 at the 1% level, and the ARCH effects have been dramatically reduced in both.

Table 4

Specification Tests after Autometrics

	Model 1	Model 2
AR	0.0270	0.0088
ARCH	0.5484	0.3065
Normality	0.5107	0.6291
Hetero	0.0101	0.0044
RESET	0.0008	0.0684

Caption: The figures represent the p-values for the respective tests and models.

The diagnostics still leave something to be desired, however, if aiming to use the model to reliably infer the validity of alternative economic theories. Attempts to force Autometrics to deliver a better-specified model (for example, with a threshold of 5% for the specification tests) generates a proliferation of shifts in the constant term.

As we show next, this difficulty stems from constraining structural change to occur solely in the constant term. Once we allow for shifts in all other parameters, we achieve a well-specified model that captures the structural change more parsimoniously, with fewer breaks.

6 Model Selection: Structural Change

In the second step of our model-selection approach, we allow for shifts in the parameters of all variables in the specifications selected by Autometrics. We consider two models, Model 1 and Model 2, which arise from Model A1 and Model A2 in Table 3, respectively. In each model, we use the regressor variables that have been retained by Autometrics as significant after the bias correction. However, in order to examine the role of fundamentals other than the interest rate, we also add unemployment to Model 2, which was the only other fundamental retained in that model before bias correction.

6.1 Misspecification Arising from Structural Change

Autometrics estimates a set of step indicators that capture structural change with shifts in the constant term. Table A3 in the Appendix presents these indicator estimates for Models A1 and A2 of Table 3. The number of retained step indicators provides a measure of structural instability. By this measure, Model A1 appears to be very unstable: Autometrics required 20 step indicators and more than 20 impulse indicators to pass specification tests at the 1% level.

Although allowing for only the shifts in the constant term is quite restrictive, it is crucial to Autometrics' ability to mitigate misspecification. In order to illustrate this point, we run the specifications suggested by Autometrics without the indicator dummies. The results are displayed in Tables 5 and 6.

Table 5

Specifications Suggested by Autometrics without IIS and SIS

	Model 1	Model 2
$Exp_{t-1\|t+11}$	0.802 [44.13]	0.769 [19.52]
$\Delta \ln(P_t/D_t)$	202.544 [23.05]	
$R_{t\|t-12}$		0.997 [10.10]
$R_{t-1\|t-13}$		−0.850 [−8.30]
Δi_t	−3.137 [−3.80]	−3.057 [−3.25]
$\Delta \ln(D_{t-2})$	156.919 [2.48]	
Δu_{t-2}		0.748 [0.24]
c	1.811 [3.46]	0.802 [44.13]

Caption: t-values are presented in brackets underneath the coefficients

Table 6

Specification Tests for Models Ignoring Structural Change

	Model 1	Model 2
AR	0.0012	0.0012
ARCH	0.0000	0.0177
Normality	0.0000	0.0000
Hetero	0.0035	0.1327
RESET	0.0000	0.0015

Caption: The figures represent the p-values for the respective tests and models.

These diagnostics make clear that constraining the models' structures to be unchanging over time results in a grossly misspecified model. The comparison between Tables 4 and 6 shows that even allowing for shifts only in the constant term improves the model specification substantially. However, as is evident from Table A3 in the Appendix, constraining other parameters of the model to be unchanging over time forces Autometrics to retain a large number of breaks in the constant term, especially in Model A1. Moreover, it seems difficult to improve model specification, as judged by the standard battery of tests, without having to retain even more of the step and outlier dummies.

We attempt to address this problem by allowing for structural change in all of the parameters. To this end, we subject Models 1 and 2 in Table 5 to the test devised by Bai and Perron (1998).[16] We then estimate the piece-wise linear specifications with the estimated timing of structural breaks between the linear segments of approximate parameter constancy, as judged by the structural-change test.

Table 7 presents the results of the Bai-Perron test. In contrast to the step indicators estimated by Autometrics, allowing for structural change in all of the model parameters results in only a few breaks.

Table 7

Intervals of Approximate Parameter Constancy from the Bai Perron Test

Model 1	63:03-73:11	73:12-99:10	99:11-15:06
Model 2	63:03-80:03	80:04-99:10	99:11-15:06

It is noteworthy that the timing of the structural breaks detected by the Bai-Perron test seems to coincide with major events, including the OPEC oil crisis of 1973, the Volcker disinflation and the bottom of the bear market in 1980, and the near peak of the IT bubble in 1999. Such historical events are at least in part unique. Thus, the Bai-Perron test indicates the empirical relevance of unforeseeable change.[17]

6.2 Structural Change Analysis as Model Selection

We rely on structural-change analysis as a key step in our approach to model selection. We consider a piece-wise linear model resulting from testing for

[16] Ericsson (2012) proposes an alternative to allow for changing betas over time, referred to as multiplicative indicator saturation (MIS), which uses regressors interacted with step indicators and conducts model selection with Autometrics. Kitov and Tabor (2016) investigate the properties of MIS, which nests the Bai Perron approach in that it allows for but does not force re-estimation of all parameters following a given structural break. This constraint may be desirable in some applications, however.

[17] Frydman *et al.* (2015) provide extensive evidence that events that are at least partly non-repetitive underpin movements in the US stock market: 20% of stock-price changes over the 20-year period spanning the 1990s and 2000s involved such events.

structural change to be well specified if *each* of its linear segments passes a battery of standard specification error tests.

Table 8

Specification Tests for Model 1

	Time-invariant Model	63:03-73:11	73:12-99:10	99:11-15:06
AR	0.0012	0.1132	0.0103	0.4324
ARCH	0.0000	0.7504	0.0000	0.6597
Normality	0.0000	0.6583	0.7583	0.4938
Hetero	0.0035	0.1506	0.0000	0.0129
RESET	0.0000	0.0155	0.0000	0.0140

Caption: The figures represent the p-values for the respective tests and models.

Table 8 presents the results for each linear segment resulting from testing for structural breaks in Model 1 (in Table 5). Although the model passes the specification tests for the first and third subperiods, it is misspecified during the middle subperiod from 73:12 to 99:10 and this can't be resolved simply through the inclusion of an extra lag for the dependent variable. We recall that Autometrics indicated that this model undergoes many structural breaks while the Bai-Perron test does not detect any breaks during the nearly 26-year, middle subperiod. It seems plausible that the model's poor performance during the middle subperiod stems at least in part from structural changes that the Bai-Perron test has not detected.

This misspecification could also arise from reliance on variables that do not adequately proxy the determinants of investor expectations. Indeed, we pointed out that $\Delta \ln(P_t/D_t)$ confounds the extrapolative and fundamental influences on investor expectations.

The failure of Model 1 to pass specification tests in all subperiods leaves Model 2 as the remaining candidate for a well-specified model of investor expectations.[18] This model arose from applying Autometrics to the unrestricted model that excluded the change in the log price-dividend ratio.

[18] For completeness, we present coefficient estimates for Model 1 in Table A2 in the Appendix.

In order to ensure reliable inference in each subperiod, we required that the AR test be passed at the 5% level. As the model fell short of this criterion in the third subperiod (from 99:11 to 15:06), we added an extra lag of the dependent variable to the model for this subperiod. This extra lag would be insignificant in the other subperiods and would negligibly alter the results. Table 9 presents the results of the specification tests.

Table 9

Specification Tests for Model 2

	Time-invariant Model	63:03-80:03	80:04-99:10	99:11-15:06
AR	0.0012	0.0971	0.0705	0.1935
ARCH	0.0177	0.9827	0.4577	0.0922
Normality	0.0000	0.0425	0.3598	0.6705
Hetero	0.1327	0.4978	0.1702	0.0996
RESET	0.0015	0.0191	0.3638	0.3914

Column 2 of the table repeats earlier results for the time-invariant version of Model 2. Clearly, constraining the model parameters to be unchanging over time results in gross misspecification. However, allowing for structural change provides a substantial remedy, with AR, ARCH, Normality, and RESET tests turning from highly significant to insignificant.

7 Model's Qualitative Predictions

Having shown that the piece-wise linear version of Model 2 passes specification tests for *all* of its linear segments, we now examine whether it generates predictions concerning *qualitative* co-movements between investor expectations and regressor variables that represent determinants of these expectations. These predictions can be used to assess the empirical relevance of alternative theoretical accounts of how participants form expectations and how stock prices move over time. To this end, Table 10 displays the estimates and test statistics for the piece-wise linear version of Model 2.

Table 10

Estimates of Parameters of Linear Segments Comprising Model 2

	63:03-80:03	80:04-99:10	99:11-15:06
$Exp_{t-1\|t+11}$	0.711 [15.02]	0.769 [19.52]	0.777 [11.86]
$Exp_{t-2\|t+10}$			-0.161 [-2.43]
$R_{t\|t-12}$	1.209 [6.29]	0.997 [10.10]	0.690 [6.87]
$R_{t-1\|t-13}$	-1.000 [-5.06]	-0.850 [-8.30]	-0.577 [-5.95]
Δi_t	-10.924 [-4.69]	-3.057 [-3.25]	4.888 [1.59]
Δu_{t-2}	-12.412 [-2.25]	0.748 [0.24]	-7.150 [-1.95]
c	3.138 [2.97]	-0.231 [-0.29]	8.089 [6.82]
Adj. R^2	0.707	0.775	0.724

Based on these estimates, we group the qualitative regularities predicted by a piece-wise linear model in Table 10 into three categories. The first concerns the degree of persistence of investor expectations. The second and third involve predictions about the role of extrapolation and fundamentals in driving these expectations.

7.1 Persistence of Investor Expectations

Investor expectations tend to be persistent. The lagged dependent variable is highly significant with t-values of over 10. Remarkably, the estimates of the coefficient for the proxy of lagged expectations are not only all positive; they also lie in a rather narrow range, between 0.7 and 0.8.

7.2 Extrapolation

The results in Table 10 show that investors' expectations are in part extrapolative. However, they also indicate that extrapolation did *not* drive a sustained swing in investor expectations during any of the subperiods of the model.

As we noted above, Autometrics retains the past return and its lag. Remarkably, Table 10 shows that both are significant in each subperiod and that their estimates are approximately the same in magnitude and have the opposite sign. This means that it is the change of the past return, $\Delta R_{t|t-12} = (R_{t|t-12} - R_{t-1|t-11})$, rather than its level that matters for investors expectations. This change is highly positively correlated with a one-month change in price, ΔP_t (at 0.0000% level; with a correlation coefficient of 0.66). As both $\Delta R_{t|t-12}$ and ΔP_t are stationary and not particularly persistent (with autocorrelation coefficients of less than 0.3), the extrapolative component of investors expectations dissipates fairly quickly.

7.3 Fundamentals

Table 10 shows that trends in fundamentals were primary drivers of swings in investors' expectations and thus stock-price fluctuations during all three subperiods.

7.3.1 Interest rate

The importance of the interest-rate variable is evident in Table 10. It has a *negative* and highly significant effect on investor's expectations in the first two subperiods.

7.3.2 Unemployment Rate

The unemployment rate is significant in the first and third subperiods during which it has a significant and *negative* effect on expectations.

8 Concluding Remarks

Behavioral-finance theorists have interpreted the rejection by Shiller (1981) and others of the REH present-value model as implying that stock-market expectations are driven by factors that are largely unrelated to fundamentals. This paper's finding that trends in fundamentals are a major driver of investors' expectations is inconsistent with this interpretation.

Indeed, our findings point to a very different explanation of the failure of the REH-based present-value model: REH does *not* represent how rational, profit-seeking participants in real-world markets form expectations on the basis of news about fundamentals. Frydman and Goldberg (2013a) have traced the reason for this explanation to REH's core premise: in forming their forecasts, market participants disregard *all* changes in the process underpinning outcomes that cannot be foreseen with a probabilistic rule.

Extracting information from Bloomberg News market wraps, Frydman *et al.* (2015) provide empirical evidence that undercuts this premise. Notably, 20% of the news that is reported as driving daily stock-price movements involves historical events that are to some extent unique, with consequences that are, *ipso facto*, unforeseeable. Such events thus engender so-called Knightian uncertainty, which cannot "be reduced to an objective, quantitatively determined probability" (Knight, 1921, p. 321).

Once we recognize the importance of unforeseeable change, *both* REH and behavioral insights matter for understanding investors' expectations. This paper's econometric finding – that, although fundamentals are a major driver of investor's expectations, extrapolation also plays a role – provides support for this hypothesis.

The findings here corroborate extensive descriptive evidence in Frydman *et al.* (2015) concerning the factors that market participants consider relevant for understanding stock-market movements. As reported by *Bloomberg*, participants mention at least one of the fundamental factors as a mover of stock prices on nearly all (99.4%) of the trading days over a 17-year period (from January 1993 to December 2009). Psychological and technical considerations (such as extrapolation) were mentioned considerably less frequently than fundamental factors. Nonetheless, their significance is obvious: Participants considered them relevant on roughly half of the trading days in the sample.

Although these findings accord both REH and behavioral-finance insights a role in understanding investors' expectations, they are inconsistent with the key implications of each of the approaches taken separately. In particular, while our findings support REH models' focus on fundamentals, they contradict these models' implication that psychological and technical considerations play no role in how market participants forecast outcomes. Our findings also upend the *raison d'être* of the behavioral-finance models, which assume that stock and other asset prices are driven by psychological and other factors that are largely unrelated to fundamental factors.

Frydman and Goldberg (2013a) have shown that opening models to unforeseeable change and the Knightian uncertainty that it engenders is the key to incorporating both REH and behavioral insights into representations of rational forecasting. As Keynes understood early on,

> We are merely reminding ourselves that... our *rational* selves [are] choosing between alternatives as best as we are able, calculating where we can [on the basis of fundamentals], but often falling back for our motive on whim or sentiment or chance. [Keynes, 1936, pp. 163, emphasis added]

This view of how rational participants forecast outcomes in real-world markets when faced with change that cannot be foreseen with a probabilistic rule poses considerable challenges for both model-building and econometric methodology. The apparent empirical relevance of Knightian uncertainty and the results presented here suggest that addressing these challenges is an important objective of future research.

References

[1] Bai, J. and P. Perron (1998), "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica,* 66(1), 47-78.

[2] Barberis, N. and R. Thaler (2003), "A Survey of Behavioral Finance," in Constandines, G., M. Harris and R. Stultz (eds.), *Handbook of the Economics of Finance,* Amsterdam: North Holland, 1053-1128.

[3] Bauer M.D. and J.D. Hamilton (2015), "Robust Bond Risk Premia," Working Paper.

[4] Campos, J., N.R. Ericsson, and D.F. Hendry (eds.) (2005), *Readings on the General-to-Specific Modeling,* Chaltenham: Edward Elgar.

[5] Castle, J.L., J. Doornik, D.F. Hendry, and F. Pretis (2015), "Detecting Location Shifts during Model Selection by Step-Indicator Saturation," *Econometrics*, 3(2), 240-264.

[6] Cochrane, J. (2011), "Presidential Address: Discount Rates," *The Journal of Finance*, 66, 1047-1108.

[7] Doornik, J.A. (1999), "Object-Oriented Matrix Programming using Ox," London: Timberlake Consultants Press.

[8] Doornik, J.A. (2009), "Autometrics," Shephard N. and J. L. Castle (eds.),*The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry,* Oxford: Oxford University Press, 88-121.

[9] Doornik, J.A. and H. Hansen (2008), "An Omnibus Test for Univariate and Multivariate Normality," *Oxford Bulletin of Economics and Statistics* 70(s1), 927-939.

[10] Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity, with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50(4), 987-1007.

[11] Ericsson, N.R. (2012), "Detecting Crises, Jumps, and Changes in Regime,", Working paper, Board of Governors of the Federal Reserve System, Washington, D.C.

[12] Frydman, R. and M.D. Goldberg (2007), *Imperfect Knowledge Economics: Exchange Rates and Risk,* Princeton University Press.

[13] Frydman, R. and M.D. Goldberg (2011), *Beyond Mechanical Markets: Asset Price Swings, Risk, and the Role of the State,* Princeton University Press.

[14] Frydman, R. and M.D. Goldberg (2013a), "The Imperfect Knowledge Imperative in Modern Macroeconomics and Finance Theory," in Frydman, R. and E. S. Phelps (eds.), *Rethinking Expectations: The Way Forward for Macroeconomics*, Princeton, N.J.: Princeton University Press, 140-165.

[15] Frydman, R. and M.D. Goldberg (2013b), "Opening Models of Asset Prices and Risk to Non-Routine Change," in Frydman, R. and E. S. Phelps (eds.), *Rethinking Expectations: The Way Forward for Macroeconomics*, Princeton, N.J.: Princeton University Press, 207-247.

[16] Frydman, R. and E.S. Phelps (2013), "Which Way Forward for Macroeconomics and Policy Analysis?," in Frydman, R. and E. S. Phelps (eds.), *Rethinking Expectations: The Way Forward for Macroeconomics*, Princeton, N.J.: Princeton University Press, 1-46.

[17] Frydman, R., M.D. Goldberg, and N. Mangee (2015), "Knightian Uncertainty and Stock Price Movements: Why the REH Present-Value Model Failed Empirically," *Economics E-Journal*, 9(24), 1-50.

[18] Frydman, R., H. Frydman, J. R. Stillwagon and M. N. Tabor (2016), "Markov Switching and Structural Change," in preparation.

[19] Godfrey, L.G. (1978), "Testing for Higher Order Serial Correlation in Regression Equations when the Regressors include Lagged Dependent Variables," *Econometrica,* 46, 1303-1313.

[20] Greenwood, R. and A. Shleifer (2014), "Expectations of Returns and Expected Returns," *Review of Financial Studies*, 27(3), 714-746.

[21] Hamilton, J. D. (1988), "Rational-Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates,"*Journal of Economic Dynamics and Control,* 12, 385-423.

[22] Hamilton, J. D. (1994), *Time-Series Analysis,* Princeton, N.J.: Princeton University Press.

[23] Hendry, D. F. and J. A. Doornik (2014), *Empirical Model Discovery and Theory Evaluation,* Cambridge, M.A: The MIT Press.

[24] Hendry, D.F, S. Johansen and C. Santos (2008), "Automatic Selection of Indicators in a Fully Saturated Regression," *Computational Statistics,* 23(2), 317-35.

[25] Hendry, D.F. and H.M. Krolzig (2001), "Computer Automation of General-to-specific Model Selection Procedures," *Journal of Economic Dynamics and Control* 25(6), 831-866.

[26] Hendry, D.F. and H.M. Krolzig (2005), "The Properties of Automatic Gets Modelling," *The Economic Journal,* 115(502), C32-C61.

[27] Hoover, K. and S.J. Perez (1999), "Data Mining Reconsidered: Encompassing and the General-to-specific Approach to Specification Search," *The Econometrics Journal,* 2(2), 167-191.

[28] Johansen, S. (1995), *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.

[29] Juselius, K. (2006), *The Cointegrated VAR Model: Methodology and Applications*. Oxford University Press.

[30] Keynes, J.M. (1936), *The General Theory of Employment, Interest and Money,* Harcourt, Brace and the World.

[31] Kitov, O. and M. Tabor (2015), "Detecting Structural Breaks in Linear Models: A Variable Selection Approach using Multiplicative Indicator Saturation," Working Paper, University of Copenhagen.

[32] Knight, F.H. (1921), *Risk, Uncertainty and Profit,* Boston: Houghton Mifflin.

[33] Lovell, M.C. (1983), "Data Mining," *Review of Economics and Statistics,* 65(2), 1-12.

[34] Lucas, R. E., Jr. (1976), "Econometric Policy Evaluation: A Critique," in Brunner, K. and A. H. Meltzer (eds.), *The Phillips Curve and Labor Markets,* Carnegie-Rochester Conferenceon Public Policy, Amsterdam: North-Holland.

[35] Ludvigson and Ng (2009), "Macro Factors in Bond Risk Premia," *Review of Financial Studies* 22(12), 5027-5067.

[36] Newey, W.K., and K.D. West (1987), "A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation-consistent Covariance Matrix," *Econometrica* 55(3), 703-708.

[37] Ramsey, J.B. (1969), "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis," *Journal of the Royal Statisical Society B,* 31(2), 350-371.

[38] Said, S.E. and D.A. Dickey (1984), "Testing for Unit Roots in Autoregressive- Moving Average Models of Unknown Order," *Biometrika* 71(3), 599-607.

[39] Shiller, R.J (1981), "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?," *American Economic Review* 71(3), 421-436.

[40] Shleifer, A. (2000). *Inefficient Markets: An Introduction to Behavioral Finance.* Oxford: Oxford University Press.

[41] Spanos, A. and J.J. Reade (2016), "Heteroskedasticity/Autocorrelation Consistent Standard Errors and Reliability of Inference." Working Paper for the 17th Annual Oxmetrics Conference, George Washington University.

[42] Stillwagon, J.R. (2016), ”Non-linear Exchange Rate Relationships: An Automated Model Selection Approach with Indicator Saturation,” *North American Journal of Economics and Finance*, forthcoming.

[43] Stillwagon, J.R. and P. Sullivan (2016), "Markov Switching Models of the Exchange Rate: Are There Only Two Regimes?" working paper.

[44] White, H. (1980), “A Heteroskedastic-consistent Covariance Matrix Estimator and Direct Test for heteroskedasticity,” *Econometrica,* 48(4), 817-38.

[45] Williams, J.C. (2013), "Bubbles Tomorrow, Yesterday, but Never Today?" FRBSF Economic Letters.

Appendix

Table A1

Unit Root Tests

	ADF w/ trend	ADF w/o trend
II $\text{Exp}_{t\|t+12}$	0.0000	0.0000
AA $\text{Exp}_{t\|t+12}$	0.0000	0.0000
$\ln(\text{P}_t/\text{D}_t)$	0.4759	0.5550
$\Delta\ln(\text{P}_t/\text{D}_t)$	0.0000	0.0000
u_t	0.1599	0.0440
i_t	0.2678	0.4299
Y_t	0.2500	0.2462
E_t	0.000	0.6752
D_t	0.0026	0.9212
C_t	1.0000	0.0000

Caption: p-values from the Dickey Fuller tests of the null of a unit root

Table A2

Estimates of Model 1 for the Bai-Perron Subsamples

	63:03-73:11	73:12-99:10	99:11-15:06
$Exp_{t-1\|t+11}$	0.712 [15.66]	0.803 [26.27]	0.768 [23.74]
$\Delta\ln(P_t/D_t)$	349.605 [9.31]	202.017 [10.75]	155.723 [14.07]
Δi_t	1.842 [0.52]	-3.120 [-2.46]	0.912 [0.38]
$\Delta\ln(D_{t-2})$	345.261 [1.15]	64.649 [0.39]	159.703 [3.19]
c	2.239 [1.46]	0.906 [0.87]	4.750 [5.86]
Adj. R^2	0.740	0.821	0.827

Table A3

Step Indicators from Autometrics

Model 1			Model 2		
	Coeff.	Bias-adj.		Coeff.	Bias-adj.
S:1966(02)	−12.096 [−2.32]	0.000 [0.00]	S:1966(04)	4.745 [2.70]	0.000 [0.00]
S:1966(04)	61.552 [7.02]	61.552 [7.02]	S:1979(09)	10.698 [3.24]	0.000 [0.00]
S:1966(05)	−40.517 [−5.47]	−40.510 [−5.47]	S:1980(06)	−11.464 [−3.54]	−7.954 [−2.45]
S:1968(06)	15.020 [2.85]	0.000 [0.00]	S:2015(05)	2.178 [4.06]	−1.867 [3.48]
S:1968(08)	−16.214 [−3.11]	0.000 [0.00]			
S:1971(05)	−9.996 [−3.26]	0.000 [0.00]			
S:1971(12)	5.545 [1.56]	0.000 [0.00]			
S:1972(10)	−14.941 [−2.70]	0.000 [0.00]			
S:1972(12)	17.386 [3.23]	0.000 [0.00]			
S:1974(04)	11.817 [2.21]	0.000 [0.00]			
S:1974(06)	−18.754 [−3.27]	0.000 [0.00]			
S:1975(01)	17.981 [3.86]	14.475 [3.10]			
S:1975(05)	−18.793 [−4.58]	−17.675 [−4.30]			
S:1977(12)	11.531 [5.93]	11.505 [5.92]			
S:1981(11)	6.297 [2.00]	0.000 [0.00]			
S:1982(05)	−18.524 [−3.18]	0.000 [0.00]			
S:1982(07)	22.329 [3.70]	16.837 [2.79]			
S:1982(12)	−13.950 [−4.09]	−12.051 [−3.53]			
S:1989(12)	3.298 [2.88]	0.000 [0.00]			
S:1996(09)	−5.438 [−5.35]	−5.382 [−5.29]			