Extrapolation Bias and the Predictability of Stock Returns by Price-Scaled Variables

Stefano Cassella
Tilburg University

Huseyin Gulen
Purdue University

Using survey data on expectations of stock returns, we recursively estimate the degree of extrapolative weighting in investors' beliefs (DOX). In an extrapolation framework, DOX determines the relative weight investors place on recent-versus-distant returns. DOX varies considerably over time. The ability of price-scaled variables to predict the year-ahead equity premium is contingent on DOX. High price-scaled variables are followed by lower returns only when DOX is high. Our findings support extrapolation-based theories of the stock market and the interpretation of price-scaled variables as mispricing proxies. Our results help answer a critical question: when will an overvalued asset experience a correction? (*JEL* G02, G12, G14)

Received December 5, 2015; editorial decision June 22, 2017 by Editor Robin Greenwood. Authors have furnished an Internet Appendix, which is available on the Oxford University Press Web Site next to the link to the final published paper online.

Ample evidence suggests that aggregate stock returns are predictable. The work of Fama and French (1988), Campbell and Shiller (1988), Cochrane (1992, 2008, 2011), and Lewellen (2004), among others, documents that the dividend-price (D/P), the book-to-market (B/M), and the earnings-to-price (E/P) ratios can predict future returns. Time-series predictability of stock market returns by price-scaled variables is often attributed to time-series variation in investors' required returns, suggesting a risk-based explanation.[1] Yet behavioral theorists

We appreciate helpful comments from Adem Atmaz, Nick Barberis, Tolga Cenesizoglu, Zhi Da, Robin Greenwood, Yeejin Jang, Lawrence Jin, Mitch Johnston, Mohitosh Kejriwal, Ralitsa Petkova, Stefano Rossi, Allan Timmermann, Chris Wagner, Deniz Yavuz, and seminar participants at Caltech, Case Western Reserve University, Chicago Quantitative Alliance, Michigan State University, Purdue University, Texas A&M University, Texas Tech University, University of Illinois at Urbana-Champaign, University of Toronto, and the 2016 European Finance Association Meetings (EFA). Supplementary data can be found on *The Review of Financial Studies* web site. Send correspondence to S.Cassella, Tilburg School of Economics and Management, Tilburg University, 5000 LE, Tilburg, NL. +31(0)61-127-0936, s.cassella@uvt.nl.

[1] Time-series variation in required compensation for risk may arise because of variation in (1) risk aversion (Campbell and Cochrane 1999), (2) aggregate consumption risk (Bansal and Yaron 2004; Bansal, Kiku, and Yaron 2012), (3) rare-disaster risk (Gabaix 2008), (4) risk-sharing opportunities among heterogeneous agents (Lustig and Van Nieuwerburgh 2005), and (5) beliefs (Timmermann 1993; Detemple and Murthy 1994).

doi:10.1093/rfs/hhx139 Advance Access publication February 2, 2018

propose that predictability may arise because prices temporarily deviate from the levels warranted by fundamentals due to the existence of irrational traders who hold biased beliefs.[2] Motivated by the arguments in these behavioral models, we investigate the extent to which time series variation in biased beliefs can account for the observed predictability relation between price-scaled variables and future stock returns.

Determining the role behavioral biases play in the extant evidence of return predictability is no easy task, because a researcher must assess both the existence and the extent of bias in investors' expectations. Work by Greenwood and Shleifer (2014) fills this gap by providing evidence of one such bias in investors' beliefs: *overextrapolation*.[3] The authors show that surveys of investors' expectations of future stock market returns are a direct and reliable measure of beliefs. Additionally, survey data provides evidence of the tendency of investors to form expectations of future stock returns by relying more (less) heavily on recent (distant) stock returns. In a related work, Barberis et al. (2015) present an equilibrium model of financial markets with heterogeneous investors and biased beliefs in which an increasing reliance of extrapolative expectations on recent returns leads to stronger short-horizon stock return predictability by the dividend-price ratio.[4]

In our study, we use survey data on stock market expectations to quantify the extent to which extrapolators use recent versus distant past returns to form beliefs. This measure captures an important aspect of overextrapolation: the tendency of investors to overweight recent returns.[5] We quantify this tendency and label it *degree of extrapolative weighting* (DOX).[6] We document considerable variation in DOX over time and then test the implications of the variation in DOX for the predictability of the equity premium by price-scaled variables. In conditional forecasting regressions of excess returns on horizons up to a year, we find that price-scaled variables predict future stock returns only when our DOX measure is high, whereas these variables hold no predictive ability at short horizons when DOX is low. This result is confirmed by out-of-sample tests and applies to stock return predictability by the dividend-price, book-to-market, and earnings-to-price ratios.

[2] Mispricing as an equilibrium outcome may arise if rational investors find it optimal not to offset irrational investors' trades (De Long et al. 1990; Shleifer and Vishny 1997; Barberis et al. 2015) or if they may profit from *riding a bubble* (Abreu and Brunnermeier 2003).

[3] Throughout the paper, we use the terms extrapolative expectations, overextrapolation, and extrapolation bias interchangeably. An extrapolative investor believes that recent high returns are more likely to be followed by high returns, and similarly, recent low returns are more likely to be followed by low returns.

[4] Other studies on extrapolation in financial markets include Choi and Mertens (2013) and Hirshleifer, Li, and Yu (2015).

[5] This tendency may have multiple sources. The first is psychological and is rooted in representativeness (Barberis, Shleifer, and Vishny 1998; Bordalo, Gennaioli, and Shleifer 2017) or in a belief in the law of small numbers (Rabin 2002). The second possible source is bounded rationality (Hong and Stein 1999; Fuster, Hebert, and Laibson 2012; Glaeser and Nathanson 2017).

[6] It is important to note that DOX reflects the relative weight the extrapolators place on past returns, not the amount of extrapolators in the marketplace.

This study makes several contributions. First, the paper is a test of extrapolation-based models of the aggregate stock market, specifically of Barberis et al. (2015; hereafter BGJS), who establish a theoretical link between the predictive power of the D/P ratio and extrapolators' tendencies to rely too heavily on recent versus distant returns (i.e., high DOX). Our results provide support to such models, since we show that the ability of price-scaled variables to predict the equity premium on short horizons is contingent on high DOX. Second, this paper provides evidence in favor of the economic and statistical strength of return predictability. In particular, instead of offering a new predictor, we show that if we reexamine the common predictors of future returns through the behavioral lens of extrapolation bias, the dynamics of aggregate stock returns are better understood. Third, we show that time-series variation in the degree of extrapolative weighting can reconcile evidence of instability in the predictive relation between price-scaled variables and future returns. Fourth, we document that a survey-based state variable allows us to better understand the relationship between aggregate quantities set in equilibrium, such as returns and price-scaled variables. In doing so, we reinforce the message in Greenwood and Shleifer (2014): survey expectations contain useful information on widely held economic beliefs.

To estimate the DOX, we use a nonlinear least-squares regression in which survey expectations of future stock market returns are regressed on quarterly stock returns lagged up to 60 quarters. The DOX is measured as the relative loading of expectations on returns in recent quarters compared to returns in more distant quarters. If future index returns are only weakly correlated with recent returns, an excessive reliance of expectations on recent stock market performance (a high DOX) suggests that investors extrapolate recent returns too much into the future. For example, in the period 1992:06–2014:12, our full sample DOX estimate obtained using the Investor Intelligence Survey implies that the loading of survey expectations on the returns over the most recent quarter is 16 times higher compared to the loading only four quarters earlier. This means that when forming expectations, investors view returns realized four quarters earlier as only 6% as important as those in the most recent quarter. During the same period, serial correlation in consecutive quarterly stock market returns is only 7%, and serial correlation between consecutive yearly returns is actually a negative 5%. This lends support to the interpretation that investors on average *over*extrapolate recent market returns into the future.

In the model of BGJS, the mechanism that links DOX to stock return predictability by price-scaled variables is straightforward. Extrapolation of past returns induces investors to form biased expectations. For instance, a recent history of positive (negative) stock return realizations causes investors to revise their expectations of future returns upward (downward). Irrationally high (low) expectations induce an irrational demand for stocks, which pushes prices too high (low) relative to fundamentals. As a result, the D/P ratio declines (increases) due to misvaluation. This conjecture is in line with the negative correlation between survey expectations and price-scaled variables observed in the data. On average, this overvaluation (undervaluation) is not sustained, since extrapolators observe new returns which do not support their initial optimism (pessimism). The key insight is that when extrapolators exhibit a heightened tendency to disproportionately use recent returns to form expectations, that is, DOX is high, few new return observations can quickly lead to significant changes in expectations. The quicker correction of biased beliefs causes the D/P ratio to revert more quickly toward the mean. So low (high) prices are more likely to be followed by high (low) prices in the near future, with an ensuing pattern of short horizon predictability of future returns that is stronger when DOX is high.

To test the implications of the proposed mechanism empirically, one needs time series variation in DOX. We argue that there are reasons to believe, ex ante, that the DOX is time varying. The variation we document may occur either because the market participation rate of high-DOX individuals varies over time or because recent returns appear more salient than distant ones in some circumstances, and less so in others. These are not competing explanations, since one can induce or reinforce the other. For instance, when the market rises, high-DOX extrapolators may decide to join the market, pushing DOX up. When the market collapses, they become bearish and exit the market, lowering DOX. Empirically, we recursively estimate DOX using survey expectations of future returns. After uncovering time-series variation in DOX, we conduct formal tests to understand why the DOX changes over time. First, DOX increases with the share of young investors in the marketplace. This is consistent with the young being high-DOX individuals whose participation is time varying and may result in changes of "consensus" extrapolation through time. Second, DOX increases following a period of good stock returns, as well as with the time since the stock market crash, suggesting that return salience may fluctuate over time.

In our main tests, we run short-horizon conditional stock return predictability regressions in which the DOX acts as a state variable and is interacted with the D/P ratio. We show that stock return predictability by price-scaled variables is conditional on the DOX. In the period 1992:06–2013:12, we find that when the extrapolators' DOX is 0.71 (one standard deviation higher than its median value), a one-standard-deviation rise in D/P ratio is followed by a statistically significant 26% increase in the expected equity premium the following year. When instead the DOX is 0.31 (one standard deviation lower than its median value), the same increase in the D/P ratio is *negatively*, but insignificantly, related to future returns, and predicts a 2% lower equity premium in the upcoming year.[7] Our results are robust to known forms of small-sample bias in predictive regressions, and our conditional predictive model appears to beat the naive forecast and the traditional univariate models in out-of-sample tests.

[7] Unless otherwise stated, all the results discussed in the Introduction refer to the use of the DOX extracted from the principal component of the Investor Intelligence and the American Association of Individual Investors surveys.

We obtain similar results for other price-scaled predictors, such as the B/M and E/P ratio.

This main finding has important implications. Simply observing overvaluation (undervaluation) in the marketplace as reflected in low (high) D/P ratios does not necessarily mean that the market will soon experience a correction. Our study helps answer this age-old question: namely, when will a mispriced asset correct back to fair value? We show that when investor beliefs load on distant past returns (low DOX), mispricing is unlikely to correct soon. On the other hand, when investor beliefs load heavily on recent returns (high DOX), there is a high chance of quick correction. In this case, even one period of bad news can quickly result in a significant change in expectations. At longer horizons, however, where the D/P ratio is unconditionally a powerful predictor of future stock returns (Cochrane 2008), the predictive link does not appear to be conditional on DOX. We also note that, approximately 15% of our monthly forecasts of year-ahead excess returns are negative. A negative equity premium prediction arises when market overvaluation (i.e., a low D/P ratio) is accompanied by a high DOX (i.e., a high likelihood of mispricing correction). This evidence of a negative risk premium prediction is hard to reconcile with rational models of risk.

In our second test, we find that when the DOX is low, price-scaled variables are more persistent. In particular, at a one-year horizon when the DOX is one standard deviation below (above) its median value, the D/P ratio has an autoregressive coefficient of approximately 0.88 (0.43). In other words, in periods of relatively low DOX, the half-life of a shock to the D/P ratio is approximately five years, while it is only 10 months when the DOX is relatively high. This is consistent with the existence of a linkage between mean-reversion in the D/P ratio and stock return predictability, and with the view that a high DOX quickens the former and thus increases the latter.

In our third test, we assess the extent to which extrapolation bias can explain the evidence in prior literature that the relationship between price-scaled predictors and future returns varies over time (Paye and Timmermann 2006; Lettau and Van Nieuwerburgh 2008; Henkel, Martin, and Nardari 2011). We argue that variation in the extent of time series predictability of stock returns may arise because investors' expectations are derived from a few recent returns more heavily in some periods (resulting in a subsequent correction), and less heavily in other periods. We not only confirm prior evidence of parameter instability, but also find that better predictability is obtained in periods characterized by higher average DOX. Furthermore, variation in average DOX across sample periods can explain 70% of the documented instability in the univariate predictability relation.

This study shares its general topic of inquiry with Bacchetta, Mertens, and Wincoop (2009), Amromin and Sharpe (2013), and Koijen, Schmeling, and Vrugt (2014), who study irrationality in expectations, and with Baker and Wurgler (2000), Yuan (2015), who provide a behavioral explanation for return predictability. Unlike prior literature, and motivated by Greenwood and Shleifer (2014) and BGJS, we show that the ability of price-scaled variables to predict future returns, traditionally linked to time-varying discount rates, is conditional on the degree of extrapolative weighting.

1. Hypothesis Development

1.1 Present-value model

To motivate our study, we rely on the present-value model of Campbell and Shiller (1988):

$$r_{t,t+1} = \Delta d_{t,t+1} + [dp_t - \rho dp_{t+1}], \tag{1}$$

where r is log raw return, dp is log dividend-price ratio, Δd is the log dividend-growth, and ρ is a constant whose historical value is 0.96. All quantities are demeaned. Equation (1) states that the future return on a risky security is higher because the security will pay higher dividends in the future or because the equilibrium price per unit of dividend will increase. Focusing on the conditioning information I=$\{dp, \Delta d, r\}$ one can posit that

$$\Delta d_{t,t+1} = \epsilon^{\Delta d}_{t+1} \tag{2}$$

and

$$dp_{t+1} = \Psi dp_t + \epsilon^{dp}_{t+1}. \tag{3}$$

Equation (2) states that consistent with evidence in Cochrane (2005, 2008), the best estimate of the future level of dividends is the current dividend level. Equation (3) models the dp as an AR(1) process with persistence coefficient Ψ. Substituting Equations (2) and (3) into Equation (1), we obtain

$$r_{t,t+1} = dp_t(1-\rho\Psi) + \epsilon^{\Delta d}_{t+1} - \rho\epsilon^{dp}_{t+1}. \tag{4}$$

Equation (4) suggests that if one runs a univariate linear predictive regression of future one-period returns on current dp, the predictability coefficient is $(1-\rho\Psi)$.

This coefficient is a function of the mean-reverting behavior exhibited by the dp ratio. On short horizon, if the dividend-price ratio is persistent, that is, Ψ is high, an increase (decline) in prices today is less indicative of lower (higher) prices tomorrow. Consequently, a univariate predictive regression of year ahead stock returns shows that the dividend-price ratio has little predictive ability, and the best forecast of the future return is the unconditional mean. When instead the dividend-price ratio mean-reverts more quickly, that is, Ψ is low, a shock to prices will quickly mean revert, and the predictability coefficient is larger. For a sufficiently long horizon, as the dividend-price ratio is bound to fully revert to its mean, the link between current D/P and future cumulative returns strengthens, and the predictability coefficient becomes insensitive to Ψ.[8]

[8] See equation 12 in Cochrane (2008). Under the assumption of no dividend-growth predictability outlined in Equation (2), the short-horizon return predictability coefficient is $b_r = (1-\rho\Psi)$ like in Equation (4) and the long-run return predictability coefficient $b_r{}^{lr} = 1$.

On one hand, Equation (4) suggests that the extent to which a price-scaled variable such as the D/P ratio can predict future returns on short horizon depends on how quickly the ratio mean reverts. On the other hand, the equation above is silent on the economic forces behind such mean-reversion. Below, we consider extrapolation bias as one such force, and test the prediction of BGJS, who posit that the extent of short-horizon mean reversion in the D/P and the associated return predictability depend on how investors' extrapolative expectations are formed. We explain this in further detail below.

1.2 Extrapolation and return predictability

The investigation of a potential link between extrapolation of past returns by market participants and aggregate stock return predictability rests on a few assumptions. The first is that individuals have extrapolative expectations. De Bondt (1993), Clarke and Statman (1998), Amromin and Sharpe (2013), and Greenwood and Shleifer (2014) find evidence of extrapolation bias in survey-based forecasts of future returns. Similarly, Tversky and Kahneman (1974) and Andreassen and Kraus (1990) offer evidence of extrapolation bias in experimental settings.

The second assumption is that individuals act in accordance with their extrapolative beliefs. Given that most evidence of overextrapolation is based either on surveys or on experimental results, and in both environments there may be lack of sufficient incentive to elicit true expectations, a discrepancy between individuals' beliefs and their subsequent actions is possible. Gennaioli, Ma, and Shleifer (2016) use individual-level responses to the Graham and Harvey survey of CFOs' expectations to provide evidence of consistency of CFOs' extrapolative forecast of future firm growth, and their subsequent planned and realized investments. Similarly, Greenwood and Shleifer (2014) show that aggregate survey expectations of future market returns, which display an extrapolative nature, correlate positively with aggregate mutual fund inflows. This again suggests that survey responses and subsequent actions are aligned.

The last assumption is that rational investors fail to instantly correct the mispricing caused by extrapolative investors, and therefore extrapolation matters in equilibrium. Some investors' extrapolative beliefs may affect their decisions and still not matter in equilibrium, since a rational investor may act immediately to correct the mispricing caused by the extrapolators who populate the market. This argument critically relies on rational investors' willingness and ability to correct such mispricing.

In this respect, a vast theoretical literature that includes De Long et al. (1990), Shleifer and Vishny (1997), and Abreu and Brunnermeier (2003), among others, calls into question the notion that it is always in a rational investor's best interest to bet against her irrational counterpart.

Building on this evidence of the pervasiveness of extrapolation in the economy, and the likely effect it has on equilibrium prices, the theoretical model of BGJS (2015) is the first to provide an extrapolation-based explanation for the extant evidence of predictability. In the model, if the extrapolative investors in the economy form expectations of future returns by relying more heavily on recent stock returns (i.e., the DOX is high), the dividend-price ratio mean-reverts more quickly (lower Ψ). As per Equation (4), this suggests a stronger link between current D/P and future short-horizon returns. On longer prediction horizon, the conditioning role of current DOX plays a smaller role, as the dividend-price ratio is bound to fully mean revert.[9]

While BGJS (2015) provide a novel comparative statics result, an empirical test of the implications of their model requires variation (either in the time series or in the cross-section) in the extrapolators' DOX. Below, we provide arguments for the exante assertion that the DOX is time varying. We then measure the extent of such variation and test its implications for stock return predictability.

1.3 Time-varying DOX

The aggregate extrapolators' DOX may change over time as per the experience-based evidence of Malmendier and Nagel (2011). Young individuals, who base their forecast of future returns on a shorter macroeconomic history, intrinsically extrapolate using a higher DOX. Older investors, who have instead witnessed a longer return time series, implicitly adopt a lower DOX. Over time, stock market participation rates by the two groups change as a reflection of their own experience of the stock market (Nagel 2012). This causes an alteration to the mix of stock market investors, which may cause the *consensus* DOX to change over time.

There is also a second potential channel that could be associated with time-series variation in DOX and may also be responsible for observed changes in stock market participation. Griffin and Tversky (1992) document that when making predictions based on new information, individuals are sensitive to the information's perceived strength, expressed in terms of salience, and may overreact relative to a rational forecaster. The relative salience of recent versus distant past returns may depend on features of the returns themselves, such as their extremeness. In this respect, Yuan (2015) finds that attention-grabbing events may deeply affect investor trading patterns. Similarly, the salience attached to a new return realization may be path-dependent, like in Da, Gurun, and Warachka (2014), or context-dependent, and change with the temporal proximity of a stock market or economic crash. DOX may also relate to broader sources of irrationality in the stock market, such as those tracked by the Baker and Wurgler (2006) investor sentiment. In Section 4.1, we document times-series variation in the DOX empirically, and in Section 4.2 we test on the underlying determinants of DOX.

[9] See BGJS, proposition 3, and the related table 3 for a more formal analysis of the role played by β, the model counterpart of DOX, in predictive regressions.

2. Econometric Approach

Motivated by the above arguments and following Greenwood and Shleifer (2014), we model extrapolative expectations as follows:

$$Exp_t = a + b\sum_{i=0}^{\infty} w_i R_{t-(i+1)\Delta t, t-i\Delta t},$$
$$w_i = \frac{\lambda^i}{\sum_{k=0}^{\infty}\lambda^k}, \; 0 \leq \lambda < 1, \tag{5}$$

where Exp_t refers to extrapolators' expectations as of time t (obtained from survey data), and $R_{i,j}$ is the return realized between time i and time j. Equation (5) states that expectations are a function of past return realizations in which the weights placed on historical returns feature a geometric decay. Δt determines the frequency of return observations. Following prior literature, we choose $\Delta t = 1/4$ and use quarterly returns. A lower λ implies that investors place higher weight on more recent observations, while earlier observations contribute less to an extrapolator's expectations. For example, when $\lambda = 0.85$, investors place twice as much weight on the most recent return realization compared to returns only four quarters earlier. The relative weight is 10 times higher compared with the weight four quarters earlier when $\lambda = 0.55$. The smoothness coefficient, λ, plays a significant role in this framework, since a lower λ is associated with both possible overreaction and with lower persistence in the beliefs of extrapolators. Figure 1 shows simulation results that illustrate these aspects of λ. In the figure that fixes the coefficient b to 1, we assume that at time $t = -1$ extrapolators' expectations of annual returns are at their long-run mean value of approximately 10%. At time $t = 0$, a quarterly return of 5% is realized and incorporated into expectations. We then report the average value of subsequent extrapolators' expectations obtained by simulating 5,000 time series of subsequent returns.[10] Figure 1 presents results for four different values of λ, and shows that while extrapolators always revise their expectations following a new return realization, the extent of their reaction as well as the speed of subsequent mean reversion in beliefs are larger when λ is low.

To assess how time-series variation in the degree of extrapolative weighting interacts with price-scaled variables in predicting stock returns, we estimate Equation (5) recursively by nonlinear least squares and extract the DOX, measured as $1 - \lambda$. The recursive estimation could be performed in every month m using a rolling window (m-l+1,m) of past l observations, so as to obtain an

10 Simulated quarterly returns are generated by matching mean and variance of the historical quarterly returns of the CRSP value-weighted portfolio in the period 1990–2014. Serial correlation in quarterly returns could potentially change the shape of investors' expectations reported in Figure 1. Nevertheless, as reported above, the historical autocorrelation of quarterly returns is low, and we choose to not consider it in our simulations.

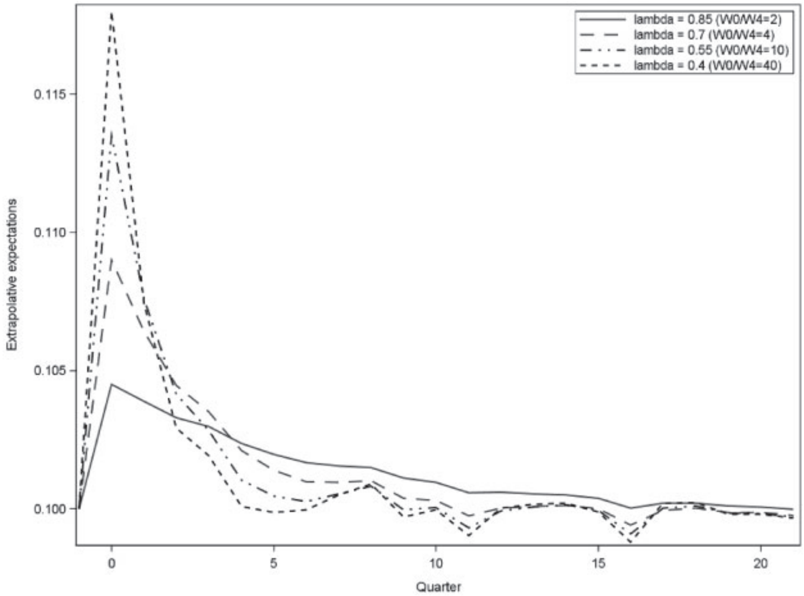

Figure 1
Time-series behavior of extrapolative beliefs
We model extrapolative beliefs as an infinite summation of current and past quarterly return realizations, like in Equation (5). Assume that at the end of quarter $t=-1$ extrapolators' expectations are at their long-run mean value of about 10%. At time $t=0$, a quarterly return of 5% is realized and incorporated into expectations. We then report the average value of extrapolators' expectations in $t \in [1,20]$, obtained by simulating 5,000 time series of subsequent returns. These returns are generated by matching the mean and variance of the historical quarterly returns of the CRSP value-weighted portfolio in the period 1987–2014. We report the resultant pattern in extrapolators' expectations for four values of the smoothness parameter λ, 0.85, 0.7, 0.55, and 0.4, which correspond to a relative weight of current quarterly returns versus quarterly returns in the previous year of 2, 4, 10, and 40, respectively.

updated estimate of DOX for month *m*. One key parameter in the recursive estimation is the length of the estimation window, *l*. Instead of arbitrarily choosing a fixed window length, we follow Pesaran and Timmermann (2007) and Capistran and Timmermann (2009) and endogenize window selection by combining estimates obtained using different window sizes.

Specifically, to assign a DOX value to month *m* we first obtain parameter estimates for months *m-12* to *m-1*. We refer to this 12-month period as the cross-validation period. For each of these 12 months, we estimate DOX over three alternative rolling window sizes based on prior 24, 36, and 48 months. For instance, in month *m-12* we use survey data in the interval [*m*-36+1,*m*-12] to obtain parameter estimates based on the 24-month rolling window. Similarly, we use [*m*-48+1,*m*-12] to obtain parameter estimates based on the 36-month rolling window, and so on. In addition to the 3 moving windows mentioned above, we also estimate the model using an expanding window (*Expanding*), which is based on all prior observations. Subsequently, for each of the months in the cross-validation period, we calculate 1-step-ahead forecast errors. Continuing on the prior example, we use the parameter estimates obtained in month m-12 with the 24-month rolling window to construct fitted survey expectations for month m-11, $\widehat{Exp}^{\{24\}}{}_{m-11}$. We compare this fitted value with the actual one Exp_{m-11}, calculate the difference $\epsilon^{\{24\}}{}_{m-11}$, and store it. We proceed in this way for each of the four alternative window lengths and for each of the 12 months in the cross-validation period. Ultimately, we obtain a set of 12 errors for each moving window, and we use them to calculate four mean-squared forecast error (MSFE) metrics, $MSFE^{\{24\}}$, $MSFE^{\{36\}}$, $MSFE^{\{48\}}$, and $MSFE^{\{Expanding\}}$, which we assign to month m. Finally, we estimate a candidate DOX for month m using each of the four moving windows, and then consolidate the four separate estimates into one, by means of a weighted average. The weights are the inverse of the MSFE (lower MSFE gets more weight) and are normalized to sum to 1.

Once we estimate the DOX time series, we use it as a conditioning variable in traditional predictive regressions in which the l-months ahead cumulative excess return $R_{t,t+l}$ is regressed on the lagged dividend-price ratio (and other price-scaled variables such as the B/M and E/P ratios). Specifically, we estimate the following linear model:

$$R_{t,t+l} = (a_0 + a_1 \mathrm{DOX}_t) + D/P_t (b_0 + b_1 \mathrm{DOX}_t) + \epsilon^R_{t,t+l}. \tag{6}$$

The null hypothesis of no effect of extrapolation on stock return predictability by the D/P ratio and the alternative one-sided hypothesis of an increase in predictability as the DOX increases are

$$H_0 : b_1 = 0 \quad H_a : b_1 > 0. \tag{7}$$

Later, we explore other implications of the alternative hypothesis. The first concerns the autoregressive behavior of price-scaled predictors of the equity premium, and posits that such predictors should revert to the mean more quickly when DOX is high. The second explores the link between parameter instability in the predictability relation and time-series variation in the DOX, and argues that periods of stronger predictability are those in which average DOX is higher. The third studies whether the conditioning role of DOX is stronger at short rather than long prediction horizon, as BGJS suggest.

3. Data

In our study, we rely on surveys of expectations of future stock market returns in the US. For statistical power and comparability with prior studies, we focus solely on the two longest available surveys. The Investor Intelligence Survey (II) collects forecasts of stock market performance since 1963 from newsletters of financial advisors in the United States. AA is the survey of retail investors from the American Association of Individual Investors, which started in 1987. In Table 1, we report general information about survey data as well as summary statistics.

Table 1
Summary statistics

Panel A. Survey general information

Survey	Period	N	Mean	Median	Minimum	Maximum	σ	ρ_3	ρ_{12}	Type	Frequency
II	1963-2014	626	13.63	15.89	−49.2	66.64	19.79	0.56	0.21	Professional	Weekly
AA	1987-2014	331	8.62	9.5	−41	50.47	15.27	0.31	0.17	Retail	Weekly
PC	1987-2014	331	0.00	0.12	−3.87	2.95	1.23	0.44	0.20		Monthly
GALLUP_ER	1998-2004	54	10.47	10.35	4.5	16.2	3.08	0.84	0.52	Retail	Monthly

Panel B. Other variables

Variable	Period	N	Minimum	Maximum	Mean	Median	σ	ρ_3	ρ_{12}
R^e	1963–2014	623	−0.44	0.53	0.06	0.08	0.16	0.75	−0.04
D/P	1963–2013	612	0.01	0.06	0.03	0.03	0.01	0.97	0.89
B/M	1963–2013	612	0.12	1.21	0.51	0.46	0.27	0.98	0.91
E/P	1963–2013	612	0.02	0.15	0.06	0.05	0.03	0.98	0.91

Panel C. Pairwise correlations

	II	AA	PC	Gallup_ER	D/P	B/M	E/P
AA	0.53						
	[0.00]						
PC	0.87	0.87					
	[0.00]	[0.00]					
GALLUP_ER	0.46	0.53	0.58				
	[0.001]	[0.00]	[0.00]				
D/P	−0.35	−0.42	−0.58	−0.81			
	[0.00]	[0.00]	[0.00]	[0.00]			
B/M	−0.25	−0.38	−0.42	−0.57	0.94		
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]		
E/P	−0.16	−0.32	−0.40	−0.83	0.91	0.89	
	[0.00]	[0.007]	[0.00]	[0.105]	[0.00]	[0.00]	
R^e	0.49	0.38	0.47	0.85	−0.24	−0.20	−0.12
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.01]

his table presents summary statistics. Panel A reports general information on the surveys of investor expectations in our analysis. *II* is the Investor Intelligence survey of U.S. financial advisors' newsletters. *AA* is the survey conducted by the American Association of Individual Investors. *PC* is the principal component of AA and II. *Gallup ER* is a version of the UBS\Gallup survey that elicits quantitative forecasts of future returns. Panel B reports summary information for the other time series used in this paper. *D/P* is the aggregate dividend-price ratio on the S&P 500. *B/M* is the aggregate book-to-market ratio on the DJIA. Both D/P and B/M are from Amit Goyal's Web site. *E/P* is the aggregate cyclically adjusted earnings-to-price ratio on the S&P 500, from Robert Shiller's Web site. Excess return (R^e) is the 12-month return on the CRSP value-weighted portfolio of US stocks, in excess of the risk-free rate. In panels A and B, $\rho(l)$ is the l-month autocorrelation and σ is the unconditional volatility. Panel C shows pairwise correlations between the time series used in the paper.

Both II and AA collect qualitative data and report the difference between the percentage of polled investors who are bullish and the percentage of polled investors who are bearish about future stock market performance. While qualitative and quantitative expectations may be different, like Greenwood and Shleifer (2014), we argue that qualitative survey data is a good proxy for quantitative expectations. To support this claim, we include in Table 1 data on a UBS/Gallup survey (*Gallup ER*) that was conducted for a short period of time between 1998 and 2004 and elicited quantitative forecasts. As the pairwise correlation statistics in panel C show, Gallup ER is highly correlated with both II (46%) and AA (53%). For this reason, we consider qualitative expectations data as a close substitute for quantitative data. Table 1 also presents summary statistics for the principal component of II and AA, (PC), which spans the period 1987:06–2014:12. Later, in our main tests, we focus on the period

1987:07–2014:12, in which both surveys and their principal component are available.

Panel B of Table 1 provides summary statistics on price-scaled variables and excess returns. The D/P ratio is a 12-month moving sum of dividends paid on the S&P 500 index, normalized by the most recent price. Fama and French (1988) find that a high D/P ratio is associated with higher returns on horizons that range from one to five years. Campbell and Shiller (1988) complement this finding by showing that the positive association between the D/P ratio and the subsequent total market return can be justified in the context of the simple present-value relation in Equation (1).

The B/M ratio, the predictive ability of which has been studied by Pontiff and Schall (1998), Kothari and Shanken (1997), and Lewellen (1999), is the ratio of book value to market value for the Dow Jones Industrial Average. Finally, the cyclically adjusted E/P ratio, considered as a predictor of aggregate stock returns by Campbell and Shiller (1988, 2001), is a 10-year moving average of earnings on the S&P 500 index, normalized by the current price.[11] The equity premium is defined as the difference between the return on the CRSP value-weighted portfolio, and the risk-free rate of return.[12] Panel B of Table 1 confirms the evidence in prior literature that price-scaled variables are persistent. Additionally, the monthly nature of our data causes yearly excess returns to be serially correlated at a quarter lag, while the correlation between nonoverlapping returns is close to zero.

Panel C of Table 1 shows high correlation among surveys, which suggests that independently collected data on investors' expectations tell a consistent story. Furthermore, the PC is approximately 90% correlated with both II and AA. This high correlation justifies its use as a representative series. The panel also shows that expectations of future returns are negatively correlated with price-scaled variables. This is consistent with the notion that improving expectations are associated with an increase in prices relative to fundamentals. Lastly, the extrapolative nature of survey-based expectations is reflected in the high correlation between survey forecasts of future returns, and the returns accumulated over the course of the year.

4. Results

4.1 DOX

For each survey, we first estimate Equation (5) for the full sample using nonlinear least squares. The infinite summation in Equation (5) is not amenable to estimation. Following Greenwood and Shleifer (2014) and BGJS, we choose

11 The D/P and B/M ratios are from Amit Goyal's Web site (http://www.hec.unil.ch/agoyal), and E/P is from Robert Shiller's Web site (http://www.econ.yale.edu/šhiller/data.htm).

12 The latter is from Kenneth French's Web site (http://www.mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

Table 2
Degree of extrapolative weighting: full sample

Coefficient	AA	II	PC
DOX (1-λ)	0.63	0.50	0.60
	[2.37]	[7.72]	[5.45]
a	5.42	6.92	−0.34
	[3.32]	[3.81]	[−4.78]
b	130.70	238.12	12.59
	[4.85]	[6.78]	[8.02]
Adj.R^2 (%)	21.90	30.10	27.50

The table presents full-sample estimates of the *degree of extrapolative weighting* (DOX) parameter. Specifically, we use nonlinear least squares to fit the following equation:

$$Exp_t = a + b\sum_{j=0}^{59} w_j R^Q_{t-j-1,t-j} + \epsilon^{Exp}_t,$$

where Exp_t is the survey-based expectation of future returns, $R^Q_{t-j-1,t-j}$ is the j-lagged quarterly return of the aggregate value-weighted stock market from CRSP, and w_j is a weight parameter whose expression is

$$w_j = \frac{\lambda^j}{\sum_{k=0}^{59}\lambda^k}.$$

We then report the corresponding extrapolators' DOX, defined as 1-λ, as well as estimates of the center and scale parameters *a* and *b*. Coefficients and *t*-statistics are based on Newey-West standard errors with a lag length of six months. Estimates are presented for II, AA, and their principal component PC. The sample period is 1987:06–2014:12.

a number of lags equal to 60. The estimated λ coefficient is then mapped onto the corresponding DOX$=1-\lambda$, which is reported in Table 2.

The full sample DOX extracted from II is 0.5, which corresponds to a weight on the most recent quarterly return that is 16 times larger than the weight assigned to the return four quarters earlier. In the case of the DOX measured from AA, there is even greater evidence of overextrapolation, since recent quarterly returns are given 55 times more weight compared to quarterly returns realized a year earlier. As expected, the principal component series exhibits a full-sample DOX of 0.6 that lies between the II and AA estimates.

Next, we use the dynamic-window combination methodology described in Section 2 to capture time-series variation in the DOX. In Table 3, we present summary statistics for the individual surveys as well as for their principal component PC. The recursive methodology generates DOX estimates for the period 1992:06–2013:12, since 5 years are needed to initialize the window selection algorithm.

Table 3 also presents summary statistics for an additional DOX time series, which we henceforth refer to as PC_{ext}. It is constructed using the DOX extracted from II during the period 1967:12–1992:05, and the DOX extracted from the principal component time series for the subsequent period. In our main tests, we focus on the short sample period, where multiple surveys are available and a common source of variation across them can be identified. We resort instead to the longer DOX time series when performing long-horizon predictability tests,

Table 3
Degree of extrapolative weighting: recursive estimation

Panel A. DOX (1-λ)

Survey	N	Mean	Median	σ
II	270	0.49	0.51	0.20
AA	270	0.58	0.57	0.20
PC	270	0.49	0.51	0.20
PC_{ext}	564	0.44	0.41	0.20

Panel B. a

Variable	N	Mean	Median	σ
II	270	−10.4	−2.2	29.20
AA	270	2.60	0.7	19.40
PC	270	−1.1	−0.51	1.93
PC_{ext}	564	−7.65	−1.10	17.85

Panel C. b

Variable	N	Mean	Median	σ
II	270	778.5	472.85	765.52
AA	270	411.73	244.27	555.35
PC	270	36.95	22.80	41.45
PC_{ext}	564	36.94	22.80	41.44

The table presents summary statistics for the recursive estimation of the *degree of extrapolative weighting* (DOX) parameter. Specifically, for every month in the interval [t-12,t-1], which we refer to as the cross-validation period, we estimate the following model of extrapolative expectations using three alternative rolling window sizes of 24, 36, and 48 months, as well as an expanding window with a starting length of 36 months:

$$Exp_t = a + b \sum_{j=0}^{59} w_j R^Q_{t-j-1,t-j} + \epsilon^{Exp}_t,$$

where Exp_t is the survey-based expectation of future returns, $R^Q_{t-j-1,t-j}$ is the j-lagged quarterly return of the aggregate value-weighted stock market from CRSP, and w_j is a weight parameter whose expression is

$$w_j = \frac{\lambda^j}{\sum_{k=0}^{59} \lambda^k}.$$

During the cross-validation period, we assess the one-step-ahead mean-squared-forecast error (MSFE) of our model for each of the four alternative window sizes listed above. We then estimate model parameters for month t using the 4 four alternative windows, and finally calculate a weighted average of the DOX estimates obtained with each window size, where the weights assigned are proportional to the inverse of the MSFE obtained in the cross-validation period. We use the same weights to combine the center and scale parameters a and b. The recursive estimation is conducted for II, AA, and their principal component (PC). A recursively estimated DOX is available for the period 1992:06–2014:12. Results are also reported for an extended version of the principal component (PC_{ext}), that uses II-based parameter estimates in the period 1967:12–1992:05, and the PC-based parameter estimates afterward.

as well as in tests of parameter instability in the predictability relation, which are conducted in sections 4.6 and 4.7, respectively.

Figure 2 plots the estimated PC DOX time series. The graph shows considerable variation in DOX over time. DOX estimates span the entire range of the coefficient, which can only lie between 0 and 1. Additionally, the DOX time series appears to move in lockstep with salient events in the recent history of the stock market. For instance, the DOX progressively increases in the decade leading up to the dot-com bubble burst, reaches a peak during the first half of 2000, and later declines back to its prebubble levels by the end of 2004.

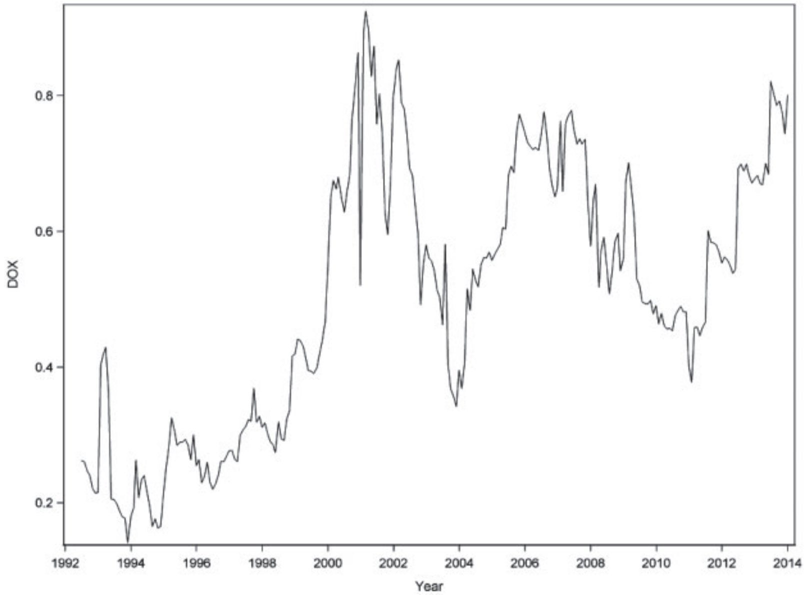

Figure 2
Time-varying DOX
The figure plots the recursively estimated DOX time series for the principal component (PC) of the II and AA surveys, in the sample period 1992:06–2013:12. (Details on estimation are in Section 2.)

The extrapolators' DOX reaches another peak in 2007, right before the Great Recession, and declines again by mid-2010.

4.2 Potential determinants of DOX

In this section, we follow up on our finding above that DOX changes over time. The first potential explanation for this variation is the time-varying composition of the pool of stock market investors. As the types of investors who participate in the stock market change over time, the consensus DOX may change over time, if various investors rely on recent versus older stock market return realizations differently. young individuals, who have experienced a shorter return history, may place more weight on recent returns (higher DOX), relative to the weight assigned to such returns by older individuals.[13] When the participation rate of younger individuals increases, and their presence in the market relative to the older investors grows, the average DOX also increases. On the other hand, DOX may also change due to salient features of any given return realization, as well as the features of the recent distribution of stock returns, or the state of the economy. Importantly, the stock market composition and salience based

[13] Recent literature suggests that age and lifetime experiences may play a key role in expectations formation (Nagel 2012; Malmendier and Nagel 2015) and in portfolio allocation decisions (Malmendier and Nagel 2011).

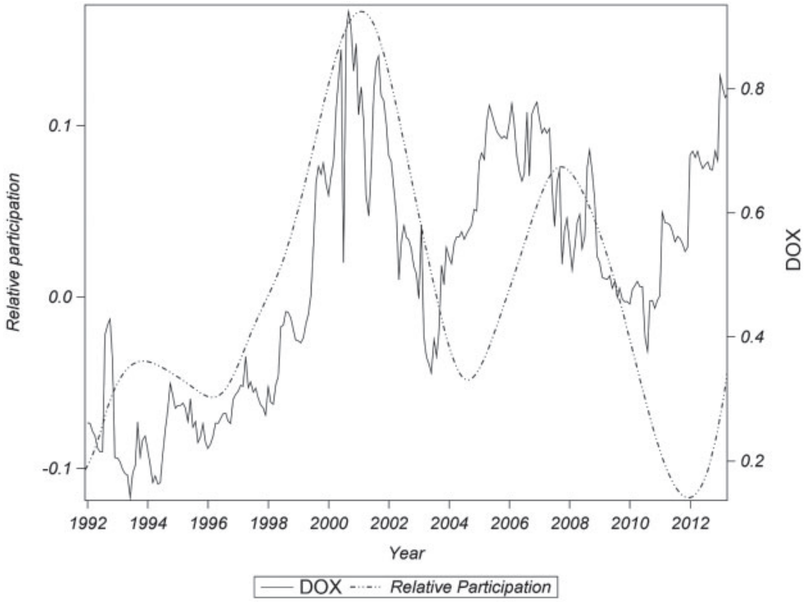

Figure 3
Potential determinants of DOX: stock market participation
In the sample period 1992:06–2013:12, we recursively estimate the extrapolators' DOX from survey data. During the same period, we report the relative participation of the young versus the old in the stock market.

explanations may not be mutually exclusive, since investors' decision to enter or exit the stock market may itself be rooted in salient information about stocks and stock returns.

Figure 3 provides graphical evidence in support of the role played by stock market participation demographics. In the figure, we overlay extrapolators' DOX with the ratio of the number of young (50 years of age or younger) to old (older than 50 years of age) investors with direct holdings of stocks.[14] The data are converted to monthly frequency by spline interpolation and detrended. Figure 4, compares instead DOX with the level of the S&P 500. We note that not only does DOX seem to respond to an increase in the fraction of young stock market investors whose expectations may be characterized by higher DOX, but also that DOX increases with stock prices. The lead-lag relation between stock prices and DOX, and the comovement of DOX and stock market demographics, point to an interplay of salience and stock market demographics. For instance, when the market starts climbing, like in the late 1990s, short-term extrapolators (possibly the young) may be compelled to enter

[14] Data on stock market participation are from the Survey of Consumer Finances (SCF) http://www.federalreserve.gov/econrDOXata/scf/scfindex.htm. Data on population demographics are from Census https://www.census.gov/hhes/families/data/households.html.

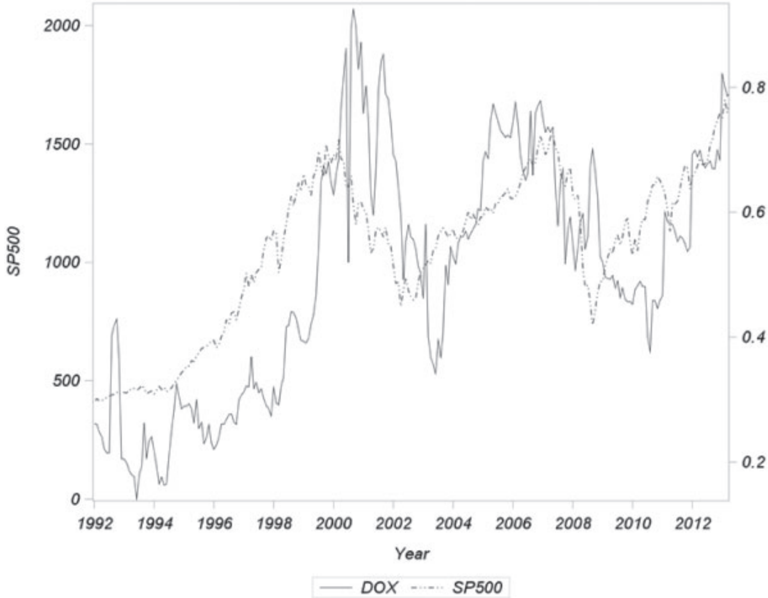

Figure 4
Potential determinants of DOX: Stock prices
In the sample period 1992:06–2013:12, we recursively estimate the extrapolators' DOX from survey data. The figure compares DOX with the level of the S&P 500 index.

the market, pushing the consensus DOX up. When the market collapses, they become bearish and exit, resulting in an ensuing decline in DOX. This cycle repeats with the boom and bust cycle from 2004 to 2008.

We complement this observation with formal tests. In these tests, a large set of age-based as well as salience-rooted potential determinants of DOX is considered. Since the observed patterns in stock market participation may be driven by past returns, we construct a 5-year cumulative stock market returns variable, lagged one year. Variation in DOX may also be linked to changes in the perceived stability of the stock market. We attempt to measure such stability by counting the number of months since the last stock market crash. A crash is defined as a quarter in which aggregate stock returns fall more than two standard deviations below their unconditional mean. To investigate the link between states of the economy and DOX, we use an NBER recession dummy, as well as a measure of time elapsed since the last recessionary period. To further study whether DOX changes with perceived uncertainty in stocks, we form two proxies for volatility. The first measures the 5-year trailing volatility of monthly returns, while the second estimates intra-quarter volatility from daily stock returns like in French, Schwert, and Stambaugh (1987). A related possibly salient feature of the return from stocks is return extremeness, which speaks more directly to attention-grabbing events and may therefore induce

Table 4
Potential determinants of the DOX

Coefficient	(1)	(2)	(3)	(4)	(5)	(6)
INT	0.001	0.001	−0.001	−0.012	0.001	−0.001
	[0.055]	[0.059]	[−0.004]	[−0.56]	[0.04]	[−0.220]
REL_PART	1.35					0.89
	[5.49]					[4.03]
PART_YOUNG		0.066				
		[3.43]				
PART_OLD		−0.029				
		[−1.40]				
CUM5YR_RET_LAG			0.115			−0.038
			[2.39]			[−1.25]
REC				0.094		0.014
				[1.10]		[0.19]
TIME_FROM_REC				0.001		0.001
				[0.99]		[2.14]
TIME_FROM_CRASH				0.001		0.001
				[3.45]		[1.51]
MONTHLY_VOL					4.22	8.26
					[2.28]	[5.80]
QVOL					4.85	0.68
					[1.88]	[0.83]
EXTRQRET					−0.29	0.113
					[−1.72]	[1.11]
SENT						0.111
						[4.02]
Adj.R^2 (%)	49.07	44.37	18.74	32.30	8.52	68.32

Regressions of DOX on a set of potential explanatory variables. PART_YOUNG is the stock market participation rate of families with a head of household 50 years of age or below. PART_OLD is the stock market participation rate of families with a head of household above 50 years of age. The relative participation of young to old REL_PART is the ratio of number of the young to old families with direct holdings of stocks in the US. CUM5YR_RET_LAG is the lagged 5-year stock market cumulative return. REC is an economic recession dummy. TIME_FROM_REC and TIME_FROM_CRASH measure the time elapsed since the end of the last economic recession and a stock market crash, respectively. A stock market crash is a quarterly return realization which is two standard deviations below its unconditional mean. MONTHLY_VOL measures the 5-year volatility of monthly return, while quarterly return volatility QVOL is measured from daily returns within a quarter like in French, Schwert, and Stambaugh (1987). Extreme quarter return EXTRQRET is a dummy variable that tracks quarterly returns that are more than two standard deviations away from the unconditional mean. SENT is the Baker and Wurgler (2006) investor sentiment. Return-based variables are from CRSP. Participation-related variables are constructed jointly from the Survey of Consumer Finances and Census data on household demographics. Data on economic recessions is from the National Bureau of Economic Research, and sentiment is from Jeffrey Wurgler's Web site. All variables are detrended. Results are based on the DOX extracted from the principal component of AA and II, in the sample period 1992:06–2013:12. t-statistics in brackets are robust to heteroscedasticity and serial correlation.

changes in DOX. We construct extremeness as a dummy variable that is equal to one when a quarterly return is more than two standard deviations away from its unconditional mean, and zero otherwise. As a final check of the relationship between DOX and prior literature on irrationality in the equity market, we include the Baker and Wurgler (2006) measure of investor sentiment in our set of explanatory variables. We work with detrended data used in Nagel (2012). Our main analysis refers to the sample period 1992:06–2013:12, when both II and AA are available, and a DOX time series extracted from their principal component can be constructed.

The results of our analysis are reported in Table 4. In univariate regressions, we find support for the hypothesis that the variation in DOX is related to the variation in the stock market participation rate of young versus old investors. In Column 1, we document that when the number of young investors in the stock market increases relative to the number of older investors, the DOX increases.

We then decompose this overall effect by regressing DOX on the participation rates of the young and the old. In Column 2, we show that when controlling for the participation rate of the young, an increase in the participation of the old corresponds to a decline in DOX. On the contrary, holding the participation rate of the old constant, an increase in the participation of the young is associated to higher DOX. Column 3 shows that DOX increases following a recent period of good stock returns. This is consistent with higher return salience after prolonged positive returns, which induces high-DOX individuals to enter the stock market. In Column 4 we document that DOX increases significantly with the time since a stock market crash. However, DOX is not significantly different from its unconditional mean during recessions, and it only marginally increases with the time since the end of the most recent recession. These findings suggest that DOX is high entering into a recession, declines afterwards, and then increases again as time from the recession passes. Column 5 shows that the relationship with volatility is positive after controlling for extreme events, which seems to result in a declining DOX. In Column 6, where we run a multivariate regression, we confirm the link to stock market demographics, as well as the association with sentiment and a positive link with past stock market volatility.

In conclusion, it appears that the variation in DOX captured recursively using survey data reflects changes in the participation of different groups of investors in the stock market. Such changes may reflect underlying dynamics that generate procyclical stock market participation by high-DOX investors (possibly the young). DOX is also sensitive to recent volatility and appears related to previously documented forms of sentiment about stocks. These results notwithstanding, one needs to interpret our findings with caution, since they point to association, not causation. Shedding light on the causal drivers of extrapolators' DOX is a challenging task that is beyond the scope of this paper.

4.3 Conditional stock return predictability

Once the time-series variation in the extrapolators' DOX is unveiled, we can run a formal test of the hypothesis that the extrapolation bias may be a determinant of stock return predictability by price-scaled variables. To this end, we estimate the conditional model in Equation (6). We focus on excess return predictability, and Table 5 presents the main results. The main sample period is 1992:06–2013:12. Panel A refers to the use of the D/P ratio, and in panels B and C, we replace it with the B/M and the E/P ratio, respectively. We perform tests for prediction horizons of three and 12 months, and run predictive regressions both with and without controls.[15] The controls include a deterministic trend and the squared D/P ratio. Table A1 in the Appendix extends the set of controls to include a larger set of competing state variables. We show results for the

[15] For brevity, we report 3-month prediction horizon results only in panel A and focus on the 12-month results in the remainder of the paper. Results for the 3-month prediction horizon are available on request.

Table 5
Conditional stock return predictability with extrapolation

Panel A. D/P

Horizon (months)	Coefficient	Baseline	II		AA		PC		PC_{ALL}	
3	a_0	−0.037	0.248	1.134	0.182	1.394	0.202	0.291	0.241	1.306
		[−1.237]	[2.87]	[1.129]	[2.883]	[1.36]	[3.187]	[0.222]	[3.255]	[1.224]
	a_1		−0.616	−0.746	−0.368	−0.235	−0.448	−0.514	−0.57	−0.439
			[−3.88]	[−2.721]	[−2.997]	[−1.988]	[−3.639]	[−2.584]	[−3.61]	[−2.732]
	b_0	2.847	−9.921	−21.553	−5.738	−52.043	−7.237	6.784	−8.358	−47.534
		[1.824]	[−2.645]	[−0.45]	[−1.94]	[−1.13]	[−2.787]	[0.111]	[−2.448]	[−0.971]
	b_1		29.315	38.944	14.523	7.365	19.528	25.649	22.925	16.562
			[3.984]	[2.665]	[2.182]	[1.214]	[3.08]	[2.352]	[2.829]	[2.17]
	$Adj.R^2$ (%)	2.871	12.775	19.076	14.451	14.494	12.782	12.569	12.735	13.009
	Controls	N	N	Y	N	Y	N	Y	N	Y
12	a_0	−0.186	0.467	6.322	0.757	3.884	0.694	2.998	0.677	5.951
		[−1.837]	[1.755]	[3.478]	[6.269]	[1.829]	[5.325]	[1.528]	[4.638]	[3.216]
	a_1		−1.355	−0.977	−1.61	−1.295	−1.637	−1.344	−1.775	−1.117
			[−2.825]	[−1.873]	[−8.35]	[−4.99]	[−7.054]	[−3.802]	[−6.731]	[−3.42]
	b_0	13.654	−14.459	−209.552	−26.634	−132.259	−23.003	−98.645	−22.079	−221.859
		[2.859]	[−1.258]	[−2.77]	[−4.529]	[−1.641]	[−3.437]	[−1.116]	[−2.575]	[−3.178]
	b_1		61.372	48.814	69.431	53.464	70.264	56.953	73.367	40.616
			[2.633]	[1.606]	[6.487]	[4.06]	[5.048]	[2.546]	[4.03]	[2.238]
	$Adj.R^2$ (%)	16.04	22.965	39.256	49.926	54.926	43.181	45.04	35.764	45.64
	Controls	N	N	Y	N	Y	N	Y	N	Y

Panel B. B/M

Horizon (months)	Coefficient	Baseline	II		AA		PC		PC_{ALL}	
12	a_0	−0.157	0.57	8.445	0.849	5.841	0.741	4.898	0.918	7.741
		[−1.398]	[2.354]	[1.999]	[5.635]	[1.843]	[4.711]	[1.313]	[5.026]	[2.218]
	a_1		−1.43	−1.013	−1.705	−1.597	−1.657	−1.538	−2.233	−1.747
			[−3.656]	[−2.307]	[−6.608]	[−5.641]	[−5.046]	[−3.938]	[−6.443]	[−4.836]
	b_0	0.886	−1.452	−23.853	−2.319	−16.21	−1.908	−14.384	−2.522	−22.036
		[2.545]	[−1.777]	[−2.026]	[−5.07]	[−1.818]	[−3.435]	[−1.416]	[−4.149]	[−2.198]
	b_1		4.59	3.47	5.464	5.445	5.114	4.994	7.041	5.633
			[3.17]	[2.29]	[7.057]	[5.941]	[4.486]	[3.664]	[5.894]	[4.27]
	$Adj.R^2$ (%)	10.591	26.68	40.398	42.743	50.685	45.411	48.755	35.737	45.952
	Controls	N	N	Y	N	Y	N	Y	N	Y

(*continued*)

DOX extracted from II, AA, and their principal component, PC. To show that the results extend to the use of other surveys, we also perform conditional tests for the DOX obtained from an alternative principal component, PC_{ALL} which incorporates other investor expectations surveys as they become available over time.[16] Since we use monthly observations, one concern is serial correlation in the error-terms due to the overlapping nature of our return observations. We address this issue by adopting Newey-West (1987) standard errors with 3 and 12 lags. Finally, Table 5 reports the result of a baseline univariate predictability regression for each sample period. This replicates prior findings of stock return predictability, and hence is the natural benchmark for our new model.

Panels A through C of Table 5 present four main findings. First, the coefficient estimate on the interaction term, b_1, is always positive and statistically

16 These are the UBS/Gallup Survey, which began in 1996:10 (with gaps), and the World Economic survey for the U.S. stock market from the Ifo Institute, which began in 1998:03 at a quarterly frequency.

Table 5
Continued

Panel C. E/P

Horizon (months)	Coefficient	Baseline	II		AA		PC		PC_{ALL}	
12	a_0	−0.235	0.69	15.624	1.076	11.302	0.905	11.941	0.993	14.532
		[−1.998]	[1.961]	[3.538]	[5.474]	[3.055]	[4.401]	[2.175]	[3.43]	[3.638]
	a_1		−1.807	−0.994	−2.209	−1.632	−2.075	−1.576	−2.54	−1.581
			[−2.993]	[−1.68]	[−6.52]	[−4.81]	[−5.462]	[−3.398]	[−4.401]	[−2.747]
	b_0	7.78	−12.442	−290.577	−21.184	−208.92	−16.794	−210.27	−19.047	−274.92
		[3.153]	[−1.532]	[−3.483]	[−4.838]	[−2.868]	[−3.297]	[−2.046]	[−2.671]	[−3.522]
	b_1		40.083	24.19	49.008	37.754	45.034	39.814	55.453	36.722
			[2.652]	[1.604]	[6.451]	[4.558]	[4.725]	[3.735]	[3.787]	[2.464]
	$Adj.R^2$ (%)	15.806	27.103	54.061	49.221	60.023	46.122	55.586	37.355	54.514
	Controls	N	N	Y	N	Y	N	Y	N	Y

Panel D. Economic magnitudes

Predictor	DOX	Conditional coefficient $b_t = b_0 + b_1 \text{DOX}_t$	t-stat	p-value
D/P	0.31	−1.924	[−0.535]	(.7)
	0.51	12.128	[3.638]	(.00)
	0.71	26.181	[5.261]	(.00)
B/M	0.31	−0.374	[−1.554]	(0.94)
	0.51	0.649	[4.500]	(.00)
	0.71	1.672	[5.649]	(.00)
E/P	0.31	−3.284	[−1.384]	(0.92)
	0.51	5.723	[5.282]	(.00)
	0.71	14.73	[7.381]	(.00)

Using GMM we estimate the following monthly time-series regression of l-month-ahead aggregate U.S. excess stock returns $R^e_{t_0,t_0+l}$:

$$R^e_{t_0,t_0+l} = a_0 + a_1 \text{DOX}_{t_0} + c_0' X_{t_0} + D/P_{t_0}[b_0 + b_1 \text{DOX}_{t_0} + c_1' X_{t_0}] + \epsilon^R_{t_0,t_0+l},$$

where D/P is the aggregate dividend-price ratio, and DOX is the degree of extrapolative weighting extracted from a survey of U.S. stock market expectations. Results are reported for DOX estimated from II, AA, and their principal component (PC), as well as an alternative principal component (PC_{ALL}) which incorporates other surveys as they become available (details are in Section 4). Panel A reports the results for a prediction horizon of three and 12 months, and includes the univariate predictability regression as a baseline specification, as well as conditional predictability regressions with and without controls, X_{t_0}. Such controls include a deterministic trend and the squared D/P ratio. The main sample is 1992:06–2013:12. t-statistics are based on Newey-West standard errors with a lag length of l months. Panel B and C replace D/P with B/M and E/P, respectively and report 12-month-ahead predictability tests. Panel D reports estimates and statistical significance of the one-year conditional coefficient of predictability, $b_t = [b_0 + b_1 DOX_t]$, for the sample period 1992:06–2013:12, and the DOX extracted from the PC time series. Standard errors for b_t are constructed following Aiken and West (1991). p-values refer to the one-sided null hypothesis of zero or negative b_t. Further details on data construction are in Section 3.

significant. This is consistent with the model of BGJS and our hypothesis that stock return predictability by price-scaled variables increases with DOX.

Second, when moving from the standard univariate regression to the conditioning predictability model, there is a considerable improvement in goodness of fit, as captured by the adjusted-R^2.[17] For example, at the year horizon, a predictive regression that features the D/P ratio alone has an adjusted-R^2 of 16%, while the conditional specification that uses the DOX data from the II series brings the goodness of fit to 23%. The improvement is even stronger when the AA or the PC DOX is used as a conditioning variable. In these cases, the adjusted-R^2 increases to 50%. Third, when return predictability

[17] In unreported tests we also run a nested model, where $b_0 = 0$, that is, no stock return predictability when DOX is zero. Results are unchanged.

is assessed in the traditional univariate regression framework, the resultant predictability coefficient always lies somewhere between the minimum (b_0) and maximum (b_0+b_1) values of the conditional predictability coefficient($b_t = b_0+b_1\text{DOX}_t$). Intuitively, given a theory that posits different degrees of stock return predictability across DOX states, a univariate regression measures an average relationship that includes episodes of high and low DOX. In so doing, the use of the univariate predictive regression framework may understate or overstate the extent of predictability at any given time. Fourth, in 15% of our monthly predictions of year-ahead aggregate excess returns, our conditional model predicts a subsequent negative equity premium. The negative equity premium prediction delivered by the conditional predictability model arises in cases of high DOX and low D/P ratio (i.e., in an overvalued market with highly transitory beliefs). In these instances, we test the null hypothesis that such prediction is positive or zero against the alternative of a negative equity premium prediction, and we reject it at the 1% level in 70% of the cases.[18] A statistically significant negative equity premium prediction by low D/P ratio in high DOX states and the subsequent realized negative excess returns is hard to reconcile with rational models of risk in which the ex ante equity premium is positive. Similar results are obtained using other price-scaled predictors. Panels A to C also show that when we introduce controls or extend the set of used surveys using PC_{ALL}, the obtained predictability results are qualitatively similar. In panel D of Table 5, we follow Aiken and West (1991), and present the conditional coefficient of predictability $b_t = [b_0+b_1\text{DOX}_t]$ as well as its related t-statistic for a set of representative values of the DOX state variable. In what follows, we refer to high DOX as a value of the degree of extrapolative weighting that is one standard deviation above the DOX sample median, and we refer to a low DOX as a value that is one standard deviation below the DOX sample median. For brevity, panel D only focuses on the DOX extracted from the principal-component series, and similar results are obtained when the DOX from II or AA is used.

The results are striking, and confirm that the predictive ability of price-scaled variables is contingent on extrapolators' DOX. When the DOX is high, a one-standard-deviation (0.01) increase in the D/P ratio is followed by a statistically significant 26% increase in the expected equity premium the following year. When instead the DOX is low, a one-standard-deviation increase in the D/P ratio is *negatively*, but insignificantly related to future returns, and predicts a 2% lower equity premium in the upcoming year. Similar results to the ones reported in panel A are obtained using other price-ratios in panels B and C. For instance, when we replace the D/P ratio with the B/M ratio, a one-standard-deviation (0.27) rise in B/M predicts a 47% higher excess return the next year

[18] See Baker and Wurgler (2000) for a similar test.

when the DOX is high, while the same rise corresponds to a 16% decline in future returns if the DOX is low.

In conclusion, Table 5 shows that the the D/P ratio, as well as other price-scaled predictors, forecasts future returns differently across DOX-states, and in a statistically significant way only when the DOX is high.

To complement our discussion of the results in Table 5, Figure 5 presents evidence of conditional one-year predictability.[19] In the figure, we provide scatter plots of the relationship between D/P, B/M, or E/P and future 12-month excess returns, conditioned on values of the DOX parameter that are either below (left panel) or above (right panel) the media DOX. We focus on the PC DOX time series. The figure shows a large wedge between stock return predictability in states of high and low DOX. For instance, there is a strong and positive association between current the D/P ratio and future aggregate returns in high DOX states, in which variation in the D//P ratio accounts for 35% (panel A) of variation in the subsequent equity premium. On the other hand, there is little or no predictability in states of low DOX, where the R^2 is only 0.7%. Similar results can be inferred for the other price-ratios considered. Therefore, the figure supports the findings in Table 5.

While the results presented here are consistent with a behavioral interpretation of the evidence of stock return predictability, one needs to interpret them with caution. The combination of a short sample period, persistent regressors, high contemporaneous correlation between price-scaled predictors and future returns, as well as overlapping observations, may raise concerns of small sample bias and size distortions in our inference. Later we address these issues via bootstrapping techniques and show that, while the aforementioned concerns are valid, correcting for small sample bias alters neither the sign nor the magnitude of our coefficient of interest on the interaction of DOX and a given price-scaled predictor.

In conclusion, our results suggest that the time-series predictability of stock returns by price-scaled variable is a phenomenon contingent on the DOX and on the time-varying tendency of extrapolators to predominantly rely on too few recent return realizations. Furthermore, the traditional univariate approach to time-series predictability of stock returns only measures an average relationship between D/P and future returns, and the strength of this relation improves when DOX is high.

4.4 Mean reversion in price-scaled predictors

In Section 2 we argue, based on the Campbell and Shiller (1988) decomposition, that the predictability of aggregate excess stock returns by the D/P ratio should decline as the D/P ratio becomes more persistent. The results of our

[19] Figure A1 replicates Figure 5 with annual-frequency data and year labels.

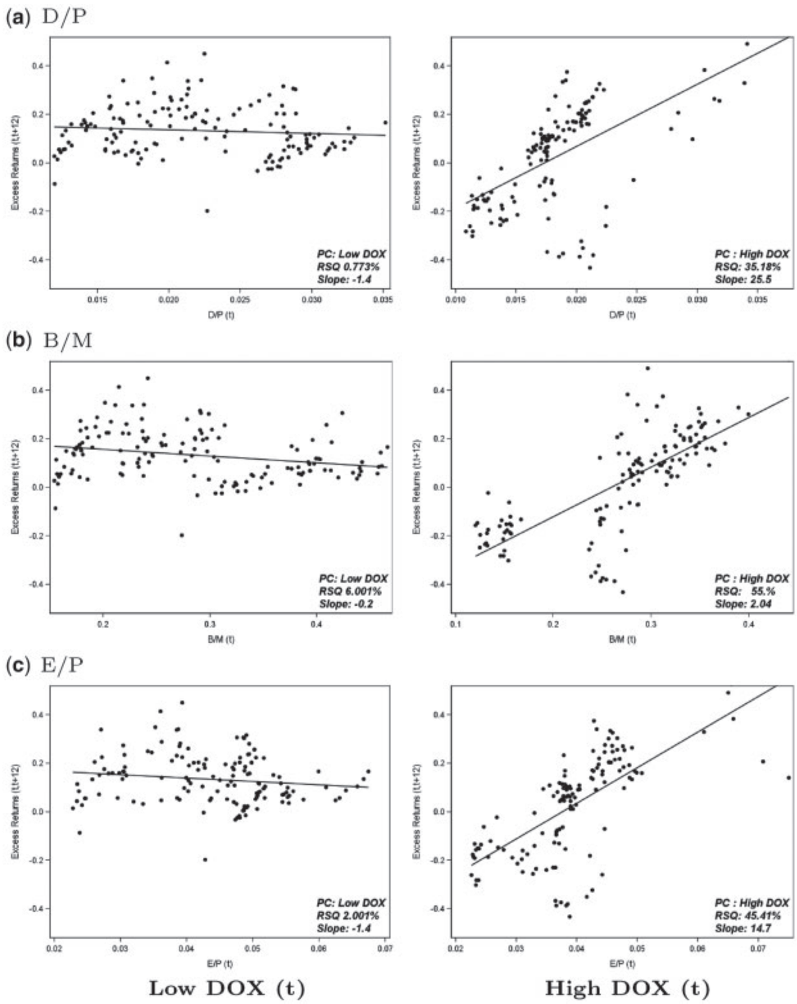

Figure 5
Conditional stock return predictability
In the period 1992:06–2013:12, we classify each month *m* based on the degree of extrapolative weighting (DOX) extracted from the principal component (PC) of the II and AA surveys. The left panel reports a scatter plot of the relationship between the D/P ratio (panel A), the B/M ratio (panel B), and the cyclically adjusted E/P ratio (panel C) in month *m* and the subsequent 12-month stock market excess returns, when the DOX in month *m* is low (i.e. below its median value of 0.51). The right panel reproduces a similar scatter plot for the months in which the DOX is high (i.e. above the median). Each figure also reports R^2 and coefficient of predictability of the corresponding univariate predictive regression.

conditional predictability regressions suggest that the predictive ability of the D/P ratio is apparent only in states of high DOX. Therefore the D/P ratio should exhibit different persistence across DOX-states. It is expected to be more (less) persistent in low (high) DOX states.

$$D/P_{t+l} = (a_0 + a_1 \mathrm{DOX}_t) + D/P_t(b_0 + b_1 \mathrm{DOX}_t) + \epsilon^{dp}_{t,t+l} \tag{8}$$

In particular, the null and alternative hypotheses are

$$H_0 : b_1 = 0 \quad H_a : b_1 < 0. \tag{9}$$

Under the null hypothesis, the coefficient b_1 on the interaction term in Equation (8) is positive or zero. The alternative hypothesis postulates that as the DOX increases, the D/P ratio reverts to the mean more quickly, hence the negative sign of the coefficient b_1.

Panel A of Table 6 presents the results at a yearly forecast horizon for the sample period 1992:06–2013:12. In panel B, we provide further evidence by reporting the half-life of a shock to a price-scaled variable in states of high versus low DOX. Like in prior tables, we refer to high DOX as the extrapolators' DOX that is one standard deviation above its sample median, and similarly we define a low DOX as one standard deviation below its sample median.

The table results warrant a rejection of the null hypothesis, since in all cases the coefficient on the interaction term is negative, and in all but one case, the coefficient is statistically significant at the 5% level. Panel B reveals that shocks to price-scaled variables under a low DOX are persistent, and their half-life has a median value of five years. When in states of high DOX, shocks to price-scaled variables appear to be reabsorbed much more quickly, since their half-life decreases to approximately 10 months.[20]

Based on the results in Tables 5 and 6, and consistent with the intuition from the Campbell and Shiller (1988) model, we can establish a parallel between the autoregressive behavior of price-scaled predictors and their ability to forecast future returns: an increase in extrapolation bias linked to DOX is associated with both better predictability and quicker mean reversion in predictor variables. Conversely, if the DOX is low, price-scaled predictors exhibit a high autocorrelation, which reduces their ability to predict future equity premium.

4.5 Out-of-sample results

In this section, we investigate whether the in-sample evidence on the conditioning role of extrapolators' DOX for stock return predictability is confirmed out-of-sample. Consider the three forecasting models M_0, M_1, and

20 In unreported results, we also bootstrap the regression in Equation (8) to correct for the Kendall (1954) small sample attenuation bias of the autoregressive coefficient b_0. As expected, correcting for the bias increases the AR(1) coefficient estimate further, whereas it has little to no effect on the interaction coefficient b_1, thus lending even further support to our findings.

Table 6
Price-scaled predictors: conditional mean-reversion with extrapolation

		Panel A. Conditional mean reversion					Panel B. Half-life (months)	
Survey	predictor	a_0	b_0	a_1	b_1	$Adj.R^2(\%)$	Low DOX	High DOX
II	D/P	0.00	0.926	0.012	−0.571	55.2	29	13
		[0.054]	[3.978]	[1.059]	[−0.907]			
					(0.18)			
	B/M	−0.03	0.996	0.256	−0.734	48.4	32	11
		[−0.534]	[5.104]	[2.112]	[−1.626]			
					(0.05)			
	E/P	−0.009	1.175	0.045	−1.054	50.1	51	10
		[−0.726]	[4.281]	[2.037]	[−2.006]			
					(0.02)			
AA	D/P	−0.01	1.363	0.028	−1.258	57.5	77	9
		[−2.837]	[8.977]	[3.861]	[−3.633]			
					(0.00)			
	B/M	−0.033	1.077	0.237	−0.837	41.9	31	10
		[−0.43]	[4.094]	[1.601]	[−1.622]			
					(0.05)			
	E/P	−0.022	1.365	0.059	−1.164	56.3	122	11
		[−3.089]	[7.37]	[4.403]	[−3.308]			
					(0.00)			
PC	D/P	−0.007	1.216	0.026	−1.1	55	62	10
		[−2.022]	[6.757]	[2.89]	[−2.138]			
					(0.017)			
	B/M	−0.043	1.029	0.271	−0.774	44.6	35	11
		[−0.715]	[4.852]	[2.024]	[−1.662]			
					(0.05)			
	E/P	−0.017	1.271	0.054	−1.079	54.2	127	12
		[−2.213]	[6.909]	[3.645]	[−2.846]			
					(0.00)			

The table reports GMM estimation results for the following monthly conditional autoregressive model:

$$Predictor_{t+12} = a_0 + a_1 * \mathrm{DOX}_t + Predictor_t [b_0 + b_1 \mathrm{DOX}_t] + \epsilon_t^P.$$

The *predictor* variable is one of the following aggregate price-scaled variables: dividend-price (D/P) ratio, the book-to-market (B/M) ratio, or the cyclically adjusted earnings-to-price (E/P). The DOX state variable is extracted from II, AA, or their principal component (PC).The sample period is 1992:06–2012:12. Newey-West t-statistics with six lags are reported in brackets. p-values for the one-sided test of zero of positive b_1 are reported in parenthesis. In panel B we present the implied half-life of a shock to the predictor in two states: low DOX (DOX one standard deviation lower than the median) and high DOX (DOX one standard deviation above the median).

M_2 defined below:

$$M_0 : R^e_{t_0,t_0+12} = \mu_{t_0} + \epsilon^{M_0}_{t_0,t_0+12}$$

$$M_1 : R^e_{t_0,t_0+12} = a_0 + b_1 \frac{D}{P}_{t_0} + \epsilon^{M_1}_{t_0,t_0+12}$$

$$M_2 : R^e_{t_0,t_0+12} = (a_0 + a_1 \mathrm{DOX}_{t_0}) + \frac{D}{P}_{t_0} (b_1 + b_2 \mathrm{DOX}_{t_0}) + \epsilon^{M_2}_{t_0,t_0+12}$$

M_0 is the naive forecasting model that uses the historical average excess return as a forecast of the future return. M_1 is the univariate prediction specification addressed by Goyal and Welch (2008), who document its poor performance in pseudo out-of-sample tests. M_2 is the augmented forecasting specification that is the focus of this study.

First, we follow Clark and West (2007), who develop a one-tailed test of equal predictive accuracy of nested linear specifications well-suited for our case. To implement the Clark and West test, every month m we estimate each of the models using all observations available between the date of the first DOX estimate and m. We then use the estimated coefficients in conjunction with the value of the right-hand-side variables in month *(m+12)*, to forecast excess returns over the interval *(m+12,m+24)*. The first out-of-sample month is 1997:06, so as to allow a five-year period for the first in-sample parameter estimation. Using the forecast errors (FE) of a given model, Clark and West's test statistic for comparing model M_i and M_j is the intercept obtained by regressing $(2FE^i_{m+12}[FE^i_{m+12}-FE^j_{m+12}])$ on a constant. This regression-based approach allows us to produce standard errors that are corrected for heteroscedasticity and serial correlation.

In panel A of Table 7, we compare models M_i and M_j ($M_i\, vs\, M_j$) and test the one-sided null hypothesis of equal prediction accuracy of M_i and M_j versus the alternative of higher accuracy of model M_j. A positive value of the statistic suggests that M_j fares better than M_i. Like in prior literature, we find that the standard univariate prediction model never beats the historical average in a statistically significant way. As for the conditional prediction model M_2, the results confirm that it provides better prediction accuracy than its univariate counterpart. In all cases, the improvement in forecasting accuracy is always statistically significant at least at the 5% level. Additionally, the conditional model also provides a statistically significant accuracy improvement over the naive forecast.

The next piece of out-of-sample evidence is based on the out-of-sample R^2 proposed in Campbell and Thompson (2008) and Rapach, Strauss, and Zhou (2010). It is calculated as follows:

$$R^2_{OOS_{M_i vs M_j}} = 1 - \frac{\sum_{t=m_0}^{T}[(r_{t,t+12}) - r^{\hat{M_j}}_{t,t+12}]^2}{\sum_{t=m_0}^{T}[(r_{t,t+12}) - r^{\hat{M_i}}_{t,t+12}]^2}, \tag{10}$$

where $r_{t,t+12}$ is the ex post 12-month cumulative returns in the interval (t,t+12), and $r^{\hat{M_j}}_{t,t+12}$ is the out-of-sample (OOS) prediction of model M$_j$. When $R^2_{OOS_{M_i vs M_j}}$ is positive, M$_j$ fares better relative to model M$_i$. When $R^2_{OOS_{M_i vs M_j}}$ is close to 0, M$_i$ and M$_j$ have comparable OOS performance. Negative values of the statistic point to a decline in accuracy when migrating from M$_i$ to M$_j$. Panel B of Table 7 presents results of pairwise (i, j) comparison. Results confirm prior findings on the weak ability of univariate forecasting models to subsume the historical average. The OOS R^2 statistic is approximately 1%. At the same time,the DOX-based conditional model outperforms both the historical average and the univariate model.

Next, we measure the economic gains introduced by using M_2 instead of M_1, or M_0. One approach is to quantify the economic gains reaped by an expected utility maximizer when she moves from the traditional univariate forecasting models M_1, or the naive forecast M_0, to the conditional model M_2.

Table 7
Out-of-sample tests

Panel A. Test of equal prediction accuracy

Horizon	D/P			B/M			E/P		
(months)	M0vsM1	M0vsM2	M1vM2	M0vsM1	M0vsM2	M1vM2	M0vsM1	M0vsM2	M1vM2
12	3.7	49.551	48.92	3.652	53.477	56.243	6.619	55.453	57.94
	[0.846]	[1.843]	[1.706]	[0.978]	[2.027]	[1.904]	[1.278]	[2.027]	[1.828]

Panel B. Out of sample R^2

12	1.247	21.525	20.534	−0.866	36.200	36.748	1.876	28.927	27.568

Panel C. Realized economic gains

Horizon		Realized economic gains			Realized mean returns			Sharpe ratio			End-of-period wealth		
(months)	Predictor	M0vsM1	M0vsM2	M1vsM2	M0	M1	M2	M0	M1	M2	M0	M1	M2
12	D/P	0.01	7.313	7.303	5.051	6.19	11.789	0.127	0.164	0.433	150	172	361
	B/M	1.636	10.925	9.289	5.051	9.007	14.627	0.127	0.26	0.591	150	257	662
	E/P	1.661	8.915	7.254	5.051	9.207	13.126	0.127	0.265	0.498	150	237	437

The table presents results of out-of-sample tests for the following forecasting models:

$$M_0 : R^e_{t_0,t_0+12} = \mu + \epsilon^{M_0}_{t_0,t_0+12},$$

$$M_1 : R^e_{t_0,t_0+12} = a_0 + b_1 Predictor_{t_0} + \epsilon^{M_1}_{t_0,t_0+12},$$

$$M_2 : R^e_{t_0,t_0+12} = (a_0 + a_1 \mathrm{DOX}_{t0}) + Predictor_{t_0}(b_1 + b_2 \mathrm{DOX}_{t_0}) + \epsilon^{M_2}_{t_0,t_0+12}.$$

The dependent variable $R^e_{t_0,t_0+12}$ is the 12-month-ahead cumulative excess return. The *Predictor* variable is one of the following: the dividend-price ratio (D/P), the book-to-market ratio (B/M), or the cyclically adjusted earnings-to-price ratio (E/P). The DOX is extracted from the principal component of the AA and II surveys. Starting in 1997:06, every month m we estimate each model using all observations available between the date of the first DOX estimate (1992:06) and m. We use the parameter estimates available in month m and the value of the right-hand variables in month *(m+12)*, to predict the l-month cumulative return over period *(m+12,m+24)*. We then collect the ex post forecast errors for each model. Panel A reports Clark and West (2007) one-sided test statistic for the null hypothesis of equal forecasting accuracy of M_i and M_j against the alternative of forecasting improvement of model M_j. Details can be found in Section 4.5. Standard errors are robust to heteroscedasticity and serial correlation with 12 lags. Panel B presents Campbell and Thompson (2008) measure of out-of-sample R^2. The larger the statistic, the larger the forecasting accuracy improvement of model M_j versus model M_i. Panel C measures the economic gains reaped by an expected-utility maximizer (details are in Section 4.5). Economic gains are in the form of average annualized ex post realized utility gain of switching from model M_i to model M_j ($M_i \, vs M_j$), as well as the annualized average portfolio returns, Sharpe ratio, and end-of-period wealth for an initial endowment of \$100.

To accomplish this, we follow Campbell and Thompson (2008) and solve for the optimal portfolio allocation of a rational investor who lives for only one period. Her portfolio includes two assets: a risk-free and a risky asset that proxies for the aggregate market. We assume power utility with a coefficient of relative risk aversion of 3, which is consistent with the estimation in Friend and Blume (1975). At the end of month m, the portfolio share of the risky asset is given by:

$$w^r_{m,M_j} = \frac{1}{3}\frac{(r^{\hat{M_j}}_{m,m+12})}{\sigma^{2\,M_j}_{\epsilon}} = \frac{(r^{\hat{M_j}}_{m,m+12})}{3\sigma^2_r(1 - R^{2\;\;M_j}_{m_0,m})}, \tag{11}$$

where $r^{\hat{M_j}}_{m,m+12}$ is the model-dependent forecasted excess-return for the period *(m,m+12)*, $\sigma^{2\,M_j}_{\epsilon}$ is the conditional excess-return variance that characterizes model M_j, and $R^{2\;\;M_j}_{m_0,m}$ is model M_j in-sample goodness of fit until month m.

Intuitively, when a rational investor expects a higher reward per unit of risk, she will invest more in the risky asset. Similarly, when a model reduces conditional uncertainty $\sigma_{\epsilon}^{2^{M_j}}$, the risk perceived by the investor decreases, and she invests in a risky asset more aggressively. To make our scenario more realistic, we impose two boundary constraints on w^r_{m,M_j}. We conform to Campbell and Thompson (2008) and set the lower boundary to 0, that is, no short-selling, whereas the upper boundary of 1.5 sets a limit to the investors' leverage.[21]

For each model M_j we calculate investor's average ex post annualized utility:

$$U^{\bar{M}_j} = \left(r_{p,12}^{\bar{M}_j} - \frac{3\sigma^2_{p,12}}{2} \right), \tag{12}$$

where $r_{p,12}^{\bar{M}_j}$ and $\sigma^2_{p,12}$ represent the investor's ex post average portfolio return and portfolio return variance, respectively. Panel C of Table 7 presents the comparative results for the three models. In line with the first two pieces of OOS evidence, the ex post average utility improvement calculation reveals that forecasting with the dividend-price ratio does not lead to a large ex post utility improvement compared to the use of the naive forecast, and in some instances the agent is actually slightly worse off. The conditional model always beats both the naive and the univariate model, with an annualized ex post utility improvement between 3% and 10%.

To make the achieved economic benefits easier to interpret, we report the agent's portfolio mean return and Sharpe ratio under each model. The evidence is overwhelmingly in favor of the conditional model M_2. For example, when moving from the univariate prediction model to the conditional one, the agent usually sees an annualized mean return improvement between 3% to 5%. Moreover, while the conditional model always subsumes the naive forecast, the univariate model does so only when B/M and E/P are adopted, and the improvement is between 1% and 4%. A further comparison with respect to the portfolio Sharpe ratio confirms the superiority of the conditional model over the univariate and naive ones. For instance, using the D/P as a univariate predictor, the Sharpe ratio moves from 0.13 when the naive forecast is adopted, to 0.16 with the univariate model, and jumps to 0.43 when the conditional model is used.

As a final measure of the wealth effects of moving from simple to conditional forecasting models, panel C of Table 7 also presents the results of a terminal wealth calculation: the expected utility maximizer modeled above is endowed with \$100 in the month 1997:06. The investor rebalances her portfolio every 12 months, and on each such occasion, she uses either M_0, M_1, or M_2 to choose optimal portfolio weights. Terminal wealth is calculated as of the final out-of-sample forecast. Once again, the results point to the benefit of adopting the

[21] Allowing short-selling does not change our comparative results.

conditional model M_2 over the unconditional M_1 and the naive forecast M_0. Across predictors and horizons, the gain in investor's final wealth due to the conditional model ranges from 7% to more than 100%. Results are particularly strong with the B/M ratio. In unreported tests, we further compare the terminal wealth amounts with those of a buy-and-hold strategy. We find that a buy-and-hold investor would have been ex post better off than a forecaster who adopted the naive or univariate prediction model. Nevertheless, a buy-and-hold strategy would have been subsumed by the conditional model, which earned 42% more over the period. In summary, out-of-sample results support our in-sample inference, and reject the null hypothesis of the no conditioning role of investors' DOX for stock return predictability by price-scaled predictors.

4.6 Long-horizon predictability

We extend our analysis of conditional predictability by assessing how the role of DOX changes with the predictability horizon. In Figure 6 we separate the sample period into high (above median) versus low (below median) DOX states. DOX is extracted from the extended PC time series, PC_{ext}. A longer DOX time series allows us to conduct tests over horizons that span multiple years.[22] The left side of panel A, which depicts the return predictability coefficient in low (red) and high (blue) DOX states, along with the 95% OLS confidence intervals, documents the diminishing conditioning role played by DOX as the horizon increases. The predictability coefficient in low DOX states is considerably lower than in high DOX states for horizons of 4 years of less, but by year 8, there is no difference between the two. On the right side of panel A, the difference between the R^2 obtained in low and high DOX also declines with the horizon. So, on long-horizon, a low (high) D/P ratio signals lower (higher) future returns even when the current DOX is low. This short-horizon role of DOX is consistent with the theory of extrapolative beliefs of BGJS (Section 4.1 and Table 3). Panel B documents similar findings for the D/P ratio: while persistence is higher in low versus high DOX states for short horizons, by year 8 there is an equally small predictive link between the current and future D/P. As for the case of return predictability, the short-horizon role of DOX in the mean-reversion of the D/P ratio is in line with that in BGJS (Section 4.2; Table 4).

4.7 Parameter instability in the predictability relation

Prior studies of stock return predictability have claimed that the association between price-scaled predictors and future returns is time-varying. In Figure 7, we provide evidence of such instability by presenting the recursively estimated univariate predictability coefficient of the D/P ratio on future one-year excess returns. In the period 1987:12–2013:12, we obtain the recursively estimated coefficient of predictability β_{RW_m} by fitting every month, m, the following

22 The results obtained using the short-sample PC DOX are similar.

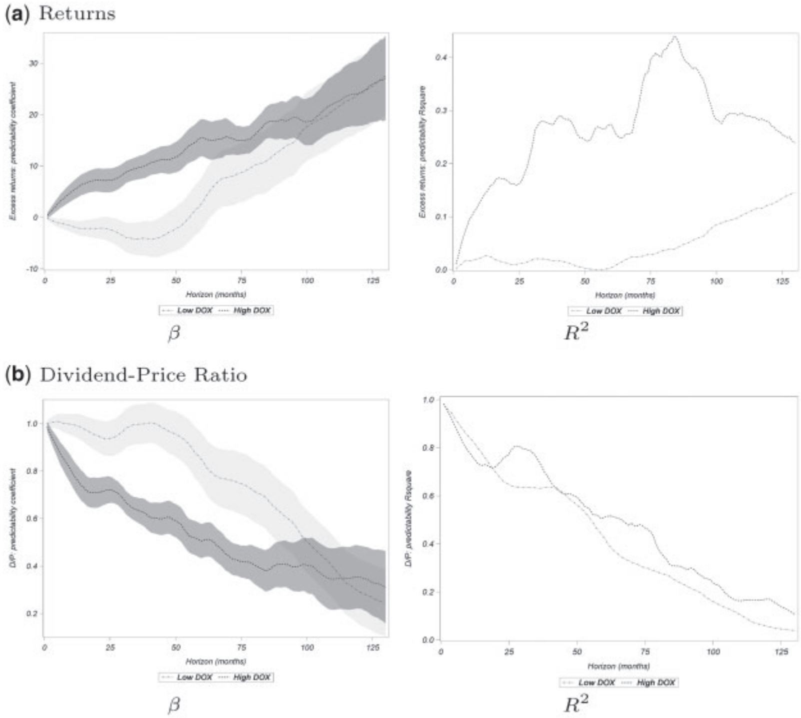

Figure 6
Conditional predictability: increasing horizon
The independent variable is the lagged D/P ratio. The dependent variable is either (1) the cumulative return on stocks in excess of the risk free rate (panel A) or (2) the D/P ratio (panel B). The predictability coefficient β and the regression R^2 are reported both for high DOX (dashed) and low DOX (dashed dotted). DOX is high if it is above its median at formation. It is low otherwise. OLS 95% confidence intervals are also reported.

predictive regression of year-ahead excess returns over a rolling window of 20 years:

$$R_{t,t+12} = a_0 + \beta_{RW\,m} D/P_t + \epsilon^{R}_{t,t+12},$$

where $t \in [m - 20 \times 12 + 1, m]$, R stands for aggregate excess return on the value-weighted portfolio of CRSP US equities, and D/P is the dividend-price ratio. Figure 7 also reports a 20-year moving average DOX ($\overline{DOX_{RW\,m}}$) extracted from the extended principal component, PC_{ext}. Under the null of no conditioning role of extrapolators' DOX for the time-series predictability of the equity premium, there should be no apparent relationship between $\beta_{RW\,m}$ and $\overline{DOX_{RW\,m}}$. In other words, a period of higher average DOX should not be characterized by better predictability compared to a period of low average DOX. Under the alternative hypothesis, the predictive ability of a price-scaled variable should be positively correlated with the in-sample prevailing DOX. Indeed, Figure 7 supports the alternative hypothesis. The correlation coefficient of the

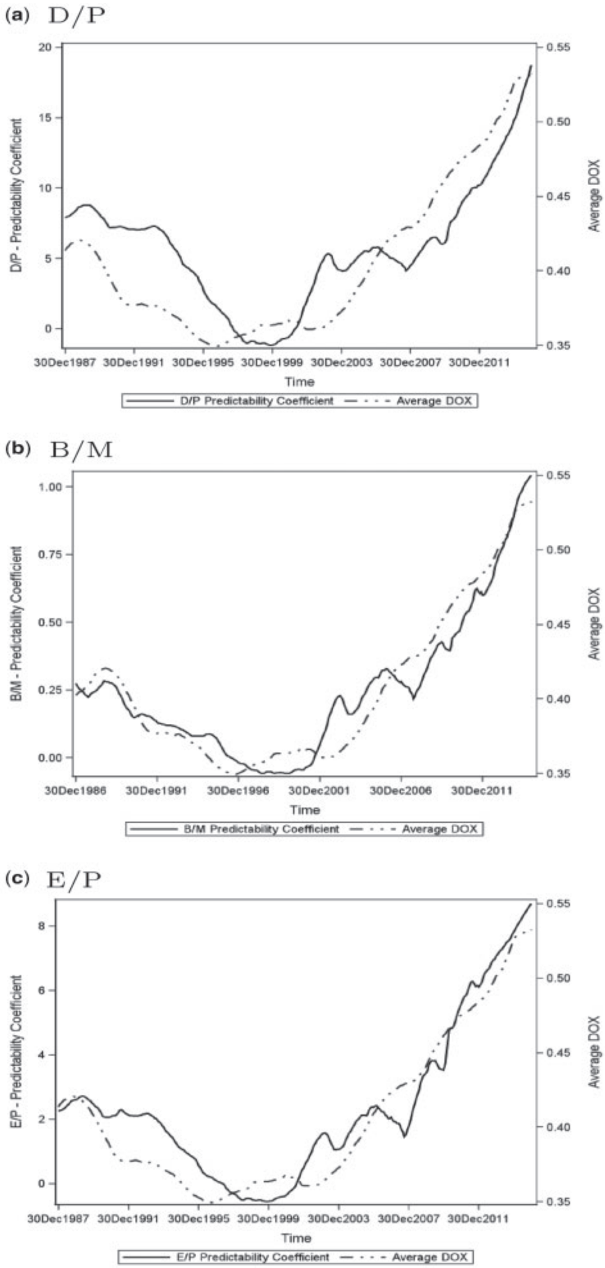

Figure 7
Stock return predictability: parameter instability
In the sample period 1967:12–2013:12, we run 20-year rolling-window univariate predictive regressions of one-year-ahead excess returns on the current D/P ratio (panel A), B/M ratio (panel B), and cyclically adjusted E/P ratio (panel C). The solid line is the recursively estimated predictability coefficient. The dashed line is a 20-year DOX moving average.

average DOX series and the recursively estimated predictability coefficient series is approximately 80% and is statistically significant at the 1% level.

A more formal test of the relation between the coefficient of predictability, $\widehat{\beta_{RWm}}$, and the average DOX, $\overline{DOX_{RWm}}$, regresses the former on the latter, with the following result in the case of predictability by the D/P ratio:

$$\widehat{\beta_{RWm}} = \underset{[-6.87]}{-21.97} \quad \underset{[9.12],}{+68.37} \quad \overline{\mathrm{DOX}_{RWm}} + \epsilon_m \qquad R^2 = 69\%$$

where Newey-West (1987) t-statistics with 12 lags are reported in brackets.[23] The result above suggests that when the average in-sample DOX grows by 10%, the coefficient of predictability increases by seven units. Additionally, an average DOX of 0.3 or above is necessary to turn the coefficient of predictability positive. This threshold is broadly consistent with the content of Table 5, in which it can be seen that the conditional coefficient of predictability, $b_0 + b_1 DOX$, becomes positive for a DOX between 0.3 and 0.45, depending on the combination of price-scaled variable and survey being selected. Additionally, time-series variation in average DOX accounts for approximately 69% of time-series variation in the predictability coefficient. Similar results apply to the use of B/M and E/P, as panels B and C of Figure 7 confirm. Consequently, it appears that the documented time-series variation in the ability of price-scaled variables to predict future returns can be explained with time-series variation in DOX.

4.8 Volatility and skewness

Having established that DOX has a conditioning role for the mean return of the stock market, we now test whether DOX is also informative about higher moments of the conditional distribution of aggregate stock returns. Specifically, we construct yearly measures of stock market realized volatility and skewness. Both are obtained from daily aggregate stock returns. Realized volatility is calculated like in French, Schwert, and Stambaugh (1987) as a sum of squared returns in the period of interest, plus an additional term that accounts for one-lag serial correlation. Skewness is constructed as the time-series mean of cubic daily returns. We then estimate conditional predictability regressions of volatility and skewness in the same spirit of our analysis in Section 4.3, as well as univariate models which serve as nested baseline specifications. We present the results in Table 8. In Column 1, we document that higher DOX is associated with higher future volatility. The link between DOX and skewness is instead only marginally significant. Column 2 shows that there is a weak unconditional association between the D/P ratio and future volatility and skewness.

[23] We also performed an ARIMA analysis of the regression residuals and performed inference with a longer lag size of 24 months. The inference is unchanged.

Table 8
Aggregate volatility and skewness predictability

Panel A. Volatility

Horizon (months)	Coefficient		D/P		B/M		E/P	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
12	a_0	0.014	0.055	0.075	0.068	0.043	0.066	0.077
		[1.844]	[3.486]	[3.144]	[7.624]	[2.38]	[4.592]	[3.52]
	a_1	0.036		−0.058		0.046		−0.029
		[1.927]		[−0.972]		[1.374]		[−0.602]
	b_0		−1.154	−2.738	−0.132	−0.099	−0.824	−1.442
			[−1.257]	[−2.629]	[−4.077]	[−1.403]	[−2.14]	[−2.521]
	b_1			4.303		−0.052		1.432
				[1.238]		[−0.354]		[0.977]
Adj.R^2(%)		5.05	3.69	7.89	10.59	14.59	6.53	9.72
Panel B. Skewness								
12	a_0	−4.463	−0.986	−18.138	−0.103	−12.439	−0.239	−14.959
		[−2.236]	[−0.278]	[−2.357]	[−0.027]	[−1.655]	[−0.062]	[−1.592]
	a_1	7.208		29.05		22.741		25.649
		[1.765]		[1.819]		[1.738]		[1.423]
	b_0		8.443	622.782	−2.668	30.55	−14.284	255.4
			[0.056]	[2.019]	[−0.212]	[1.186]	[−0.172]	[1.213]
	b_1			−1009.72		−59.89		−457.076
				[−1.33]		[−1.39]		[−1.09]
Adj.R^2(%)		4.62	−0.38	7.96	−0.29	5.90	−0.34	5.39

The table reports estimation results for the following monthly conditional model:

$$Y_{t+12} = a_0 + a_1 * \text{DOX}_t + Predictor_t [b_0 + b_1 \text{DOX}_t] + \epsilon_{t+12},$$

where Y is either the realized volatility of aggregate stock market returns (panel A) or the realized skewness, scaled by a 10^7 factor (panel B). Volatility is constructed like in French, Schwert, and Stambaugh (1987) from daily returns. Stock market return skeweness is the average of the cubic daily returns on the stock market. The *predictor* variable is one of the following aggregate price-scaled variables: dividend-price (D/P) ratio (columns 2 and 3), the book-to-market (B/M) ratio (columns 4 and 5), or the cyclically adjusted earnings-to-price (E/P) (columns 6 and 7). DOX is the degree of extrapolative weighting extracted from the principal component (PC) of the II and AA surveys. The sample period is 1992:06–2013:12. Coefficients and t-statistics are based on Newey-West standard errors.

In Column 3, where we test the conditional model, the interaction term of D/P and DOX is only weakly significant. The results in Columns 4 to 7 replicate these findings when the D/P ratio is replaced by the B/M or E/P ratio. Overall, DOX appears as a univariate predictor of volatility, and evidence of DOX as conditioning variable for higher moments of the return distribution is limited.

5. Robustness

In this section, we perform additional robustness tests. For brevity, we discuss results here but report the related tables and figures in the appendix.

We aim to document the robustness of our conditional predictability results to the inclusion of other potential state variables. To this end, we consider a set of k sources of conditional stock return predictability, which include our DOX variable, and extend our earlier specification in Equation 6 as follows:

$$R_{t,t+l} = (A' Z_t) + D/P_t (B' Z_t) + \epsilon^R_{t,t+l}, \tag{13}$$

where A and B are two (k+1)×1 vectors of coefficients, whose first row corresponds to the unconditional coefficient a_0 and b_0 in Equation 6, and Z_t is a (k+1)$\times$1 vector of potential state variables in the predictive regression, which includes a constant in its top row, and stacks all the k state variables, including the DOX, in the remaining rows. The competing state variables we consider are annual returns, Baker and Wurgler's (2006) measure of investor sentiment, growth in industrial production, and the NBER recession indicator.[24] If the DOX proxies for reversal of extreme realizations of the return on the stock market, controlling for the most recent annual returns should largely attenuate its marginal effect on predictability. Controlling for the Baker and Wurgler (2006) sentiment is important, given its prominence in the literature and its documented ability to predict future stock market returns (Baker, Wurgler, and Yuan 2012; Huang et al. 2015). Including state variables that capture the business cycle also is key, given the evidence in recent literature, which ties time-varying predictability to the state of the economy and to countercyclical risk premiums (Henkel, Martin, and Nardari 2011; Dangl and Halling 2012). We initially run conditional regressions which feature only one of the competing state variables mentioned above. We then choose k=2, as we conduct a "horse race" of our DOX with any one competing source of conditional predictability. Finally, we perform a kitchen-sink regression in which all the competing state variables are included in the model. Table A1 reports the results for the 12-month prediction horizon and the use of the D/P ratio. In each of the pairwise comparisons, the coefficient on the interaction term of D/P and the DOX is significant and subsumes the competing state variable. Furthermore, in the kitchen-sink regression, the DOX beats all the other proposed state variables combined, and none of the competing interaction terms are significant. This is evidence that DOX captures the most relevant source of variation in the predictability relation. Results are similar when we use other price-scaled variables.

To further confirm the robustness of the DOX to the inclusion of other competing state variables, we replicate the evidence presented in Section 4.7, and replace the DOX with the NBER recession dummy, and with log growth in industrial production. In Figure A2, panel B, we show that there is indeed a positive relation between the occurrence of economic contractions during the sample period, and the observed stock return predictability during that same period. The extent of the relation is stronger in the early portion of the sample period, while the association between number of recessions and predictability becomes less prominent toward the end of the sample. In panel A, we show evidence of the relation between cumulative industrial production growth and predictability. A more formal test of the relationship between predictability and potential state variables is presented in Table A2. The table shows that the presence of recessions in the sample period over which stock return predictability is assessed is positively correlated with the extent of

[24] Sentiment is from Jeffrey Wurgler's Web site (http://people.stern.nyu.edu/jwurgler). Growth in industrial production and the NBER recession indicator are from the Federal Reserve Bank of Saint Louis.

predictability, as found in prior literature. This lends support to a theory that relates predictability to time-varying risk premiums. Nevertheless, the variation in the predictability coefficient explained by this state variable is less than 2% of the overall observed variation in the predictability coefficient. A slightly higher variation is matched by cumulative growth in industrial production, which exhibits the wrong sign, however. Finally, in a specification that includes the DOX as a state variable, almost all the observed time-series variation in the predictability coefficient is accounted for, and the DOX is the strongest explanatory variable in statistical terms.

We then perform a falsification test that addresses the concern voiced by Lamont (2003) and Cochrane (2011) that surveys are a noisy proxy for investors' beliefs. If this is the case, estimating the model of extrapolation in Equation (5) equates to overfitting, and the DOX time series we estimate is noise. Using Monte Carlo simulations, we generate 50,000 fictitious DOX sequences, whose values lie between 0 and 1, and whose mean and variance match the mean and the variance of the DOX extracted from the principal component (PC) of the II and AA surveys. For each artificially created DOX series, we estimate our conditioning predictive model 6 and draw an empirical distribution of the regressions coefficients. Finally, we place our actual coefficient estimates presented in Table 5 on the empirical distribution so as to have an indication of the likelihood of obtaining our results if the DOX were pure noise. Results in Figure A3, which refers to the case of stock return predictability by the D/P ratio, reveal two important findings. First, if the DOX were pure noise, the expected coefficient estimate on the interaction (b_1), as well as the coefficient estimate on the DOX (a_1), would both be centered at zero. At the same time, the coefficient estimate on the D/P ratio (b_0) would be centered around the value we estimate in our baseline regressions in Table 5. Second, if the DOX were pure noise, it would never deliver coefficient estimates as extreme as those in our sample results, since the actual coefficients estimated always lie far from the empirical distribution. We obtain similar results for B/M and E/P. These results point to the informativeness of survey data, and confirm the information content of our DOX measure.

In our main return predictability analysis, we restricted the sample period to 1992:06–2013:12. In this period, we have focused on two independently collected surveys, II and AA, which are available at monthly frequency and with no gaps. Having two independently collected surveys allows us to better measure the common component of investor expectations. The data on the II survey extends back to 1963, which allow us to test the validity of our findings in this "out-of-sample" period to alleviate concerns that our results may be valid only for the selected sample period. Using the DOX estimated from the PC_{ext} time series described above, Figure A4 in the Appendix shows the long-sample DOX along with the relative fraction of young versus old in the stock market. Table A3 extends our main conditional predictability results to the longer sample period. The results are qualitatively in line with those presented in

Table 5, but quantitatively smaller for the analysis conducted over the extended sample period. In Table A4, where we address the potential determinants of the long-sample DOX, we notice that the relationship between DOX and its proposed covariates is somewhat different from the one documented in the short sample. A comparison with the short-sample analysis shows a weaker relation between DOX and age-based proxies for market composition in the long sample. This is to be expected, since II maintains a constant cross-section of survey participants over time and thus cannot track directly changes in aggregate expectations and DOX that stem from shifts in stock market demographics.

We also test our hypothesis on the conditional role of DOX in stock return predictive regressions outside of the United States. We chose Germany due to availability of survey data over a sufficiently long period. The ZEW financial markets survey conducts monthly interviews with 300 German institutional investors in regard to their short-term expectations for the German stock market. The survey starts in 1991:11 and is conducted monthly. Figure A5 shows the time-series variation in DOX estimated for Germany, alongside the price level of the German DAX index. In our predictability regressions, we focus on a 6-month predictability horizon, since this matches the horizon of the forecasts collected from survey participants. In Table A5 we first show that, unconditionally between 1996:10 and 2013:12, the D/P ratio has no statistically significant link to future returns. Once we allow for a conditional relationship, and use the DOX extracted from the ZEW survey, we confirm our findings from the US data, showing evidence of predictability only in states of high DOX. Our results are robust to the inclusion of controls, such as a deterministic trend, the squared D/P ratio, and lagged returns. Figure A6 provides graphical evidence of conditional predictability for German equities, which parallels those presented in Figure 5 for the U.S. stock market.

In summary, this section shows that our main results (1) are robust to the inclusion of other competing state variables that proxy for countercyclical risk premiums and investor sentiment; (2) would not arise if the DOX held no information content; and (3) can be extended to a longer sample, as well as to other stock markets. Further robustness tests are discussed in the Internet Appendix. These tests are focused on addressing issues such as small-sample bias in predictive regressions and robustness of our results to different window lengths. The Internet Appendix also includes extensions, such as the application of our conditional predictability framework in the bond market.

6. Conclusion

The evidence of time-series predictability of stock returns by price-scaled variables goes back more than 30 years. While there is a considerable amount of research that attributes predictability to time-series variation in required returns set by rational agents, less has been done to evaluate the extent to which irrational beliefs result in predictable patterns in stock returns. Motivated by the recent works of Greenwood and Shleifer (2014) and Barberis et al. (2015), we investigate how one such bias, overextrapolation, can explain the relationship between price-scaled variables and future stock returns.

In this study, we use survey data to quantify a bias in investor expectations of stock market returns in the United States. Specifically, we measure the relative weight extrapolators place on recent versus distant past returns (DOX) in forming expectations. We argue ex ante that the DOX is time-varying in nature. Through a recursive estimation methodology, we confirm our intuition and document considerable time variation in the DOX. We subsequently show that the predictability of the equity premium by price-scaled variables is contingent on the DOX. In particular, on short horizon, the relation between the D/P, the B/M, or the E/P ratio and future returns is positive and significant in economic and statistical terms when the DOX is high, while the predictability becomes weaker when the DOX is low, and the sign of the predictive relation turns from positive to negative. Furthermore, consistent with the intuition from the Campbell and Shiller (1988) present-value identity, we link the evidence of conditional stock return predictability to conditional mean-reversion in the price-scaled predictors. Specifically, we find that states of high DOX and stronger stock return predictability are associated with quicker mean reversion of price-scaled variables. Conversely, these price-ratios are much more persistent in states of low DOX, and such persistence weakens their ability to predict future returns. Finally, we reconcile evidence of parameter instability in predictive regressions in which price-scaled variables are used to predict future returns. We show that by accounting for time-series variation in the DOX, we are able to explain approximately 70% of the witnessed variability in the coefficient of predictability. These results are confirmed both in and out of sample, extend to the prediction of raw returns and capital gains, and are robust to the inclusion of other potential state variables from both the behavioral literature and the literature on time-varying risk premiums.

Our findings have important implications. The results lend support to an extrapolation-based theory of financial markets, whereby price-scaled variables not only proxy for time-varying risk premiums, but also capture the degree of mispricing in the stock market. More importantly, our study addresses an age-old question, namely: when will an overvalued asset correct back to fair value? We entertain the possibility that extrapolative beliefs are responsible for aggregate stock market misvaluation, and further hypothesize that if this is the case, mispricing should be corrected more quickly when extrapolative beliefs are more transitory, that is, when DOX is high. Our conditional regressions provide results consistent with this conjecture. Furthermore, our specification delivers accurate negative equity premium predictions. This is difficult to reconcile with traditional models of risk in which expected and required returns coincide and, therefore, expected equity returns can never be lower than the risk-free rate. Overall, our evidence suggests that DOX can be interpreted as a market-timing device which complements information on the existence of mispricing contained in the aggregate D/P, B/M, and E/P ratio.

While this new evidence in favor of a behavioral explanation of return predictability speaks to the specific effect that extrapolation may have on prices and their dynamics, there is still much to examine. Finding evidence of extrapolation does not necessarily mean that the required compensation for risk remains constant during the business cycle, and the nature of the interplay between these two features has not been studied. Furthermore, extrapolation is one of many known behavioral biases which may contribute to the rise and fall of fads. Entertaining richer expectation dynamics may prove useful in understanding asymmetric phenomena such as booms and busts. Finally, this study focuses on demand-side extrapolation, and leaves the issue of how extrapolation can affect the supply-side of the equity market for future research. There is evidence, such as that presented by Baker and Wurgler (2000) and Greenwood and Shleifer (2014), that corporations in the aggregate understand mispricing and use it to their advantage. On the other hand, a recent paper by Gennaioli, Ma, and Shleifer (2016) argues that at a micro-level, decision makers within corporations are prone to making mistakes similar to those traditionally attributed to less sophisticated investors. How supply and demand-side extrapolation jointly generate equilibrium prices is a question for future research.

Appendix A. Further Robustness Evidence

Table A1
Conditional stock return predictability: competing state variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
INT	−0.21	−0.136	−0.276	−0.069	0.698	0.706	0.705	1.000
	[−2.26]	[−1.15]	[−2.85]	[−0.94]	[4.11]	[4.92]	[2.22]	[4.15]
DOX					−1.718	−1.818	−1.744	−2.394
					[−5.29]	[−6.30]	[−3.39]	[−6.05]
IPG	0.194				−0.002			0.028
	[2.68]				[−0.02]			[0.59]
REC		−0.58				−0.235		−0.002
		[−2.14]				[−0.98]		[−0.01]
Y_{RET}			0.890				−0.185	−0.127
			[2.91]				[−0.37]	[−0.26]
SENT				−0.175				0.224
				[−0.65]				[1.38]
D/P	14.19	12.12	16.34	7.85	−24.754	−24.916	−25.263	−33.971
	[3.25]	[2.2]	[3.19]	[1.82]	[−2.91]	[−3.44]	[−1.86]	[−3.44]
DOX*(D/P)					77.081	84.764	78.741	104.952
					[4.10]	[5.03]	[3.04]	[5.61]
IPG*(D/P)	−6.72				2.012			−1.224
	[−2.49]				[0.45]			[−0.57]
REC*(D/P)		17.490				0.506		−15.987
		[1.59]				[0.05]		[−0.90]
Y_{RET}*(D/P)			−31.11				14.443	−20.436
			[−2.65]				[0.73]	[−0.74]
SENT*(D/P)				3.13				−11.691
				[0.16]				[−1.07]
Adj.R^2 (%)	21.20	32.72	26.36	28.30	41.40	50.27	40.49	56.99

We extend our specification in Equation (6) to include a set of k-conditioning variables in predictive regressions of year-ahead equity premium:

$$R^e{}_{t,t+12} = (A'Z_t) + D/P_t(B'Z_t) + \epsilon^R_{t,t+12}.$$

A and B are two (k+1)×1 column vectors of coefficients, whose first row corresponds to the unconditional coefficient a_0 and b_0 in Equation (6), and Z_t is a (k+1)×1 column vector of potential state variables in the predictive regression, which includes the constant 1 in its top row, and stacks all the k state variables, including the DOX from the PC time series, in the remaining rows. The competing state variables we consider are annual returns (Y_{RET}), the Baker and Wurgler's (BW, 2006) measure of investor sentiment (SENT) growth in industrial production (IPG), and the NBER recession dummy (REC). Specifications (1) to (4) set k=1 and replace the DOX with one of the competing state variables in the conditional predictability regressions. Specifications (5) to (7) set k=2, as they horse race the DOX measure with one competing source of conditional predictability. Specification (8) performs a kitchen-sink regression in which all the competing state variables are included in the model. The regressions that include the BW measure of sentiment are run over the period 1992:06–2010:12. All other regressions are run in the period 1992:06–2013:12. t-statistics are based on Newey-West (1987) with 12 lags.

Table A2
Parameter instability in the predictability relation: competing state variables

	(1)	(2)	(3)
INT	4.345	−6.266	−29.141
	[2.80]	[−1.56]	[−8.75]
$\overline{CUMREC_{RW}}$	0.140		0.293
	[1.48]		[3.13]
$\overline{CUMIPG_{RW}}$		0.464	0.161
		[3.02]	[1.93]
$\overline{DOX_{RW}}$			68.426
			[8.4]
Adj.R^2 (%)	1.30	15.80	78.60

In the period 1987:12–2013:12, we obtain the recursively estimated coefficient of predictability $\widehat{\beta_{RW_m}}$ by fitting every month m the following monthly predictive regression of year-ahead excess returns over a rolling window of 20 years:

$$R^e{}_{t,t+12} = a_0 + \beta_{RW_m} D/P_t + \epsilon^R_{t,t+12} \qquad t \in [m-20\times 12+1, m],$$

where R^e stands for aggregate excess return on the value-weighted portfolio of CRSP U.S. equities and D/P is the aggregate value weighted dividend-price ratio. At the same time, we estimate in each month m a 20-year moving average DOX ($\overline{DOX_{RW}}$) extracted from the extended principal component PC_{ext}, a 20-year moving sum of NBER recessionary months ($\overline{CUMREC_{RW}}$), and a 20-year cumulative growth in industrial production ($\overline{CUMIPG_{RW}}$). We then run the following regressions:

$$\widehat{\beta_{RW_m}} = a + b(\overline{CUMREC_{RW_m}}) + \epsilon_m, \tag{A1}$$

$$\widehat{\beta_{RW_m}} = a + b(\overline{CUMIPG_{RW_m}}) + \epsilon_m, \tag{A2}$$

$$\widehat{\beta_{RW_m}} = a + b_1(\overline{CUMREC_{RW_m}}) + b_2(\overline{CUMIPG_{RW_m}}) + b_3\overline{\mathrm{DOX}_{RW_m}} + \epsilon_m. \tag{A3}$$

t-statistics are based on Newey-West (1987) with 12 lags.

Table A3
Long sample conditional stock return predictability with extrapolation

Horizon (month)		D/P			B/M			E/P		
		Baseline	(1)	(2)	Baseline	(1)	(2)	Baseline	(1)	(2)
3	a_0	0.00	0.059	−0.117	0.013	0.06	0.093	0.001	0.045	0.252
		[−0.032]	[2.358]	[−0.305]	[1.127]	[3.305]	[0.278]	[0.089]	[2.212]	[0.982]
	a_1		−0.107	−0.129		−0.092	−0.096		−0.083	−0.084
			[−2.449]	[−2.49]		[−2.733]	[−2.221]		[−2.209]	[−1.85]
	b_0	0.514	−0.993	−0.661	0.005	−0.06	−0.431	0.227	−0.274	−6.95
		[1.195]	[−1.43]	[−0.045]	[0.248]	[−2.055]	[−0.604]	[1.237]	[−1.021]	[−1.331]
	b_1		2.613	2.877		0.123	0.113		0.917	0.789
			[1.965]	[2.003]		[2.081]	[1.718]		[1.632]	[1.297]
	$Adj.R^2$ (%)	0.653	2.50	5.30	−0.12	2.25	4.01	0.71	2.07	5.06
	Controls	N	N	Y	N	N	Y	N	N	Y
12	a_0	0.001	0.308	−0.951	0.047	0.293	0.143	0.012	0.234	0.589
		[0.022]	[3.51]	[−0.894]	[0.987]	[4.637]	[0.165]	[0.226]	[2.919]	[0.768]
	a_1		−0.549	−0.634		−0.476	−0.491		−0.416	−0.423
			[−3.32]	[−4.208]		[−3.682]	[−4.111]		[−2.787]	[−2.864]
	b_0	1.884	−5.689	14.261	0.021	−0.315	−1.389	0.745	−1.734	−19.217
		[1.122]	[−2.541]	[0.362]	[0.27]	[−3.082]	[−0.722]	[1.059]	[−1.751]	[−1.283]
	b_1		13.039	14.485		0.636	0.608		4.442	4.186
			[2.886]	[3.299]		[3.008]	[2.918]		[2.226]	[2.109]
	$Adj.R^2$ (%)	2.946	11.65	26.73	0.213	13.04	24.24	2.232	9.90	20.892
	Controls	N	N	Y	N	N	Y	N	N	Y

Using GMM we estimate the following monthly time-series regression of l-month-ahead aggregate US stock excess returns $R^e_{t_0,t_0+l}$:

$$R^e_{t_0,t_0+l}=a_0+a_1\mathrm{DOX}_{t_0}+c_0'X_{t_0}+Predictor_{t_0}[b_0+b_1\mathrm{DOX}_{t_0}+c_1'X_{t_0}]+\epsilon^R_{t_0,t_0+l},$$

where *Predictor* is one of the following price-scaled variables: D/P, B/M, or E/P. DOX is the degree of extrapolative weighting extracted from the extended principal component PC_{ext} time series. Results are reported with and without controls, X_{t_0}. Such controls include a deterministic trend and the squared predictor variable ratio. Results are also reported for the traditional univariate predictability regression, which serves as a baseline. The sample period is 1967:12–2013:12. t-statistics are based on Newey-West standard errors with a lag length of l months.

Table A4
Potential determinants of the DOX: long sample

Coefficient	(1)	(2)	(3)	(4)	(5)	(6)
INT	0	0	−0.001	−0.021	−0.001	−0.01
	[0.010]	[−0.03]	[−0.004]	[−0.87]	[−0.01]	[−0.39]
REL_PART	0.119					0.13
	[0.36]					[0.78]
PART_YOUNG		0.056				
		[2.04]				
PART_OLD		−0.056				
		[−2.34]				
CUM5YR_RET_LAG			0.015			−0.122
			[0.28]			[−2.96]
REC				0.144		0.17
				[2.32]		[2.64]
TIME_FROM_REC				−0.001		0
				[−0.85]		[0.03]
TIME_FROM_CRASH				0.008		0
				[2.51]		[1.54]
MONTHLY_VOL					−0.17	2.06
					[−0.06]	[0.83]
QVOL					5.83	5.73
					[2.68]	[3.17]
EXTRQRET					−0.18	−0.1
					[−1.83]	[−2.43]
SENT						0.049
						[2.01]
$Adj.R^2$ (%)	0.08	12.88	0.01	16.96	4.75	26.41

Regressions of DOX on a set of potential explanatory variables. PART_YOUNG is the stock market participation rate of families with a head of household 50 years of age or below. PART_OLD is the stock market participation rate of families with a head of household above 50 years of age. The relative participation of young to old REL_PART is the ratio of number of the young to old families with direct holdings of stocks in the US. CUM5YR_RET_LAG is the lagged 5-year stock market cumulative return. REC is an economic recession dummy. TIME_FROM_REC and TIME_FROM_CRASH measure the time elapsed since the end of the last economic recession and a stock market crash, respectively. A stock market crash is a quarterly return realization which is two standard deviations below its unconditional mean. MONTHLY_VOL measures the 5-year volatility of monthly return, while quarterly return volatility QVOL is measured from daily returns within a quarter like in French, Schwert, and Stambaugh (1987). Extreme quarter return EXTRQRET is a dummy variable that tracks quarterly returns that are more than two standard deviations away from the unconditional mean. SENT is the Baker and Wurgler (2006) investor sentiment. Return-based variables are from CRSP. Participation-related variables are constructed jointly from the Survey of Consumer Finances and Census data on household demographics. Data on economic recessions is from the National Bureau of Economic Research, and sentiment is from Jeffrey Wurgler's Web site. All variables are detrended. Results are based on the DOX extracted from the extended principal component PC_{EXT}, in the sample period 1967:06–2013:12. t-statistics are based on Newey-West (1987) standard-errors, with 12 lags.

Table A5
Conditional stock market predictability in Germany

Coefficient	(1)	(2)	(3)	(4)	(5)	(6)
a_0	0.007	0.052	3.520	0.483	0.670	6.880
	[0.110]	[2.370]	[2.090]	[2.070]	[1.690]	[3.320]
a_1		−1.135	−1.790	−1.060	−1.275	−1.852
		[−2.350]	[−2.60]	[−2.20]	[−2.10]	[−2.47]
b_0	1.73	−17.070	−97.760	−15.140	−27.060	−181.489
	[0.67]	[−2.10]	[−1.380]	[−1.680]	[−1.100]	[−2.30]
b_1		43.520	69.010	39.670	48.950	69.380
		[2.25]	[2.52]	[2.03]	[1.97]	[2.31]
$Adj.R^2$ (%)	0.65	4.34	11.28	3.89	4.15	12.62
Controls			t	Ret	D/P^2	All

Using the survey of 6-month-ahead German stock market expectations collected by the Center for European Economic Research (ZEW), we estimate a DOX time series (DOX^G). We then perform a test of conditional stock predictability of the subsequent 6-month return on the MSCI Germany Index ($R^G_{t,t+6}$):

$$R^G{}_{t,t+6} = a_0 + a_1 \mathrm{DOX}^G{}_t + c_0' X_t + D/P_t [b_0 + b_1 DOX^G{}_t + c_1' X_t] + \epsilon^R_{t,t+6},$$

where the main predictor is the aggregate dividend-price ratio for the MSCI German Index. The sample period is 1996:10–2014:12. The control vector X contains the following alternative conditioning state variables: a deterministic trend (t), the squared D/P ratio (D/P^2), and the past 6-month return on the stock market *(Ret)*. t-statistics (in brackets) are robust to heteroscedasticity and serial correlation.

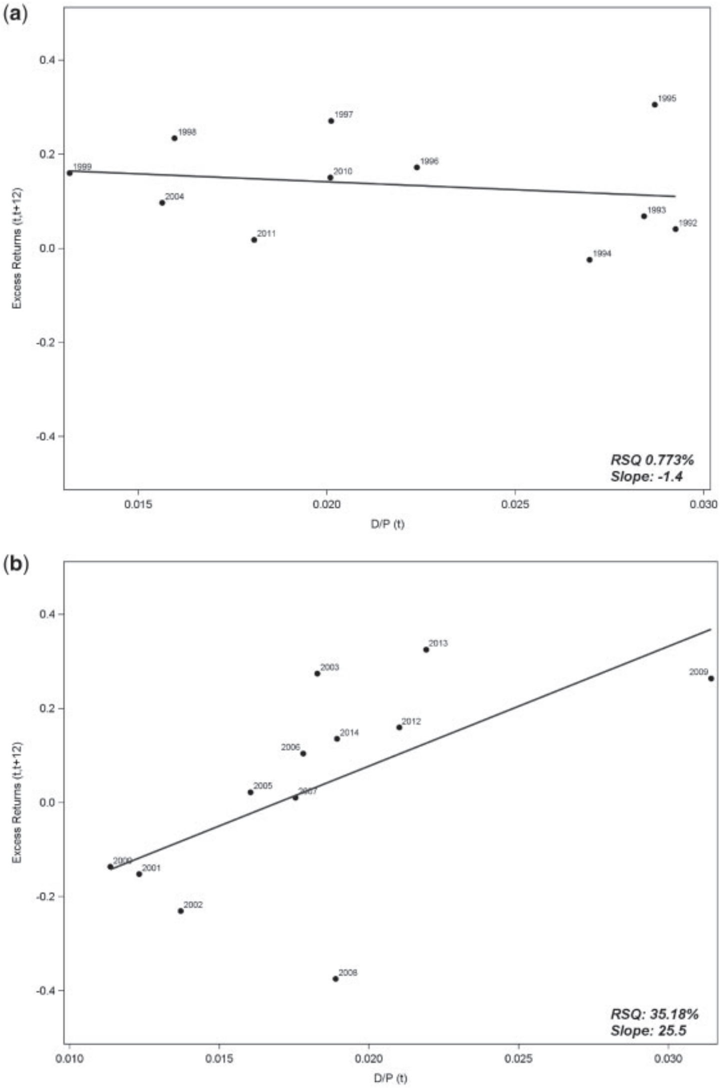

Figure A1
Conditional year-end stock return predictability
We classify each year-end *Y* based on the degree of extrapolative weighting (DOX) extracted from the principal component (PC) of the II and AA surveys in at the end of year *Y-1*. Panel A reports a scatterplot of the relationship between the D/P ratio in year *Y-1* and the future stock market excess return in year *Y*, when DOX_{Y-1} is low (i.e., below its median value of 0.51). Panel B reproduces a similar scatterplot for years with a high DOX_{Y-1} (i.e., above the median). Observations are labeled with the corresponding year *Y*. Each figure also reports the R^2 and coefficient of predictability for the corresponding univariate predictive regression.

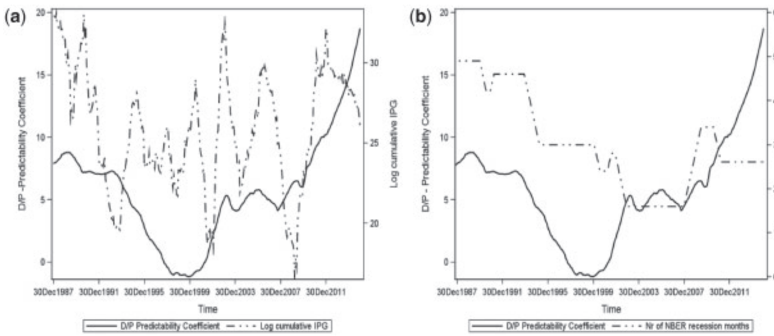

Figure A2
Parameter instability in the predictability relation: competing state variables
In the sample period 1987:12–2013:12, we run monthly univariate predictive regressions of one-year-ahead excess returns on current D/P ratio over a rolling window of 20 years (bold line). At the same time, in panel A we construct a 20-year log cumulative industrial production growth time series, and in Panel B we construct a 20-year moving count of the number of recessionary months from the NBER data on the business cycle (dashed line).

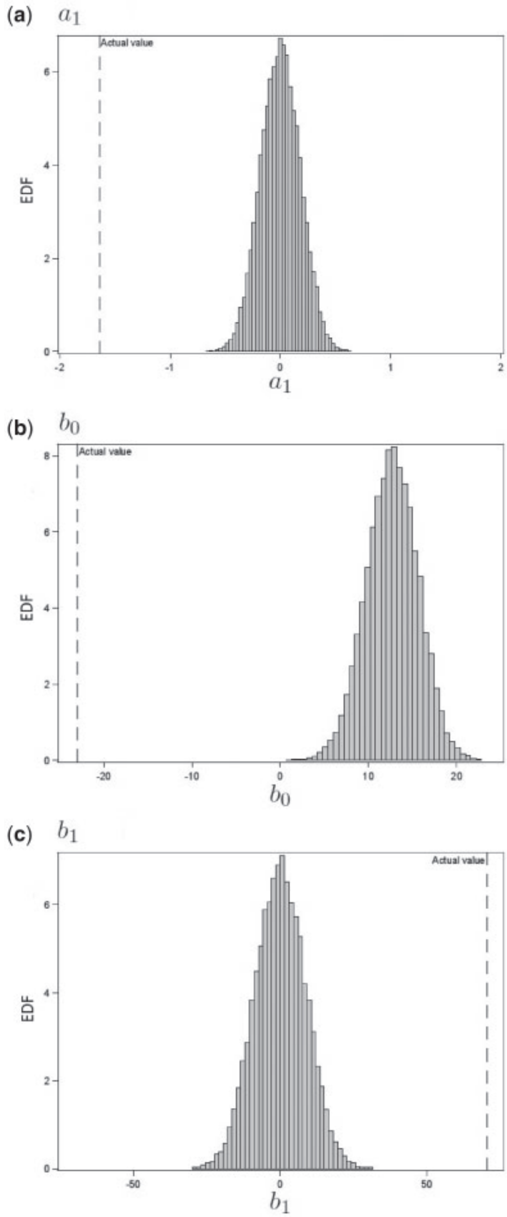

Figure A3
Conditional stock return predictability: falsification test
Using Monte Carlo simulations, we generate 50,000 fictitious DOX sequences, whose mean and variance match the mean and the variance of the DOX extracted from the principal component (PC) of the II and AA surveys. We then fit our monthly conditioning predictive model of one-year ahead excess returns

$$R^e{}_{t,t+12} = (a_0 + a_1 \mathrm{DOX}_t) + D/P_t(b_0 + b_1 \mathrm{DOX}_t) + \epsilon^R_{t,t+12},$$

using the artificially created DOX and draw an empirical distribution of the regressions coefficients. Finally, we place our actual coefficient estimates {a1, b0, b1} presented in Table 5 on the obtained empirical distribution.

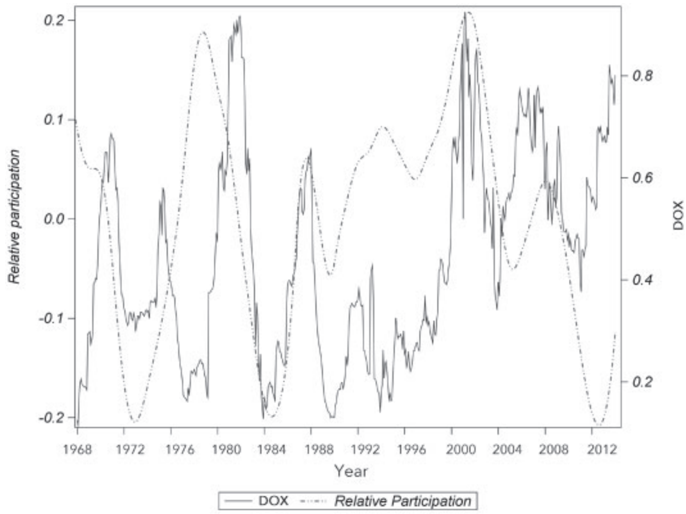

Figure A4
Long-sample DOX and stock market demographics
The figure overlays the recursively estimated DOX time series for the extended principal component (PC_{Ext}) of the II and AA surveys, with the ratio of young to old families in the United States with direct holdings of stocks. The sample period is 1967:12–2013:12.

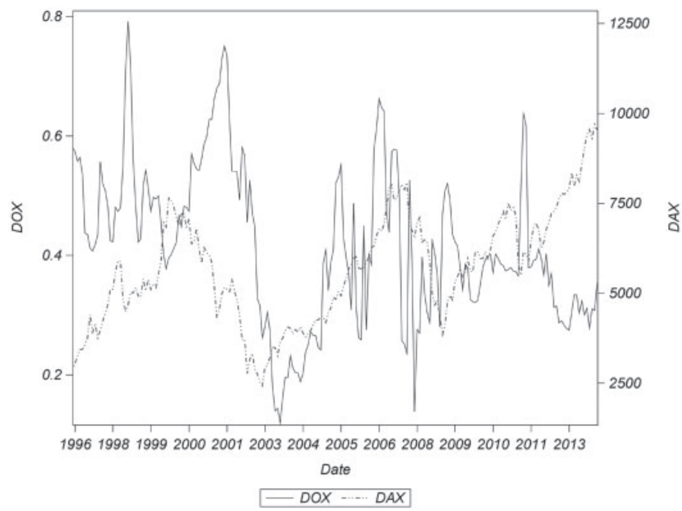

Figure A5
German investor expectations of future stock returns: time-varying DOX
This figure compares the recursively estimated DOX time series of the German investor expectations of the future 6-month stock market returns with the level of the German DAX index.

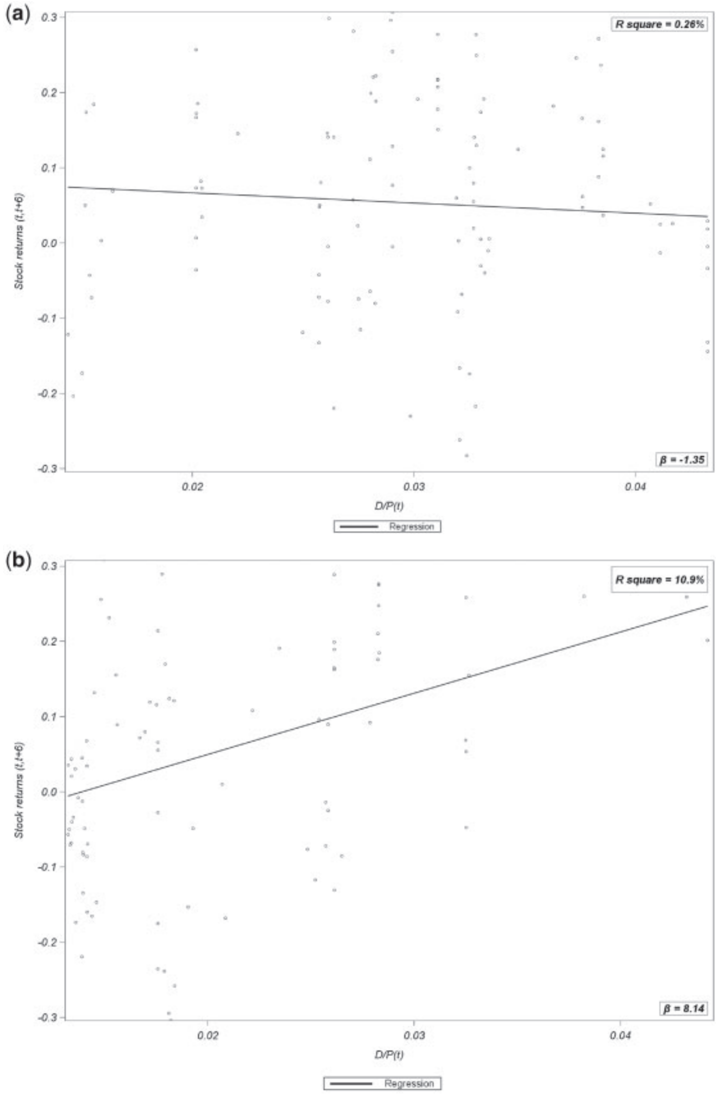

Figure A6
Conditional stock return predictability: Germany
For the period 1996:10–2014:12, we classify each month *m* based on the degree of extrapolative weighting (DOX) extracted from the ZEW survey of 6-month-ahead German stock market outlook. The top panel reports a scatter plot of the relationship between the D/P ratio in month *m* and the subsequent 6-month MSCI Germany stock return, for months characterized by a DOX that is lower than the sample median (0.43). The bottom panel reproduces a similar scatterplot for the months in which the DOX is high (i.e., above the median). Each figure also reports the R^2 and coefficient of predictability for the corresponding univariate predictive regression.

References

Abreu, D., and M. K. Brunnermeier. 2003. Bubbles and crashes. *Econometrica* 71:173–204.

Aiken, L. S., and S. G. West. 1991. Multiple regression: Testing and interpreting interactions. Thousand Oaks, CA: Sage.

Amromin, G., and S. A. Sharpe. 2013. From the horse's mouth: Economic conditions and investor expectations of risk and return. *Management Science* 60:845–66.

Andreassen, P. B., and S. J. Kraus. 1990. Judgmental extrapolation and the salience of change. *Journal of Forecasting* 9:347–72.

Bacchetta, P., E. Mertens, and E. Van Wincoop. 2009. Predictability in financial markets: What do survey expectations tell us? *Journal of International Money and Finance* 28:406–26.

Baker, M., and J. C. Stein. 2004. Market liquidity as a sentiment indicator. *Journal of Financial Markets* 73: 271–99.

Baker, M., R. Taliaferro and J. Wurgler. 2006. Predicting Returns with Managerial Decision Variables: Is There a SmallâŁłSample Bias? *Journal of Finance*, 61:1711–30.

Baker, M., and J. Wurgler. 2000. The equity share in new issues and aggregate stock returns. *Journal of Finance* 55:2219–57.

Baker, M., and J. Wurgler. 2006. Investor sentiment and the cross−section of stock returns. *Journal of Finance* 61:1645–80.

Baker, M., Wurgler, J. and Yuan, Y., 2012. Global, local, and contagious investor sentiment. *Journal of Financial Economics* 104:272–87.

Bansal, R., and A. Yaron. 2004. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance* 59:1481–509.

Bansal, R., D. Kiku, and A. Yaron. 2012. An Empirical evaluation of the long-run risks model for asset prices. *Critical Finance Review* 11:183-221.

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer. 2015. X-CAPM: An extrapolative capital asset pricing model. *Journal of Financial Economics* 115:1–24.

Barberis, N., A. Shleifer, and R. Vishny. 1998. A model of investor sentiment. *Journal of Financial Economics* 49:307–43.

Bordalo, P., Gennaioli, N. and Shleifer, A., 2017. Memory, attention, and choice. Working Paper, National Bureau of Economic Research.

Campbell, J. Y., and J. Cochrane. 1999. By force of habit: A Consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy* 107:205–51.

Campbell, J. Y., and R. J. Shiller. 1988. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1:195–228.

———. 1988. Stock prices, earnings, and expected dividends. *Journal of Finance* 43:661–76.

Campbell, J.Y. and Shiller, R.J., 2001. Valuation ratios and the long-run stock market outlook: an update (No. w8221). *National bureau of economic research*.

Campbell, J. Y., and S. B. Thompson. 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies* 21:1509–31.

Capistran, C., and A. Timmermann. 2009. Forecast combination with entry and exit of experts. *Journal of Business and Economic Statistics* 27:429–40.

Choi, J. J., and T. Mertens. 2013. Extrapolative expectations and the equity premium. Working Paper, Yale University.

Clark, T. E., and K. D. West. 2007. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics* 138:291–311.

Clarke, R. G., and M. Statman. 1998. Bullish or bearish? *Financial Analysts Journal* 54:63–72.

Cochrane, J. H. 1992. Explaining the variance of price–dividend ratios. *Review of Financial Studies* 5:243–80.

———. 2005. *Asset pricing, Revised Edition*. Princeton, NJ: Princeton University Press.

———. 2008. The dog that did not bark: A defense of return predictability. *Review of Financial Studies* 21: 1533–75.

———. 2011. Presidential address: Discount rates. *Journal of Finance* 66:1047–108.

Da, Z., U. G. Gurun, and M. Warachka. 2014. Frog in the pan: Continuous information and momentum. *Review of Financial Studies* 27:2171–218.

De Bondt, W. P. 1993. Betting on trends: Intuitive forecasts of financial risk and return. *International Journal of Forecasting* 9:355–71.

De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann. 1990. Noise trader risk in financial markets. *Journal of Political Economy* 98:703–38.

Dangl, T., and M. Halling. 2012. Predictive regressions with time–varying coefficients. *Journal of Financial Economics* 106:157–81.

Detemple, J., and S. Murthy. 1994. Intertemporal asset pricing with heterogeneous beliefs. *Journal of Economic Theory* 62:294–320.

Fama, E. F., and K. R. French. 1988. Dividend yields and expected stock returns. *Journal of Financial Economics* 22:3–25.

French, K. R., G. W. Schwert, and R. F. Stambaugh. 1987. Expected stock returns and volatility. *Journal of Financial Economics* 19:3–29.

Friend, I., and M. E. Blume. 1975. The demand for risky assets. *American Economic Review* 65:900–922.

Fuster, A., B. Hebert, and D. Laibson. 2012. Natural expectations, macroeconomic dynamics, and asset pricing. *NBER Macroeconomics Annual* 26:1–48.

Gabaix, X. 2008. Variable rare disasters: A tractable theory of ten puzzles in macro–finance. *American Economic Review* 98:64–67.

Gennaioli, N., Ma, Y., and A. Shleifer. 2016. Expectations and Investment. NBER Macroeconomics Annual, 10.3386/w21260.

Glaeser, E. L., and C. G. Nathanson. 2017. An extrapolative model of house price dynamics. *Journal of Financial Economics* 126:147–70.

Goyal, A., and I. Welch. 2008. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* 21:1455–508.

Greenwood, R., and S. G. Hanson. 2013. Issuer quality and corporate bond returns. *Review of Financial Studies* 26:1483–525.

Greenwood, R., and A. Shleifer. 2014. Expectations of returns and expected returns. *Review of Financial Studies* 27:714–46.

Griffin, D., and A. Tversky. 1992. The weighing of evidence and the determinants of confidence. *Cognitive Psychology* 24:411–35.

Henkel, S. J., J. S. Martin, and F. Nardari. 2011. Time-varying short-horizon predictability. *Journal of Financial Economics* 99:560–80.

Hirshleifer, D., J. Li, and J. Yu. 2015. Asset pricing in production economies with extrapolative expectations. *Journal of Monetary Economics* 76:87–106.

Hong, H., and J. Stein. 1999. A unified theory of underreaction, momentum trading and overreaction in asset markets. *Journal of Finance* 54:2143–84.

Huang, D., Jiang, F., Tu, J. and Zhou, G., 2015. Investor sentiment aligned: A powerful predictor of stock returns. *The Review of Financial Studies* 28:791–837.

Kendall, M.G., 1954. Note on bias in the estimation of auto-correlation, *Biometrika* 41:403–4.

Koijen, R. S., M. Schmeling, and E. B. Vrugt. 2014. Survey expectations of returns and asset pricing puzzles. Working Paper, SSRN.

Kothari, S. P., and J. Shanken. 1997. Book–to–market, dividend yield, and expected market returns: A time–series analysis. *Journal of Financial Economics* 44:169–203.

———. 2003. [Perspectives on Behavioral Finance: Does Irrationality Disappear with Wealth? Evidence from Expectations and Actions]: Discussion. NBER Macroeconomics Annual, 10.1086/ma.18.3585255.

Lamont, O. 2003. Discussion of "Perspectives on Behavioral Finance: Does Irrationality Disappear with Wealth? Evidence from Expectations and Actions" by Annette Vissing-Jorgensen, NBER Macroeconomics Annual 2003, Eds. Mark Gertler and Kenneth Rogoff.

Lettau, M., and S. Van Nieuwerburgh. 2008. Reconciling the return predictability evidence. *Review of Financial Studies* 21:1607–52.

Lewellen, J. 1999. The time–series relations among expected return, risk, and book–to–market. *Journal of Financial Economics* 54:5–43.

———. 2004. Predicting returns with financial ratios. *Journal of Financial Economics* 74:209–35.

Lustig, H. N., and S. G. Van Nieuwerburgh. 2005. Housing collateral, consumption insurance, and Risk premia: An empirical perspective. *Journal of Finance* 60:1167–219.

Malmendier, U., and S. Nagel. 2011. Depression babies: Do macroeconomic experiences affect risk taking? *Quarterly Journal of Economics* 126:373–416.

———. 2015. Learning from inflation experiences. *Quarterly Journal of Economics* 131:53–87.

Nagel, S. 2012. Macroeconomic experiences and expectations. Academic Consultants Meeting at the Federal Reserve Board.

Nelson, C. R., and M. J. Kim. 1993. Predictable stock returns: The role of small sample bias. *Journal of Finance* 48:641–61.

Newey, W. K., and K. D. West. 1987. A simple, positive semi–definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:703–8.

Paye, B. S., and A. Timmermann. 2006. Instability of return prediction models. *Journal of Empirical Finance* 13:274–315.

Pesaran, M. H., and A. Timmermann. 2007. Selection of estimation window in the presence of breaks. *Journal of Econometrics* 137:134–61.

Pontiff, J., and L. D. Schall. 1998. Book–to–market ratios as predictors of market returns. *Journal of Financial Economics* 49:141–60.

Rabin, M., 2002. Inference by believers in the law of small numbers. *Quarterly Journal of Economics* 117: 775–816.

Rapach, D., J. Strauss, and G. Zhou. 2010. Out–of–sample equity premium prediction: Consistently beating the historical average. *Review of Financial Studies* 23:821–62.

Shleifer, A., and R. W. Vishny. 1997. The limits of arbitrage. *Journal of Finance* 52:35–55.

Stambaugh, R. F. 1999. Predictive regressions. *Journal of Financial Economics* 54:375–421.

Timmermann, A. G. 1993. How learning in financial markets generates excess volatility and predictability in stock prices. *Quarterly Journal of Economics* 108:1135–45.

Tversky, A., and D. Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185:1124–31.

Yuan, Y. 2015. Market–wide attention, trading, and stock returns. *Journal of Financial Economics* 116:548–64.