Explaining Returns through Valuation

Erik Johannesson*
Baruch College – City University of New York
erik.johannesson@baruch.cuny.edu

James A Ohlson
Hong Kong Polytechnic University – School of Accounting and Finance
jaohlso@polyu.edu.hk

May 21, 2019

Abstract

This paper develops an analytically coherent yet parsimonious framework to explain market returns in a novel way: valuation is connected to returns and vice versa. The framework requires two components. First, an explicit valuation model that maps information to an estimate of value. Second, the assumption that differences between firms' actual values and values-per-model—valuation gaps—follow an autoregressive stochastic process with a single attenuation parameter and no intercept. Empirical analyses evaluate the framework's robustness and validity. The attenuation parameter can be estimated simply and efficiently. Given an estimate of this sole parameter, the framework yields implied returns which can be correlated with realized returns. The framework's explanatory power compares favorably to that of traditional OLS regressions, even though the former requires only one degree of freedom. In a setting with pooled annual data, the correlation between implied and realized returns are 73% and 64%, respectively.

Index words: valuation, returns, valuation gap, autoregressive framework.

*Corresponding author. Baruch College, Department of Accounting, 55 Lexington Ave, New York, NY 10010. Telephone: (917) 858-9131.

All data used in this study is publicly available via WRDS. We thank workshop participants at Columbia University for helpful comments and suggestions. All errors are our own.

1. Introduction and overview

A large body of accounting research studies the extent to which contemporaneous accounting data relate to changes in stock prices, i.e. market returns. Aside from linearity, most analyses do not impose structural restrictions or limit the degrees of freedom. Our paper continues this stream of research, but with the qualification of not being open-ended: it uses an information-to-value model to explain the information-to-return relation.

The prevailing modus operandi of explaining market returns motivates the research topic and our strategy. For decades, work in this area has followed a simple two-step procedure: identify some "value-relevant" independent information variables and use OLS estimation techniques. Regardless of information variables, researchers have found that the goodness-of-fit, as measured by R^2, has remained stubbornly low. Although the number of independent variables often increases, material discrete improvements tend to lack. Well-known papers expressing disappointment in this lack of explanatory power include Black (1986), Roll (1988) and Lev (1989). More contemporary research has not fundamentally changed this judgement. One can ask whether the explanatory power can be improved by deviating from the prevailing modus operandi.

To answer this question, we develop an analytically cohesive framework anchored in the precept that valid a priori assumptions can reduce the degrees of freedom needed to successfully explain returns.[1, 2] We start from the observation that, in an ideal world, if prices could be perfectly explained at the endpoints of a return interval, the same would be true for returns. Absent such perfection, a central question emerges: is there a coherent way of thinking about the valuation gap, the difference between a stock's price and an estimate of its value per valuation model? In other words, since a completely accurate model of value is impossible, even if one assumes efficient markets, it is of interest to consider the difference between actual price and model value, and how this gap can be expected to change over time.

Our framework can be described in terms of its two components.[3] First, we stipulate an exogenous information-to-value model, V_t. Second, we model the valuation gap, $G_t = P_t - V_t$, where P_t denotes price: $G_{t+1} = C \times G_t + noise_{t+1}$, and C specifies the attenuation parameter, $0 < C < 1$. That is, the simplest possible autoregressive process models the dynamic—a one-parameter process

[1] The large literature on so-called returns-on-earnings regressions includes numerous variations, for example, the inclusion of variables other than earnings, short vs. long return intervals, discontinuities due to losses, etc. The richness of variations is illustrated by Ali and Zarowin (1992), Barth, Beaver, and Landsman (1992), Cheng, Hopwood, and McKeown (1992), Copeland, Dolgoff, and Moel (2004), Easton, Harris, and Ohlson (1992), Francis, Schipper, and Vincent (2003), Hayn (1995), Livnat and Zarowin (1990) and Ohlson and Penman (1992). It is worthwhile to note that none of these papers consider the RHS specifications as derivatives of some formal valuation framework. As far as we know, only Liu and Thomas (2000) have done so using RIV. However, this paper does not consider the implications of RIV's imperfections as an empirical matter. More generally, none of the papers mentioned above identify what we view as the crux of the matter, namely, a framework that captures the difference between value per model and price.

[2] It is well to note that an improved explanatory power compared to OLS is a necessary but insufficient condition for success: the analyses also make the case that the proposed framework appeals conceptually and empirically.

[3] Our modelling can be viewed as being similar to Lee et al., but it differs in several respects that goes beyond implementation differences. It differs in purpose because Lee et al. addresses implications of "incorrect" pricing: market inefficiency is a focal point. It introduces three variables, price P and valuation V, but also V^* which represents the "true" unobservable intrinsic value. $P - V^*$ is assumed to satisfy a dynamic process. The autoregressive process without an intercept assumed in our paper would be a special case. Setting $P = V^*$ in Lee et al (1999), along with an assumption about autoregressive properties of the valuation gap, would lead to similar setups. Given the focus on market inefficiencies, this angle is not explored in Lee et al (1999). More subtly, motivated by the need to ensure a process satisfying cointegration, the P, V, and V^* in Lee et al (1999) are in logarithmic form. This would not work in our setup insofar that it would greatly reduce the simplicity of concepts and implementation that we view as a desirable core attribute of our framework. In a log setting, there are no longer any reasons to believe that the various residual noise terms have zero mean, and the use of log-modelling forces the analysis to consider the variances of noise terms. Given the assumption that $\log(P) - \log(V)$ satisfies an auto-regressive process, it is a complicated matter to calculate the distribution of $P - V$ and its expected value. Further, the use of log returns and log-transformed independent variables deviates from prior research that explains returns empirically.

specifying the time-series behavior of G. Any gap, whether positive or negative, is expected to gradually disappear, so long as the valuation model is unbiased in the long run.

An estimate of C, in addition to V_t, V_{t+1}, P_t, implies the expected return, conditional on contemporaneous information: $\hat{r} = (V_{t+1} - C \times V_t)/P_t + C - 1$. These model-implied expected returns can be correlated with realized returns. We denote explanatory power in return space "dynamic goodness-of-fit". This metric is complemented by static goodness-of-fit—accuracy in valuation space. Static goodness-of-fit is measured by the median absolute value percentage error relative to the actual price, median($|((P\text{-}V)/P)_t|$), a commonly used metric. Our framework allows for empirically linking valuation models' (static) goodness-of-fit to the (dynamic) goodness-of-fit of the relation that explains returns. Consider two exogenous valuation models, *V1* and *V2*. If *V2* dominates *V1* in terms of static goodness-of-fit, one can hypothesize that *V2* also dominates *V1* in explaining returns. This is not a foregone conclusion and yields the following proposition: attempts at improving the explanation of returns are intrinsically connected to the relative accuracy of the underlying valuation models.

Empirical evaluations of our framework require explicit valuation models. We choose to focus on valuation models based on analysts' forecast of EPS because forward EPS is more informative about value than current EPS realizations, as noted by Kothari (2001) and many others. Liu, Nissim and Thomas (2002) presents strong evidence that, in terms of valuation accuracy, analysts' consensus forecasts of EPS two years ahead in time ("EPS2") are more informative than forecasts of EPS one year ahead ("EPS1"). This evidence motivates our choice of *V1* and *V2*. The former is based on EPS1 and the latter on EPS2, both scaled by their respective market multiples. Empirical analyses confirm the finding in Liu, Nissim, and Thomas (2002) and show that the dominance of *V2* over *V1* carries over to the dynamic goodness-of-fit: *V2* explains also returns better than *V1*.[4]

The estimation of the attenuation parameter C plays an important role in all empirical applications of our framework. An attractive estimation method ought to satisfy both simplicity and robustness. The assumed dynamics of the valuation gap provide a solution. Based on the equation $G_{t+1}/G_t = C + (noise_{t+1}/G_t)$, we estimate C using the median of G_{t+1}/G_t across firms (firm indices are implicit). Alternative estimators of the attenuation parameter C, including one that maximizes the dynamic goodness-of-fit, differ little in terms of resulting magnitudes. For annual returns, and for both *V1* and *V2*, the estimate of C centers on 0.85, with only modest variation across years.

We evaluate the robustness and validity of our proposed framework from several angles. Of obvious interest is to ensure that the inherent parsimony of our framework does not come at the cost of explanatory power. Accordingly, we compare the explanatory power of returns derived from our framework to that of implied returns derived from standard OLS regressions, conditional on the same information. Realized returns, r_{t+1}, are regressed on V_{t+1}/P_t and V_t/P_t, and an intercept; the two independent variables follow from the assumed valuation gap dynamics. Although this regression comprises three unrestricted parameters, as compared to only one in our framework, its goodness-of-fit never dominates. When valuation model *V2* is applied in a pooled sample, OLS yields a dynamic goodness-of-fit of 64.3%, as compared to 73.1% for our single-parameter framework (applied across years and firms). Including year fixed effects in the OLS regression leads to a trivial improvement in goodness-of-fit (64.4%). Only the additional inclusion of firm fixed effects yields a dynamic goodness-of-fit close to 73.1%. In other words, OLS achieves an

[4] Section 2 provides the specifics ensuring unbiased valuation heuristics.

explanatory power comparable to our single-parameter framework only if one allows for a very large number of free parameters.[5]

Notwithstanding these empirical results, this paper's main contribution is conceptual: it adds to the literature by explicating the link between valuation space and returns space through an analytically cohesive framework. This aspect is novel and useful as a practical matter since it gives researchers another tool to evaluate valuation heuristics empirically. The framework we contribute is truly parsimonious, requiring only an information-to-value model and that a single easy-to-interpret parameter has been estimated, and its overall validity is supported by extensive data analysis. From this perspective, that fact that traditional OLS regressions do not outperform the framework in explaining returns serves as an important illustration but is not a goal per se. The point to be made is that simple but valid a priori assumptions can successfully replace more elaborate, unrestricted modes of data analysis when explaining returns.

2. The autoregressive framework

This section develops the autoregressive framework which is applied in subsequent cross-sectional empirical tests.[6] Let

P_t = price per share.

V_t = an estimate of price per share, a function of date t information.

Define

[5] Results are even more distinct for valuation model *V1*.

[6] Firm subscripts, related to *V*, *P*, etc., are implicit. That said, it is well to note that the empirical analyses assume that the attenuation parameter *C* is the same for all firms,

$G_t \equiv P_t - V_t$ = the valuation gap.

A single dynamic equation underpins the model. It is assumed that G_t satisfies an autoregressive process without an intercept:

$$G_{t+1} = C \times G_t + u_{t+1} \qquad \text{(ARF)}$$[7]

where $E(u_{t+1}; t) = 0$. The parameter C satisfies $0 < C < 1$. In the development of the model, C is presumed a known constant.[8]

The distribution of u_{t+1} may depend on date t information; thus, ARF allows for heteroscedasticity. Empirical applications additionally impose the condition that the noise term u_{t+1} and V_{t+1} are uncorrelated. It poses no technical problem given the ARF framework.[9]

The lack of intercept in (ARF), combined with $0 < C < 1$, ensures that, over time, the sequence of G_t gravitates toward zero, in expectation. Analytically, $E(G_{t+s}; t) = C^s \times G_t$, s > 0, which goes to 0 as s goes to infinity. The setup implies that, in the long run, V_t is unbiased relative to P_t. Note further that if one additionally assumes a symmetric error-term distribution, then, over time, the medians of $(V/P)_t$ and $(G/P)_t$ approximate 1 and 0, respectively. The G variable is central in the ARF framework due to its expected average of zero; V and P hence converge over time, in expectation, but they never have to equal each other.

[7] The acronym "ARF"—Auto-Regressive Framework—is used as a proper noun throughout the paper, even though this may be tautological in some instances. "(ARF)" refers to the actual equation above.
[8] The ARF framework is of course restrictive, the cost of only having to estimate a single parameter. A standard two-parameter generalization of ARF expands it into an adaptive expectations model: $G_{t+1} = C \times w \times G_t + (1\text{-}w) \times E(G_{t;\, t-1}) + noise_{t+1}$. In turn, the latter model is a special case of the general autoregressive model $G_{t+1} = C_1 \times G_t + C_2 \times G_{t-1} + \ldots + noise_{t+1}$, that is, the set of all higher-order autoregressive models with, in principle, an unlimited number of parameters.
[9] This assumption does not contradict that u_{t+1} and G_{t+1} connect perfectly given G_t.

While ARF's parsimony motivates its use, this attribute also raises the question of whether the framework is overly restrictive or even contradicts basic economics. The key requirement of unbiasedness—in the long run, the gap between *V* and *P* is expected to be zero—is perhaps the most critical.[10] *V* must therefore satisfy this property by construction, which is unproblematic so long as the information on which *V* depends can be scaled to become unbiased, in the long run, for any one firm. A scaled *V* can hence act as an unbiased valuation anchor and thereby effectively eliminate the need for an intercept in ARF.[11] Relatedly, the parameter *C* captures that the idea that the gap between *V* and *P*, in expectation, will become smaller with the passage of time as each firm becomes more "typical".

The magnitude of *C* captures the speed of the adjustment and could in principle be analytically connected to firm-specific characteristics, such as risk. However, any attempt to do so would likely diminish ARF's usefulness due to increased complexities. In this paper, we do not provide any such analytical reasoning but instead view the magnitude of *C* as an empirical matter.

It should be noted that ARF's contraction property allows for prices which gravitate toward, as opposed to equate, some concept of "intrinsic" value. In principle ARF's valuation gap does not preclude inefficient markets. That said, there is no such assumption required, and markets may be assumed efficient. In this regard, the ARF framework differs from the modelling in Lee et al (1999), where the dynamic modeling depends on the possibility of discrepancies between *P* and intrinsic value. Efficient markets in our context simply implies that *V* represents an imperfect assessment of firm value, without any preconceptions as to the source of the valuation gap. *V* may, or may not, be based on information measured with systematic error. This poses no problem so

[10] The unbiasedness assumption eliminates the need for an intercept in the ARF framework.
[11] To illustrate, while a firm's book value cannot act as a valuation anchor (due to conservatism), if adjusted for the market multiple this should materially mitigate the biasedness problem, in the very long run.

long as V's imperfections—regardless of source—on average is zero across firms in the long run. While the sources behind the gap between V and P may differ depending on assumptions regarding market efficiency, there will still be a contraction, in expectation, as required by ARF.[12]

We refer to V as a *valuation heuristic*. The terminology indicates the absence of a priori restrictions on V (aside from the long-run unbiasedness with respect to P); V could be no more than a common-sense guesstimate of value, without connections to textbook precepts like RIV or PVED. V is exogenous and effectively chosen by the researcher. It depends on "new" date t information, and potentially also on "old" information, revealed at dates t-1, t-2, et cetera. In principle, V can depend on a range of inputs, e.g. subsets of current financial data, such as earnings, book values, dividends; forecasts of future financial outcomes; the date t beta; the risk-free rate; the bid-ask spread. A relatively sophisticated scheme would use EPS1 times a multiplier that depends on risk and a measure of growth expectations, much like the Gordon Growth model. Though such a specification poses no implementation problems per se, it goes beyond the scope of this paper. In general, one should expect a valuation heuristic with some power to include earnings and forecasted earnings, and forecasted earnings should be more important than current and past earnings.[13]

In our empirical analyses, we use capitalized forward earnings as valuation heuristics. For each year, we calculate the cross-sectional market median *P/EPS1* and *P/EPS2* ratios as capitalization factors.

$$V1_t \equiv EPS1_t \times (\text{median } P_t/EPS1_t \text{ across firms})$$

[12] On a more basic level, the valuation gap recognizes that real-world valuation is inherently complex given the underlying information and how investors interpret it.
[13] A model like $q_1 \times BVPS_t + q_2 \times EPS_t + q_3 \times EPS1_t$ falls within the admissible set, but it must be normalized so that P - V on average approximates zero.

$V2_t \equiv EPS2_t \times$ (median $P_t/EPS2_t$ across firms).

Both valuation heuristics satisfy ARF's unbiasedness requirement; the EPS-multiplier construct implies an average G_t of zero. In a cross-sectional setting, if G_t has a symmetric distribution across all firms then G_t is positive for half of the firms and negative for the other half. Hence, every firm has a fifty percent chance of having an above-average valuation gap in the very long run, since in the very long run every firm is expected to be an average firm. This logic applies to both the EPS1-based valuation heuristic *V1* and the EPS2-based valuation heuristic *V2* (in a rational world, at date t, EPS2 = E($EPS1_{t+2}$ | info date t+1)).[14]

The next issue considers goodness-of-fit. For any given firm at date t, the absolute value of the term $(G/P)_t$ measures the valuation heuristic's error. This naturally leads to the following metric pertaining to the cross-section of all firms, referred to as the *static goodness-of-fit:*[15]

$$V\text{'s static goodness-of-fit} \equiv \text{Median}(|(G/P)_t|)^{16} \qquad (1)$$

Referring back to Equation (ARF), if the firm-specific G variables are relatively small (in absolute value terms), due to a relatively accurate valuation heuristic, it follows that the noise terms are also relatively small. The theoretical boundary case $G_t = 0$ corresponds to $u_{t+1} = 0$. More generally, the variation in G_t relates directly to the variation in u_{t+1}; in this way, the static goodness-of-fit ties into the accuracy and uncertainty inherent in the ARF framework.

[14] We make the following theory-based observation: consistent with Modigliani-Miller, the *V2* valuation heuristic should be corrected by adding the term m × *DPS1*, where m is the marginal earnings rate foregone due to the dividends paid at the beginning of the upcoming year. One can take the parameter m to equal the "typical" *EPS1*/*P* yield (Appendix I develops this case). Back-of-the-napkin calculations suggest that the error due to the missing term is unlikely to exceed 5% and is more likely about 2.5% (*DPS1*/*EPS1* ratios rarely exceed 0.5).

[15] Note that static goodness-of-fit is defined as the valuation model error, and hence it ranges from 0 to infinity. 0 implies perfect static goodness-of-fit.

[16] The use of the absolute percentage valuation error is common in research concerning the relative quality/accuracy of competing valuation models. See for example Callen and Segal (2005) and Bach and Christensen (2016).

We next develop the dynamic goodness-of-fit metric. Restating (ARF), let P serve as the dependent variable:

$$P_{t+1} = V_{t+1} + C \times G_t + u_{t+1} \tag{2a}$$

By assumption, u_{t+1} does not correlate with either V_{t+1} or G_t, a desirable property since it conforms with standard assumptions related to linear models.

Expression (2a) also explains returns. Dividing by P_t yields:

$$P_{t+1}/P_t = V_{t+1}/P_t + C \times (G/P)_t + u'_{t+1} \tag{2b}$$

$$= V_{t+1}/P_t - C \times (V/P)_t + C + u'_{t+1}$$

By setting u'_{t+1} to zero, one can express the model's implied expected return in terms of the following components:

(i) "new" information, V_{t+1}/P_t;

(ii) "old" information, $(V/P)_t = 1 - (G/P)_t$, identified at the start of the period, date t;

(iii) C, the intercept.

In other words,

$$\hat{R}_{t+1} = V_{t+1}/P_t - C \times (V/P)_t + C \tag{3}$$

Equation (3) leads to the *dynamic goodness-of-fit*. We define this metric as the rank correlation between realized returns and model-implied expected returns:[17]

$$V\text{'s dynamic goodness-of-fit} \equiv \text{corr}(return_{t+1}, \hat{R}_{t+1})$$

[17] Empirical analyses report on the rank-correlations rather than the Pearson R-coefficients (or R-squares) because the latter metric is significantly more sensitive to outliers. A close look at the data shows that Pearson Rs are typically lower than the rank correlations, and, more distinct, have a greater variation across years. Robustness tests reported on in Section 7 provide information that illustrates this point.
All our analyses refrain from any kind of winsorization or trimming of data, and estimation methods use robust methods. Thus, only our applications of OLS winsorizes the data; it is essential to avoid results that make little sense.

$$= \text{corr}(return_{t+1}, (V_{t+1} - C \times V_t)/P_t + C) = \text{corr}(return_{t+1}, (V_{t+1} - C \times V_t)/P_t) \qquad (4)$$

Analogous to the case of static goodness-of-fit, the uncertainty in u'_{t+1} relates directly to the dynamic goodness-of-fit. A perfect correlation between realized and model-implied returns obtains if and only if $u'_{t+1} = 0$.

The additive constant C in Equation (4) can be deleted since correlations are unaffected by linear transformations of variables. To be sure, C remains important since it is still part of Equation (4).

Because of the irrelevance of linear transformations, the dynamic goodness-of-fit by itself does not bear on whether the realized returns, on average, are close to the model-implied returns. It could be that model-implied returns correlate perfectly with realized returns, i.e. return = $a + b \times$ (*implied return*) for some (a, b), even though realized return ≠ model-implied return. If (a, b) ≠ (0, 1) the difference can be material. Section 8 revisits this issue.

The final point in this section pertains to how C and the dynamic goodness-of-fit depend on the return interval. Note that

$$G_{t+s} = C^s \times G_t + noise_{t+s} \qquad \text{(ARF')}$$

The contraction parameter, now identified by C^s, declines as the interval, measured by s, increases. In the limit, the term goes to zero. It follows that for very long return intervals the variable V_{t+s}/P_t alone explains returns. Under reasonable conditions, one should also expect the dynamic goodness-of-fit to increase as the return interval increases; for very long intervals, the metric approaches 1.

3. Unexpected realizations of V_{t+1}/P_t, conditioned on V_t/P_t.

This section connects the ARF framework with familiar returns-earnings concepts, including reverse regressions, and it develops additional implications which guide subsequent validation tests. We derive the precise structure of reverse regressions, in which V_{t+1}/P_t defines the dependent variable and V_t/P_t and an intercept comprise the right-hand side. The load factors are shown to depend on firms' growth rates, a new parameter in this extended ARF modelling, as well as on C. We extract concrete empirical implications which highlight the role of the length of the return interval (evaluated in Sections 7 and 8). Additionally, in the context of reverse regressions, the issue of why V_t/P_t is relevant for explaining V_{t+1}/P_t (and returns) arises. This section accordingly starts out by relating ARF to classical regression theory and identifies the roles of "new" and "old" information in explaining returns.

Assuming C exceeds zero, Equation (3) raises the question of why date t ("old") information, V_t/P_t , in addition to "new" information helps explain period (t, $t+1$) returns. After all, old information does not correlate with subsequent returns given a presumption of efficient markets. Classical regression theory resolves the issue. Given corr($return_{t+1}$, V_t/P_t) = 0, the variable V_t/P_t is relevant if and only if corr(V_{t+1}/P_t, V_t/P_t) $\neq$ 0. To elaborate, consider the EPS1 based valuation heuristic, $V1$. C differs from zero assuming that $EPS1_t/P_t$ does not correlate with the next-period market return—efficient markets—*and* that $EPS1_{t+1}/P_t$ in the cross-section correlates with $EPS1_t/P_t$. The latter correlation typically exceeds 80%. In sum, $C > 0$ takes on a specific role in Equation (3) even though V_t/P_t does not correlate with the dependent variable; the variable is relevant because it corrects the main right-hand side variable, V_{t+1}/P_t, by its effect on expectations.

The ARF framework can be recast in the spirit of reverse regressions. In doing so, we show a connection to yet another traditional concept, namely that unexpected earnings explains returns. One can replace “unexpected earnings” with the more generic expression “unexpected realizations of the valuation heuristic”. We assume that markets are efficient and that $P_{t+1} = R \times P_t + v_{t+1}$, where the noise term is unpredictable. R accordingly corresponds to the expected market return minus the expected dividend yield (R is in the order of 1.06 in an annual setting). Inserting this expression into Equation (ARF) and rearranging gives:

$$V_{t+1} = P_{t+1} - G_{t+1} = (R \times P_t + v_{t+1}) – (C \times G_t + u_{t+1}) \tag{5a}$$

$$= (R - C) \times P_t + C \times V_t + residual_{t+1} \tag{5b}$$

In (5b), the $residual_{t+1}$ equals the sum of the two sources of uncertainty, $u_{t+1} + v_{t+1}$. R, like C, is assumed to be a known parameter.

The expected value of the left-hand side of (5b), given the date t information, equals

$$\mathrm{E}(V_{t+1}; t) = (R - C) \times P_t + C \times V_t \tag{6}$$

Combining (5b) and (6) yields the unexpected realizations of V_{t+1}:

$$V_{t+1} - \mathrm{E}(V_{t+1}; t) = residual_{t+1}. \tag{7}$$

Expression (7), normalized by P_t, connects with realized return. Specifically,

$$\mathrm{corr}(realized\ return_{t+1}, residual_{t+1}/P_t) = \mathrm{corr}(realized\ return_{t+1}, V_{t+1}/P_t - ((R - C) + C \times V_t/P_t)) = \mathrm{corr}(realized\ return_{t+1}, V_{t+1}/P_t - C \times V_t/P_t).$$

The last equality goes back to the observation that correlation coefficients do not change if one adds a constant to one of the variables. The ARF framework thus picks up on the idea that unpredictable “new” information—a single source of uncertainty identified by the residual in

Equation (7)— reflects all available information that can explain returns. Although Equation (7) results in only one independent variable, the underlying two stages retain the same information content as the two-variable Equation (2a).

The parameter R determines not only a stock's ex-dividend expected return (since $E(P_{t+1}/P_t = R)$), but also P's growth in the short and long run. This means that V, too, must in the long run grow at the rate of R, in expectation. The claim follows from the ARF framework, since $E(P_s - V_s; t)$ goes to zero as s goes to infinity. The near-term expected growth rate of V can differ from R, a point made obvious by equation (6).

Equation (6) shows that the sum of the two coefficients related to P_t and V_t must add up to R, regardless of C. Moreover, to bring out this aspect and the role of C, note that if one defines $w = (C/R)$, then Equation (6) can be expressed as:

$$E(V_{t+1}; t) = R \times ((1-w) \times P_t + w \times V_t) \quad (8)$$

Because the weight on V_t increases as C increases, the forecasting of next period's V attaches an increasing weight on the current V as the persistence of the valuation gap increases. Moreover, the weight on P_t provides increasingly relevant information in the forecasting of V_{t+s} when the horizon s becomes longer; it follows because the persistence of the valuation gap decreases as C^s, a function of s, decreases. In a similar vein, the scalar R will become larger for longer horizons. The empirical part of the paper revisits these intuitively appealing implications by allowing the return interval to vary.[18]

[18] From a strict theory perspective, one may consider what happens as the return interval approaches zero, in which case C's limit equals one. Now it would appear that explaining returns and explaining values are disconnected. Given $C = 1$, the variance of change in G_{t+1}, normalized by P_t determines the key metric corr(*realized return*, $V_{t+1}/P_t - C \times V_t/P_t$). But this change is stochastically unrelated to the gap G_t/P_t. Conversely, the latter variable does not connect with

A final point in this section concerns the potential role of valuation theory; its absence may suggest that ARF lacks a solid foundation. In fact, as shown in Appendix I, ARF can be identified using fundamental assumptions consistent with a basic valuation setup. Two core assumptions underpin the analysis: (i) an EPS information dynamic that ensures dividend-policy irrelevance, and (ii) the classical precept *P* = PVED. The key takeaway shows how the persistence of EPS in the dynamic (i) determines the persistence parameter *C*.

4. Estimating the parameter *C*

This section concerns the practical issue of estimating the contraction parameter *C*. We consider four distinct estimation methods, all related directly to prior analytical analyses. Given our research objective of explaining returns parsimoniously and effectively, our principal choice of *C* estimator ideally satisfies simplicity without being less effective than the three competing methods. Furthermore, the use of multiple methods becomes relevant when one assesses ARF's general validity.

Our first method, M*, serves as an indispensable benchmark because it maximizes the dynamic goodness-of-fit over *C*. The second method, referred to as M, is a truly simple non-parametric method that can be assigned a principal role so long as it does not underperform any other methods. Ideally, it should be only marginally worse than M*. This method depends solely on firms' valuation gaps, *G*. The third method, M', relates to the expression explaining returns, Equation (3), and it captures the spirit of conventional regression analysis. The fourth method, M'',

---

the former except in the peculiar singularity case, *G* = 0 for all t. The conclusion is that in case of very short return intervals, one cannot have it both ways: a choice between explaining returns or explaining values is imposed on the researcher. (Nota bene: a full, rigorous theory of these issues requires a considerable analytical machinery to work out the precise claims. This is outside the scope of this paper.)

implements the reverse regression in Equation (5b). Both M' and M'' are parametric and will be more thoroughly discussed below, along with M* and M.

*Method M**: C is estimated by maximizing the dynamic goodness-of-fit over C: $\text{Max}_C[\text{corr}(\textit{realized return}, V_{t+1}/P_t - ((R - C) + C \times V_t/P_t))]$. The resulting estimate of C is referred to as C^*. Under mild concavity conditions, C^* will be unique. Numeric methods verify the concavity and uniqueness of the solution.

As noted, M* provides an upper-bound benchmark for the dynamic goodness-of-fit. This allows for a comparison of estimates of C derived from other estimation methods, to get a sense of the extent to which these methods are inferior, if at all.

The M* estimation method by itself has three limitations. First, conceptually, it lacks parsimony since it involves a mechanical numerical maximization. Second, it is silent regarding the relevance and validity of the ARF framework. Third, M* raises the specter of statistical overfitting: the "true" dynamic goodness-of-fit is likely to be lower, since M* adapts to sample idiosyncrasies. There is thus a need for an estimator of C that does not connect logically to the dynamic goodness-of-fit metric.

Method M: C is estimated using the cross-sectional median of G_{t+1}/G_t, inherent in Equation (ARF).[19]

Method M serves as the principal method through which we estimate the C parameter, as noted above. Setting aside the issue of how it compares to M*, M's simplicity will be a questionable

[19] An obvious advantage of using the median instead of the mean is that the denominator of individual observations may be close to zero without causing problems due to outliers.

virtue if more sophisticated methods perform better in terms of dynamic goodness-of-fit. The remaining two estimation methods, M' and M", are used to address this aspect.

Method M': C is estimated via Equation (2): $P_{t+1} = \mathrm{k}_1 \times V_{t+1} + \mathrm{k}_2 \times V_t + \mathrm{k}_3 \times P_t + noise_{t+1}$. The ratio $\hat{\mathrm{k}}_2/\hat{\mathrm{k}}_1$ provides an estimate of C; it ensures that corr(realized returns, implied returns) = corr(*realized return*, V_{t+1}/P_t - $C \times V_t/P_t$).

Method M' may seem arbitrary insofar that Equation (2b), no less than Equation (2a), can be used to estimate k_1, k_2, and k_3. To circumvent this dilemma, we apply Theil-Sen estimation (TS) rather than OLS. Estimations (2a) and (2b) will yield identical results (Wilcox (2010) provides a textbook treatment of TS. Ohlson and Kim (2015) discusses the method's efficiency advantages over OLS in various valuation settings).

Method M'': C is estimated via Equation (5b): $V_{t+1} = \mathrm{k}_1 \times P_t + \mathrm{k}_2 \times V_t + noise_{t+1}$. Per Section 3, the estimate of C equals the estimate of k_2. As in the case of M', we use TS to avoid the issue of having to choose between (5a) and (5b).

Finally, setting aside the issues of estimating C and assessing the resulting dynamic goodness-of-fit metric, methods M' and M'' can be used to address the validity of the ARF framework more generally. M' and M'' lead to predictions regarding the estimated coefficients, both in terms of signs and relative magnitudes, which are of interest in their own right. Subsequent analyses evaluate these aspects.

5. Data

The data requirements are unusually straightforward as only three variables are involved: price per share, and forward EPS for one and two years ahead ("EPS1" and "EPS2", respectively).

Consensus forward EPS data derive from the I/B/E/S database. Our basic sample uses companies in the S&P500 index and spans 2003 through 2014; it ensures high-quality forecasts, a desirable attribute. The starting year was chosen to avoid the dot-com bubble and its aftermath. The relatively short sample period allows us to show results for individual years, in addition to averages. Such detail helps the reader to make up his/her own judgements about the stability of results across years. To ascertain that our results are not sample specific, we also provide supplementary summary results for firms in the S&P1500 index, going back to the year 1990.

Starting in 2003, we take the constituents of the S&P500 index per March 31 and define these as the 2003 sample. At the end of the third month for each calendar year, we revise this list of companies according to changes in the S&P500 index. This means that our yearly samples change somewhat over time, in terms of constituent companies, but that the size remains almost constant.

We obtain consensus analyst forecast of EPS1 and EPS2, along with price data, from I/B/E/S and CRSP, respectively. To mitigate concerns that different fiscal year-ends may skew the analyses in one way or another, we identify each company's fiscal year-end and then retrieve the consensus forecasts nine months prior to this. The procedure allows us to distinguish between EPS1 and EPS2 as clearly as possible. If we had alternatively used calendar time, the companies whose fiscal year-ends differ from calendar year-ends would have muddied this distinction.

While the data analysis involves shorter and longer return periods than 12 months, the starting point is always the same: nine months before fiscal year-end. Regarding the practical specification of return interval length, at the lower end it is set to 3 months. A shorter period may increase the impact of estimation errors and cause more erratic outcomes. This is familiar from returns-earnings regressions with very short windows (e.g. a few days) around the earnings announcement date.

These settings tend to yield low explanatory power and small ERCs, with large associated standard errors. The longer interval maintains symmetry; it is set to 21 months.

Table 1 provides descriptive statistics for $EPS1_t/P_t$ and $EPS2_t/P_t$ for each fiscal year. As should be expected, and consistent with prior literature, the distribution of $EPS2_t/P_t$ shifts to the right compared to the distribution of $EPS1_t/P_t$—the median is approximately 12% greater. Furthermore, EPS2 exceeds EPS1 in every single year. This outcome is expected in a growing economy. However, the relatively large magnitudes of EPS growth compared to economic growth presumably also reflects the widely known—and material—optimistic bias in analysts' EPS2 forecasts.

TABLE 1: Forward E/P ratios: descriptive statistics

Table 1 also shows the 25th percentile and the 75th percentile. The interquartile range is sufficiently large to imply that any use of (the inverse of) $EPS1_t/P_t$ and $EPS2_t/P_t$ as valuation heuristics leave out much valuation-relevant information. It guarantees non-trivial valuation gaps. Finally, Table 1 shows that the interquartile ranges for $EPS1_t/P_t$ and $EPS2_t/P_t$ are similar and steady over time. This aspect suggests that C will be reasonably similar for valuation heuristics $V1$ and $V2$. More importantly, it also suggests that year-specific estimates of C will not vary unduly across the years.

6. Empirical analyses

This section addresses the three main empirical research questions derived from our ARF framework. Broadly speaking, these pertain to the one-to-one connection between static and dynamic goodness-of-fits, the restricted scope of parsimoniousness, and, finally, our proposed framework's efficiency in explaining returns. Specifically,

i) With respect to connecting valuation to returns, we evaluate the issue of dual dominance: relative static goodness-of-fit should connect with relative dynamic goodness-of-fit. In other words, if valuation heuristic *V2* dominates *V1* in terms of static goodness-of-fit, does *V2* also dominate *V1* in terms of dynamic goodness-of-fit?

ii) With respect to parsimony, we evaluate whether the simple method M results in estimates of *C* and goodness-of-fit similar to those deriving from M*, the maximizing of dynamic goodness-of-fit over *C*. Do M and M* yield similar results?

iii) With respect to explaining returns efficiently, we evaluate how the ARF framework compares to traditional OLS regressions. Do OLS regressions, with wide latitude in using unrestricted parameters, outperform ARF in terms of dynamic goodness-of-fit?

Before addressing these three questions, a preliminary assessment of the central parameter *C* is informative. The ARF framework is unlikely to provide a useful mode of analysis to explain returns unless estimates of *C* remain stable across years, for both *V1* and *V2*. Using estimation method M, we accordingly proceed by investigating this stability aspect.

Table 2: Estimates of *C* using method M

Table 2 shows low variation in estimates of *C* across all years and for both valuation heuristics. Approximately 85% of all estimates fall within a range of 0.65 and 0.95, with the average yearly estimate being 0.830 for *V1* and 0.838 for *V2*. These results point to a distinct attenuation in the valuation gap *G*, but more importantly also to intertemporal stability. Overall, it is our assessment that method M provides robust estimates of *C*. As a numeric reference point it generally come close 0.85.[20]

Next, consider question i). We formalize it as a proposition which relates the relative performance of static valuation (price perspective) to the relative performance of models explaining returns.

> PROPOSITION 1. – *The valuation heuristic V2 dominates V1 on both the static and the dynamic goodness-of-fit metrics.*

Table 3, Panel A compares the average yearly dynamic goodness-of-fit of *V1* to that of *V2*. It shows an increase from 0.575 (*V1*) to 0.645 (*V2*). With respect to individual years, *V2 always* scores better dynamic goodness-of-fit than *V1*. There is little ambiguity about *V2*'s superior performance regarding dynamic goodness-of-fit.

TABLE 3, Panel A: Dual dominance, *V2* compared to *V1*

With respect to the complementary performance dimension, the static goodness-of-fit, prior research provides evidence that *V2* dominates *V1* (Liu, Nissim, and Thomas 2002). One would expect the same to be true in our setting, even though the time periods and the sample of firms differ. Table 3, Panel A shows that *V2*'s average error of 20.4% dominates *V1*'s 22.9%. Focusing on the 12 individual years, *V2*'s static goodness-of-fit is better in all but one (2012). This lends strong support for P1 (using a binomial test, a 1/12 outcome implies rejection of the null of randomness at the 1% level).

[20] We do not attach standard errors to the estimated *C* parameters. While this could be done using various bootstrapping methods, it is unnecessary. As a more straightforward alternative we instead consider the variation of the estimated *C* parameters across individual years in our sample and across estimation methods. We believe the consistency of our results are strong enough to convince the reader of the conclusions drawn.

Consider next a pooling of all years. Table 3, Panel B, shows results which provide additional support for P1: both the static and the dynamic goodness-of-fit for *V2* exceed those for *V1*. The dynamic goodness-of-fits are 0.731 (*V2*) and 0.662 (*V1*), and the static goodness-of-fits are 0.203 and 0.223, respectively.

TABLE 3, Panel B: Dual dominance, pooled data

This finding holds even though the estimates of *C* are effectively the same for *V1* and *V2*, as Table 3, Panel B shows (the average *C* for *V1* and *V2* are 0.865 and 0.874, respectively). The estimates of *C* do thus not explain the differential in goodness-of-fit between *V1* and *V2*.

Proposition P1 can be evaluated for any return interval, at least in principle. While the static goodness-of-fit does not vary with the return interval length, the dynamic goodness-of-fit does. Here we consider two additional cases, in which the return interval is changed from 12 to 3 months and 21 months, respectively.

TABLE 3, Panel C: Dual dominance, 3 and 21 months return intervals

As Table 3, Panel C shows, using a 3-month return interval, V2 outperforms V1 in terms of dynamic goodness-of-fit in all but two years. Using a 21-month return interval, results are as strong as they can be: the dominance holds up for all years. Taken together, the results in Table 3 clearly support the validity of Proposition 1, regarding dual dominance.

Next, consider the second question (ii), which concerns the potential tradeoff between simplicity and accuracy when estimating *C*. We evaluate the relative performance of the two estimation methods M (the simplest method) and M* (the most accurate method). Specifically, we calculate the two method-specific implied returns, which in turn are used to calculate method-specific dynamic goodness-of-fits. We also compare the estimates of *C* to get a better sense of the source of any differences in dynamic goodness-of-fit; similarities with respect to dynamic goodness-of-fit should imply similarities in the estimates of *C*, and vice versa. An optimal outcome should suggest that the two methods are effectively no different, and equivalence with respect to *C* provides a sufficient condition for this.

PROPOSITION 2 – *Dynamic goodness-of-fit and estimated C parameters resulting from estimation method M do not deviate materially from those of M*, the maximization method.*

Although Proposition 2 is evaluated using both valuation heuristics *V1* and *V2*, our commentary focuses *V2*, due to its superior explanatory power.

TABLE 4: Dynamic goodness-of-fit and the efficacy of M

Table 4 present results related to Proposition 2. With respect to *V2*, annual averages show that the dynamic goodness-of-fit for M* and M are virtually identical: 0.650 and 0.645, respectively. Comparisons for any given year show only trivial differences. With respect to the estimates of *C*, the annual averages are also similar, but not to the same extent as the dynamic goodness-of-fits: *C**, via method M*, equals 0.872 and *C*, via method M, equals 0.838. For individual years, the average difference approximates 15%. These results presumably reflect that the dynamic goodness-of-fit is relatively insensitive to changes in *C* around the optimum.

Next, consider the results for the EPS1-based valuation heuristic, *V1*. Table 4 shows that conclusions differ little from those pertaining to *V2*. Dynamic goodness-of-fit and estimated *C* parameters do not differ materially between the methods M* and M.

We conclude that the simple method M, through which C is estimated using the cross-sectional median of G_{t+1}/G_t, yields results very similar to the maximization method M*. This supports Proposition 2—parsimony has not come at the cost of explanatory power.

One can ask what loss of dynamic goodness-of-fit takes place by disregarding any contraction of the valuation gap and putting $C = 1$. The ARF framework would hence reduce to the simplest possible benchmark: changes in V, normalized by the beginning of period price, explain returns. This aligns with a core idea in the history of returns-earnings studies, namely that the change in earnings should explain returns (Ball and Brown 1968). Compared to C^*, the goodness-of-fits show declines of, on average, 0.022 (for V2, on average the difference is 0.650 – 0.629 = 0.021, and for *V1* the difference is 0.579 – 0.556 = 0.023).

As the analytics in Section 2 showed, $C = 1$ corresponds to the boundary case when the return interval approaches zero. Although the above results indicate that setting $C = 1$ can act as a crude approximation when implementing ARF, theory asserts that this is not a generally applicable approximation. As was shown in Section 2, when the return interval increases, C will decrease, implying that setting $C = 1$ will no longer be an effective approximation. These issues will be empirically evaluated in Section 7.

Next, to appreciate the effectiveness of M, it is of interest to consider the performance of the supplementary estimation methods M' and M''. These methods' dynamic goodness-of-fit and estimates of C can be compared to those of M to evaluate the intrinsic robustness of ARF. It is unlikely that these supplementary methods will outperform M, given the latter's strong performance, documented above. They may, however, perform worse. For ARF to appeal as an empirical representation of how the world works, the relatively minor additional complexity inherent in M’ and M’’ ought not to result in a significant decrease in the dynamic goodness-of-fit.

TABLE 5: M’ and M’’

Consider first the results related to *V2*, in Table 5, Panel B. Averages show that the dynamic goodness-of-fit for M’, 0.646, is virtually identical to that of M, reported on previously. The estimates of *C* using M’ average 0.841, which also differs only marginally compared to M. With respect to M”, results show a slightly worse dynamic goodness-of-fit of 0.618, and a higher estimate of *C*, 0.924. Overall, results show that the supplementary estimation methods support ARF, although one can argue that M”, in contrast to M’, falls short on the accuracy criterion. Conclusions regarding valuation heuristic *V1*, based on the results in Table 5, Panel A, differ little from those for *V2*.[21]

Our final proposition iii) raises perhaps the most stimulating issue, because it compares the explanatory power of the parsimonious ARF framework to that of the traditional way of explaining returns, namely, OLS. To make the comparison meaningful, the regressions rely on the same information, that is, either *V1* or *V2*. For each of these valuation heuristics, to explain *t+1* returns, the right-hand side of the regressions comprise both “old” date *t* information and “contemporaneous” date *t+1* information. Beyond this information aspect, the traditional regression analyses make no reference to any valuation precepts and leave estimates of the three coefficients unrestricted. Because ARF is less restrictive than OLS, one can argue that OLS ought to perform better.

[21] The methods M’ and M’’ share the feature that they depend on the estimation of an underlying equation, Equations (2) and (5), respectively. To get additional sense of the validity of the ARF framework, it is therefore of interest to consider the stability of all coefficients estimated via M’ and M’’, not only those used to estimate *C*. Table 5 shows that they are consistently of the right sign and of reasonable magnitude, meaning that the equations estimated appear to make empirical sense. This also suggests that the estimates of *C* are reasonable.

PROPOSITION 3 – *The ARF framework does not underperform traditional OLS regressions in terms of dynamic goodness-of-fit.*

OLS-implied returns are calculated by estimating Equation (2b): an intercept and V_{t+1}/P_t and V_t/P_t, act as the independent variables, and the market return, P_{t+1}/P_t, is the dependent variable. The implied returns inferred from the estimated model are thereafter correlated with actual returns to yield a dynamic goodness-of-fit score. This score can be compared to the dynamic goodness-of-fit deriving from the ARF framework (with method M estimating C).

Table 6 provides the results. As before, the discussion focuses primarily on valuation heuristic *V2*. Average yearly dynamic goodness-of-fit for ARF and OLS is 64.5% and 63.2%, respectively.

TABLE 6: ARF compared to OLS—dynamic goodness-of-fit

Looking at individual years, the dynamic goodness-of-fits are reasonably similar but ARF outperforms OLS in 9/12 years in support of Proposition 3. In statistical terms, equivalence cannot be rejected at an alpha level of 5%. Moving from annual to pooled data shows that OLS does distinctly worse: its dynamic goodness-of-fit is 64.3% compared to 73.1% for ARF.

The results for the pooled setting raise the issue of fixed effects. To address this issue, we expand the OLS analysis to include firm and/or year fixed effects, and we again compare the OLS-based dynamic goodness-of-fit to those of ARF (using M to estimate C). Note that fixed effects have no role in our ARF analyses; the calculation of dynamic goodness-of-fit still depends only on a single estimated parameter.

Table 6 shows that year fixed effects do not change the dynamic goodness-of-fit materially. It increases from 0.643 to 0.644. By contrast, when including firm fixed effects, the dynamic goodness-of-fit increases to 0.729, which is similar to the case in which both sets of fixed effects are simultaneously included. ARF nonetheless yields a marginally better dynamic goodness-of-fit, 0.731, as previously reported.

To complete the analysis, consider finally the case of *V1* and the inclusion of year fixed effects, firm fixed effects, and year and firm fixed effects, in the OLS regressions. The resulting dynamic goodness-of-fits are 0.480, 0.231, and 0.513, respectively.[22] The dynamic goodness-of-fit for ARF, 0.662, is considerably higher.

To summarize, this section evaluates whether return analyses can achieve satisfactory (or even superior) explanatory power through an analytically coherent framework that assigns an explicit role to what estimates value. We address this question empirically using the framework ARF. Through three propositions we show that i) goodness-of-fit in valuation space connects to goodness-of-fit in returns space; ii) the single parameter *C* in ARF can be estimated robustly with close to optimal resulting explanatory power. Finally, iii), to explain returns, ARF works at least as effectively as traditional OLS regressions. Collectively, these results validate the idea that formalizing the role of valuation in a coherent and parsimonious way implies an effective mode of explaining contemporaneous returns. The benefits from additional complexities appear to be absent.

7. Additional validation of the ARF framework

The empirical analyses in the previous section consider return intervals of 12 months. This section varies this return interval and evaluates related empirical propositions, which derive from the

[22] On a more subtle note, Table 6 exhibits uncertainty about the role of the fixed effects. Depending on which valuation heuristic is used, *V1* or *V2*, the relative importance of the fixed effects shifts. For *V1*, the explanatory power increases more when including year fixed effects, whereas for *V2*, the explanatory power increases more when firm fixed effects are included.

analytical Section 3. The evidence produced hence bears directly on the validity of the ARF framework.

Section 3 yields three distinct points that pick up on implications of changing the return interval length. When it increases, it follows that:

a) The attenuation parameter *C* is expected to decline;

b) The dynamic goodness-of-fit is expected to increase;

c) The implied growth parameter, *R*, is expected to increase.

To empirically evaluate a) and b), we again rely on our main method M to estimate *C*.

With respect to c), consider Equation 5b (the basis for estimation method M''):

$$V_{t+1} = \mathrm{k}_1 \times V_t + \mathrm{k}_2 \times P_t + residual_{t+1}$$

where $\mathrm{k}_1 = (R - C)$ and $\mathrm{k}_2 = C$.

Summing the estimates of k_1 and k_2 hence implies an estimate of *R*. The equation embeds the following intricacy: the sum of k_1 and k_2 is expected to increase as the return interval increases *despite the fact that k_2 is expected to decrease*, as per a) above. It is far from obvious that both a) and c) will hold.

TABLE 7: Return intervals and robustness

Table 7 provides the relevant results, and, as before, the discussion refers primarily to valuation heuristic *V2*. We first evaluate whether the average estimate of *C* decreases as the return interval increases. Consistent with point a) above, Table 7, Panel A shows that going from the shorter 3-month return interval to the baseline 12-month return interval to the longer 21-month return interval results in average estimates of *C* of 0.968, 0.838, and 0.789, respectively. Given the null of randomness, the probability of finding this monotonicity is 1/6 = 17%. Looking at individual years, the monotonicity property holds for 5 out of 12 cases. Under the null, the probability of observing 5 or more cases of the monotonicity property is less than 2%. Monotonicity deviations occur primarily when going from 3 to 12 months. Overall, per ARF's predictions, the evidence clearly points to *C* declining as the return interval increases.

Next, we evaluate whether the dynamic goodness-of-fit increases as the return interval increases.[23] As shown in Table 7, Panel A, the evidence strongly supports this: the average dynamic goodness-of-fit equals 36.5%, 64.5%, and 69.7%, respectively. For individual years, monotonicity holds for 10 out of 12 years.

Finally, we evaluate implied growth, *R*, which is expected to increase as the return interval increases. Consider first the average *R* for each of the three intervals: Table 7, Panel B shows that *R* goes from 1.031 (3 months) to 1.093 (12 months) to 1.138 (21 months). All estimates exceed 1, as expected and required by theory, with reasonable orders of magnitude. Looking at individual years, Table 7, Panel B shows that the monotonicity property holds in 8/12 years. Reversed monotonicity occurs only once. Collectively, the evidence lends compelling support for point c) above.

Moving from *V2* to *V1*, the flavor of the results for points a), b), and c) remain. Regarding point a), average estimates of *C* equal 0.984, 0.830, and 0.765, with monotonicity in 7 out of 12 years. For point b), average dynamic goodness-of-fits are 0.320, 0.575, and 0.588, with monotonicity in 6 out of 12 years. With respect to point c), the implied growth *R* equal 1.031, 1.095, and 1.144, with monotonicity in 8 out of 12 years.

[23] Static goodness-of-fit does not change, since the starting point is always the same.

To summarize this section, the empirical evaluation of points a), b), and c) shows that a large majority of elasticities go in the direction predicted by the analytical reasoning in Section 3. This reinforces the validity of the ARF framework and of our main results.

8. Robustness

This section considers two apparent robustness issues, namely, the specifics of the metric that measures the success of explaining returns and the sample from which our results derive. First, our use of rank correlations to measure dynamic goodness-of-fit may seem overly arbitrary and too narrow. It gives rise to the question of whether the thrust of our main results still hold up using some alternative metric. Second, we expand the sample and assess how sensitive our findings in the core empirical Section 6 are to sample selection. We expect that a more expansive data set results in a decline in both static and dynamic goodness-of-fit. A more expansive data set should make it harder to explain value due to the addition of smaller firms, and thus static and dynamic goodness-of-fits would decline.

An Alternative Dynamic Goodness-of-Fit Metric: Implied and Realized Returns

The rank correlation metric used to measure dynamic goodness-of-fit disregards the extent to which implied returns are close to realized returns, in terms of magnitude. As noted earlier, because corr(*realized return*; a + b × *implied return*) does not depend on a and b, it is perfectly possible that the rank correlation is high even though the implied and realized returns differ materially. As an alternative metric that rules out this possibility, consider the cross-sectional median absolute value of the difference between realized returns and model-implied returns.

Alternative dynamic goodness-of-fit ≡ Median($|r_{t+1}$-($implied\ return)_{t+1}|$).[24]

This alternative metric, conceptually equivalent to our static goodness-of-fit metric, clearly ensures that dynamic goodness-of-fit depends only on implied returns' proximity to realized returns. It is simple, robust, provides useful information, and is in the spirit of the rest of the paper. The drawback compared to the original rank correlation metric is also apparent, insofar that it shows poor performance if the implied returns are systematically biased in either direction.

TABLE 8: Median return errors

Results in Table 8 show that using the alternative dynamic goodness-of-fit metric does not change any of the previous findings:

- *V2* continues to exhibit dual dominance over *V1*. Specifically, using M as an estimator of *C*, *V2* dominates *V1* for all years, in terms of alternative dynamic goodness-of-fit. In terms of static goodness-of-fit, *V2* dominates *V1* in all years but one, as before.
- OLS-derived implied returns continue to show worse dynamic goodness-of-fit compared to model-implied returns using the ARF framework. Focusing on *V2*, the average yearly median return error is 12.7 percentage points for M, whereas the OLS-derived implied returns have an average median error of 13.8 percentage points. Turning to the pooled setting, results support the superiority of ARF, as when our original dynamic goodness-of-fit metric was used. The median return error for ARF is 12.8 percentage points, compared to 16.1 percentage points for OLS-derived implied returns. For *V1*, the median return errors are 14.3 and 18.4 percentage points, respectively.

[24] As with static goodness-of-fit, the alternative dynamic goodness-of-fit measure—the median return error—is *decreasing* when explanatory power increases. Zero implies perfect explanation.

To summarize, the evidence rejects, at least with respect to the core research issues, the notion that our findings depend on the use of rank correlations as a dynamic goodness-of-fit metric.

Expanding the sample

Our selection of sample firms has been predicated on the idea that the use of analyst forecasts makes little sense unless they are of reasonable quality. Such is unlikely to be the case if one goes far back in time, or if one expands the pool of firms that can be included. This motivated our choice of time and scope: starting our sample period in 2003 and only including S&P500 firms. Nonetheless, it is of interest to expand this narrow sample to see how ARF performs and whether the framework can guide what to expect. In particular, there are good prospects of the static goodness-of-fit being considerably worse in an expanded sample, and one should thus also expect a corresponding decline in dynamic goodness-of-fit. That said, there are no compelling reasons why the estimates of *C* should change when the sample is expanded; the ARF framework and the *C* estimator should be inherently robust. The presumed increased pervasiveness of outliers should cause no concern. Finally, the intrinsic merit of ARF when explaining returns should remain; there is no reason to believe that an expanded sample would improve the relative performance of OLS as compared to using ARF, with method M estimating *C*.

The expanded dataset requires firms to be covered by at least three analysts, with estimates of both *EPS1* and *EPS2*. The period covered starts in 1990 and ends in 2016, for a total of 27 years and a total of 45,977 firm-year observations. Table 9 displays the results which bear on the points discussed in the previous paragraph.

TABLE 9: Expanded sample

Compared to the original sample used up until this section, a material deterioration in the static goodness-of-fit takes place, for both *V1* and *V2*. *V2* goes from the previous average score of 20.4% to 25.4%. Similarly, *V1* goes from 22.9% to 28.2%. Reassuringly, the dynamic goodness-of-fits are unambiguously consistent with the static goodness-of-fits: as before, dual dominance applies. Specifically, in case of *V2*, the prior dynamic goodness-of-fit of 64.5% now becomes 63.8%; for *V1* the dynamic goodness-of-fits actually increases from 57.5% and 58.5% (which, to be sure, still implies dual dominance, since it is lower than the goodness-of-fit for *V2*).

The average estimates of *C*, based on method M, remain roughly the same even as the sample expands. The average *C* for *V1* increases by 2.1 percentage points, and the average *C* for *V2* decreases by 0.6 percentage points.

Predictably, ARF continues to perform well in explaining returns compared to OLS also using the expanded sample. For valuation heuristic *V2*, the average dynamic goodness-of-fit is 63.1% for OLS compared to 63.8% for ARF. The latter outperforms the former in 20 out of 27 years, in terms of explanatory power. Pooling all observations and redoing the analysis results in a dynamic goodness-of-fit of 51.8% for OLS compared to 68.7% for ARF, a material difference in our opinion.

Turning next to valuation heuristic *V1*, the conclusions from the previous paragraph are reinforced. For individual years, OLS gives an average dynamic goodness-of-fit of 57.7% compared to 59.0% for ARF.[25] The latter outperforms the former in 17 out of 27 years. In a pooled setting, the difference in explanatory power is again larger: 46.9% for OLS compared to 62.9% for ARF.

[25] The averages exclude 2016, since OLS renders negative dynamic goodness-of-fit in that year. Including 2016 would, of course, increase the difference in average explanatory power.

9. Summary of findings and potential extensions

In this paper, we investigate one of the central questions in accounting research: what explains stock market returns? Instead of emphasizing traditional regression methodologies, we focus on the question's conceptual foundation and the relation between firm valuation (static perspective) and stock returns (dynamic perspective). Concretely, we investigate whether an explicit role for valuation can explain returns in a more parsimonious way—simpler, using fewer degrees of freedom, without loss of explanatory power—than using traditional OLS regressions. We introduce an autoregressive framework, called ARF, which stipulates the stochastic behavior of the valuation gap—the difference between actual stock prices and a valuation model's estimated values. ARF relies on one single parameter, *C*. Through this framework, we empirically evaluate the idea that to explain returns one benefits from connecting with concepts of valuation. This leads to three propositions, the evaluation of which are summarized below.

First, goodness-of-fit in valuation space and goodness-of-fit in return space are intrinsically connected and move in tandem; high explanatory power in one dimension implies high explanatory power in the other. Second, to estimate the single attenuation parameter *C*, we rely on a simplistic, non-parametric method called M. Though truly simple, method M yields estimates of *C* and dynamic goodness-of-fit which are strikingly similar to those deriving from a method that maximizes dynamic goodness-of-fit, M*. This points to M being an efficient estimator; simplicity does not come at the cost of explanatory power. In a setting with pooled yearly observations, the dynamic goodness-of-fit exceeds 73%. Third, the explanatory power of the ARF framework should be no worse than that of traditional OLS regressions, which have the advantage of not restricting the coefficients related to the independent variables. In this regard, it is noteworthy that even if one introduces fixed effect, with hundreds of unrestricted parameters, traditional OLS regressions do not explain returns any better than the ARF framework with its single parameter *C*. In our view, these results testify to the power of the idea that to explain returns, one ought to consider the linkage to concepts of valuation.

The paper has been confined by specifying relatively unsophisticated valuation heuristics that rely solely on EPS forecasts. There ought to be potential for achieving better goodness-of-fits by specifying heuristics based on additional value-relevant information. In particular, one can hypothesize that current book values and current earnings are relevant due to their influence on future earnings and earnings growth.[26] A related improvement would be to recognize the valuation heuristic's dependence on cost of capital.[27] Finally, to the extent that one relies on analyst forecasts, these should ideally be adjusted for predictable forecast errors (see for example Gode and Mohanran 2013 and Easton and Monahan 2015 for an overview).

[26] Hou, Hugon, and Lyle (2016) show how current accounting data can be mapped into growth.

[27] The literature on the measurement of cost of capital is vast. To evaluate the performance of alternative cost of capital measures, see Li and Mohanram (2014).

REFERENCES

Ali, A., and P. Zarowin. 1992. The Role of Earnings Levels in Annual Earnings Returns Studies. *Journal of accounting research* 30 (2): 286-296.

Bach, C. and P. O. Christensen. 2016. Consumption-based equity valuation. *Review of Accounting Studies* 21: 1149.

Ball, R., and P. Brown. 1968. An Empirical Evaluation of Accounting Income Numbers. *Journal of accounting research* 6 (2): 159-178.

Barth, M. E., W. H. Beaver, and W. R. Landsman. 1992. The Market Valuation Implications of Net Periodic Pension Cost Components. *Journal of Accounting & Economics* 15 (1): 27-62.

Basu, S. 1997. The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting & Economics* 24 (1): 3-37.

Black, F. 1986. Noise. *Journal of Finance* 41 (3): 529-543.

Brown, S., K. Lo, and T. Lys. 1999. Use of R2 in Accounting Research: Measuring Changes in Value Relevance Over the Last Four Decades. *Journal of Accounting and Economics* 28 (2): 83-115.

Callen, J. L., and D. Segal. 2005. Empirical Tests of the Feltham-Ohlson (1995) Model. *Review of Accounting Studies* 10 (4): 409-429.

Cao, Y., L. A. Myers, and T. Sougiannis. 2011. Does earnings acceleration convey information? *Review of Accounting Studies* 16: 812-842.

Cheng, C. S. A., W. S. Hopwood, and J. C. McKeown. 1992. Nonlinearity and Specification Problems in Unexpected Earnings Response Regression Model. *The Accounting Review* 67 (3): 579-598.

Claus, J., and J. Thomas. 2001. Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stocks. *Journal of Finance* 56 (5): 1629-1666.

Collins, D. W., E. L. Maydew, and I. S. Weiss. 1997. Changes in the value-relevance of earnings and book values over the past forty years. *Journal of Accounting & Economics* 24 (1): 39-67.

Copeland, T., A. Dolgoff, and A. Moel. 2004 The Role of Expectations in Explaining the Cross-Section of Stock Returns. *Review of Accounting Studies* 9 (2/3): 149-188.

Cornell, B., and W. R. Landsman. 1989. Security Price Response to Quarterly Earnings Announcements and Analysts Forecast Revisions. *Accounting Review* 64 (4): 680-692.

Easton, P. D., and T. S. Harris. 1991. Earings as an Explanatory Variable for Returns. *Journal of accounting research* 29 (1): 19-36.

Easton, P. D., T. S. Harris, and J. A. Ohlson. 1992. Aggregate Accounting Earnings Can Explain Most of Security Returns – the Case of Long Return Intervals. *Journal of Accounting & Economics* 15 (2-3): 119-142.

Easton, P. D., and S. J. Monahan. (2016) Review of Recent Research on Improving Earnings Forecasts and Evaluating Accounting-based Estimates of the Expected Rate of Return on Equity Capital. Abacus, 52: 35–58.

Francis, J., and K. Schipper. 1999. Have Financial Statements Lost Their Relevance? *Journal of accounting research* 37 (2): 319-352.

Francis, J., K. Schipper, and L. Vincent. 2003. The Relative and Incremental Explanatory Power of Earnings and Alternative (to Earnings) Performance Measures for Returns. *Contemporary Accounting Research* 20 (1): 121-164.

Francis, J. and L. Soffer. 1997. The relative informativeness of analysts' stock recommendations and earnings forecast revisions. *Journal of Accounting Research* 35: 193-211.

Frankel, R., and C. M. C. Lee. 1998. Accounting valuation, market expectation, and cross-sectional stock returns. *Journal of Accounting & Economics* 25 (3): 283-319.

Gu, Z. 2007. Across-sample Incomparability of R2s and Additional Evidence on Value Relevance Changes Over Time. *Journal of Business Finance & Accounting* 34 (7/8): 1073-1098.

Hayn, C. 1995. The Information-Content of Losses. *Journal of Accounting & Economics* 20 (2): 125-153.

Holthausen, R. W., and R. Watts. 2001. The relevance of the value-relevance literature for financial accounting standard setting. *Journal of Accounting & Economics* 31 (1-3): 3-75.

Hou, Yu, A. Hugon, and M. R. Lyle. 2016. Economic Growth Forecasts Embedded in the Cross-Section of Accounting Data. *Available on SSRN*.

Lee, C., J. Myers, and B. Swaminathan, 1999, What is the intrinsic value of the Dow? Journal of Finance 54, 1693-1741

Lev, B. 1989. On the Usefulness of Earnings and Earnings Research: Lessons and Directions from Two Decades of Empirical Research. *Journal of accounting research* 27: 153.

Li, K.K. and P. Mohanram. 2014. Evaluating cross-sectional forecasting models for implied cost of capital. *Review of Accounting Studies* 19: 1152.

Liu, J., D. Nissim, and J. Thomas. 2002. Equity Valuation Using Multiples. *Journal of Accounting Research* 40 (1): 135-172.

Liu, J., and J. Thomas. 2000. Stock Returns and Accounting Earnings. *Journal of accounting research* 38 (1): 71-101.

Livnat, J., and P. Zarowin. 1990. The Incremental Information-Content of Cash-Flow Components. *Journal of Accounting & Economics* 13 (1): 25-46.

Mohanram, P. and D. Gode. Removing predictable analyst forecast errors to improve implied cost of equity estimates. *Review of Accounting Studies* (2013) 18: 443

Ohlson, J. A. and E. Johannesson, 2016. Equity Value as a Function of (eps1, eps2, dps1, bvps, beta): Concepts and Realities. *ABACUS* 52 (1): 70-99.

Ohlson, J. A., and S. Kim. 2015. Linear valuation without OLS: the Theil-Sen estimation approach. *Review of Accounting Studies* 20 (1): 395-435.

Ohlson, J. A., and S. H. Penman. 1992. Disaggregated accounting data as explanatory variables for returns. *Journal of Accounting, Auditing & Finance* 7 (4): 553-573.

Park, C. W., and E. K. Stice. 2000. Analyst Forecasting Ability and the Stock Price Reaction to Forecast Revisions. *Review of Accounting Studies* 5 (3): 259-272.

Penman, S. 2012. *Financial Statement Analysis and Security Valuation*, 5th edition, McGraw-Hill Education, London.

Roll, R. 1988. R2. *The Journal of Finance.* 43 (3): 541-566.

Wahlen, J. M., S. P., Baginski, M. T. Bradshaw, and C. P. Stickney. 2015. *Financial reporting, financial statement analysis, and valuation: A strategic perspective*.

Wilcox, R. R. 2010). *Fundamentals of modern statistical methods: Substantially improving power and accuracy* (2nd ed.). New York: Springer.

EXHIBITS

TABLE 1

Forward *E*/*P* ratios: descriptive statistics

		$EPS1_t/P_t$			$EPS2_t/P_t$		
Year	*obs.*	*p25*	*p50*	*p75*	*p25*	*p50*	*p75*
2003	503	0.04	0.062	0.087	0.053	0.073	0.095
2004	499	0.04	0.057	0.071	0.048	0.064	0.078
2005	502	0.044	0.058	0.072	0.053	0.064	0.078
2006	503	0.047	0.058	0.072	0.054	0.066	0.079
2007	506	0.049	0.059	0.071	0.057	0.068	0.081
2008	502	0.053	0.066	0.082	0.062	0.076	0.093
2009	506	0.058	0.079	0.103	0.072	0.093	0.119
2010	501	0.047	0.064	0.078	0.061	0.074	0.09
2011	499	0.054	0.068	0.084	0.066	0.079	0.093
2012	501	0.057	0.071	0.087	0.065	0.081	0.099
2013	503	0.053	0.065	0.083	0.061	0.074	0.094
2014	500	0.048	0.058	0.071	0.056	0.066	0.079
Average	502	0.049	0.064	0.08	0.059	0.073	0.09

This table provides statistics related to the two forward EPS-to-price ratios. The sample consists of S&P500 firms as of March 31, of each year. This set of firms changes every year so as to reflect changes in the index, and firms are grouped according to fiscal year. The forward earnings estimates derive from IBES and the price data is obtained from CRSP. The 25^{th}, 50^{th}, and 75^{th} percentiles are indicated by p25, p50, and p75, respectively.

TABLE 2

Estimates of *C* using method M

	C	
Year	*V1*	*V2*
2003	0.922	0.916
2004	0.827	0.796
2005	0.927	0.905
2006	0.720	0.802
2007	0.793	0.789
2008	0.460	0.451
2009	0.696	0.643
2010	0.805	0.870
2011	0.822	0.932
2012	1.084	1.049
2013	0.894	0.891
2014	1.010	1.014
Average	0.830	0.838

Table 2 shows yearly and average estimates of the contraction parameter *C*, obtained through the estimation method M. M gives *C* as the yearly median cross-sectional value of $(P_{t+1} - V_{t+1})/(P_t - V_t)$. *V1* is the valuation heuristic based on EPS1, *V2* is the valuation heuristic based on EPS2. The number of observations are approximately 500 for each year.

TABLE 3

Panel A: Dual dominance, *V2* compared to *V1*

	StatGoF		**DynGoF**	
Year	*V1*	*V2*	*V1*	*V2*
2003	0.321	0.276	0.513	0.565
2004	0.267	0.222	0.633	0.697
2005	0.241	0.186	0.657	0.704
2006	0.215	0.190	0.511	0.559
2007	0.186	0.176	0.691	0.730
2008	0.215	0.204	0.669	0.731
2009	0.247	0.237	0.518	0.678
2010	0.228	0.192	0.445	0.571
2011	0.217	0.175	0.446	0.519
2012	0.199	0.210	0.615	0.651
2013	0.220	0.213	0.572	0.658
2014	0.191	0.170	0.626	0.674
Average	0.229	0.204	0.575	0.645

Table 3, Panel A shows yearly and average dynamic and static goodness-of-fit for the two valuation heuristics *V1* (based on EPS1) and V2 (based on EPS2). The static goodness-of-fit ("StatGoF") is defined as the cross-sectional median percentage absolute value deviation of model value from actual value. Dynamic goodness-of-fit ("DynGoF") is the rank correlation between realized returns and implied returns. Dynamic goodness-of-fit is calculated using method M to estimate the contraction parameter *C*. M gives *C* as the yearly median cross-sectional value of $(P_{t+1} - V_{t+1})/(P_t - V_t)$. The number of observations are approximately 500 for each year.

Panel B: Dual dominance, pooled data

	StatGoF		DynGoF		C	
Year	*V1*	*V2*	*V1*	*V2*	*V1*	*V2*
Pooled	0.223	0.203	0.662	0.731	0.865	0.874

Table 3, Panel B shows dynamic and static goodness-of-fit for the two valuation heuristics *V1* (based on EPS1) and *V2* (based on EPS2) for the sample in which the yearly data are pooled. Panel B also shows the contraction parameter *C* for both valuation heuristics. The static goodness-of-fit ("StatGoF") is defined as the cross-sectional median percentage absolute value deviation of model value from actual value. Dynamic goodness-of-fit ("DynGoF") is the rank correlation between realized return and implied return. Dynamic goodness-of-fit is calculated using method M to estimate the contraction parameter *C*. M gives *C* as the yearly median cross-sectional value of $(P_{t+1} - V_{t+1})/(P_t - V_t)$.

Panel C: Dual dominance, 3 and 21 months return intervals

	StatGoF		DynGoF			
			3 months		21 months	
Year	*V1*	*V2*	*V1*	*V2*	*V1*	*V2*
2003	0.321	0.276	0.180	0.258	0.594	0.681
2004	0.267	0.222	0.278	0.281	0.665	0.734
2005	0.241	0.186	0.385	0.424	0.679	0.724
2006	0.215	0.190	0.296	0.266	0.641	0.744
2007	0.186	0.176	0.414	0.492	0.661	0.770
2008	0.215	0.204	0.515	0.543	0.536	0.650
2009	0.247	0.237	0.160	0.254	0.540	0.698
2010	0.228	0.192	0.332	0.369	0.398	0.553
2011	0.217	0.175	0.237	0.303	0.558	0.681
2012	0.199	0.210	0.305	0.391	0.604	0.719
2013	0.220	0.213	0.345	0.405	0.567	0.702
2014	0.191	0.170	0.395	0.389	0.615	0.711
Average	0.229	0.204	0.320	0.365	0.588	0.697

Table 3, Panel C shows yearly and average dynamic and static goodness-of-fit for the two valuation heuristics *V1* (based on EPS1) and *V2* (based on EPS2). It differs from Panel A in that the return interval used to calculate dynamic goodness-of-fit is 3 and 21 months (as opposed to 12 months in Panel A). The static goodness-of-fit ("StatGoF") is defined as the cross-sectional median percentage absolute value deviation of model value from actual value. Dynamic goodness-of-fit ("DynGoF") is the rank correlation between realized return and implied return. Dynamic goodness-of-fit is calculated using method M to estimate the contraction parameter *C*. M gives *C* as the yearly median cross-sectional value of $(P_{t+1} - V_{t+1})/(P_t - V_t)$. The number of observations are approximately 500 for each year.

TABLE 4

Dynamic goodness-of-fit and the efficacy of M

	DynGoF				C*/C			
	V1		V2		V1		V2	
Year	*M**	*M*	*M**	*M*	*M**	*M*	*M**	*M*
2003	0.519	0.513	0.573	0.565	0.859	0.922	0.767	0.916
2004	0.636	0.633	0.698	0.697	0.764	0.827	0.791	0.796
2005	0.658	0.657	0.710	0.704	0.976	0.927	0.872	0.905
2006	0.515	0.511	0.562	0.559	0.808	0.720	0.913	0.802
2007	0.692	0.691	0.731	0.730	0.820	0.793	0.799	0.789
2008	0.684	0.669	0.751	0.731	0.582	0.460	0.626	0.451
2009	0.519	0.518	0.678	0.678	0.650	0.696	0.668	0.643
2010	0.448	0.445	0.578	0.571	0.880	0.805	0.982	0.870
2011	0.463	0.446	0.526	0.519	0.957	0.822	1.012	0.932
2012	0.616	0.615	0.651	0.651	1.057	1.084	1.044	1.049
2013	0.573	0.572	0.658	0.658	0.882	0.894	0.886	0.891
2014	0.630	0.626	0.683	0.674	1.075	1.010	1.099	1.014
Average	0.579	0.575	0.650	0.645	0.859	0.830	0.872	0.838

Table 4 compares the efficacy of method M in calculating the contraction parameter *C* vis-à-vis the maximization method M*, and shows yearly and average dynamic goodness-of-fit and values of the contraction parameter *C* for the two valuation heuristics *V1* (based on EPS1) and *V2* (based on EPS2). Dynamic goodness-of-fit ("DynGoF") is the rank correlation between realized return and implied return. Method M gives *C* as the yearly median cross-sectional value of $(P_{t+1} - V_{t+1})/(P_t - V_t)$. M* gives the *C* that maximizes the rank correlation between realized return and implied return ("*C**"). The number of observations are approximately 500 for each year.

TABLE 5

Panel A: M' and M'' for *V1*

V1

	M*		M		M'					M''		
Year	*DynGoF*	*C**	*DynGoF*	*C*	*DynGoF*	*Intercept*	V_{t+1}/P_t	V_t/P_t	*Implied C*	*DynGoF*	P_t	V_t (= C)
2003	0.519	0.859	0.513	0.922	0.512	0.033	0.587	-0.463	0.788	0.503	0.248	1.012
2004	0.636	0.764	0.633	0.827	0.636	-0.120	0.548	-0.425	0.776	0.608	0.175	0.966
2005	0.658	0.976	0.657	0.927	0.657	-0.028	0.710	-0.683	0.962	0.654	0.090	1.017
2006	0.515	0.808	0.511	0.720	0.505	-0.125	0.496	-0.321	0.646	0.502	0.193	0.902
2007	0.692	0.820	0.691	0.793	0.691	-0.182	0.851	-0.677	0.795	0.689	0.172	0.776
2008	0.684	0.582	0.669	0.460	0.683	-0.395	0.509	-0.329	0.647	0.638	0.300	0.295
2009	0.519	0.650	0.518	0.696	0.513	-0.028	0.585	-0.301	0.515	0.467	0.300	1.026
2010	0.448	0.880	0.445	0.805	0.446	0.027	0.442	-0.371	0.839	0.430	0.453	0.695
2011	0.463	0.957	0.446	0.822	0.445	-0.005	0.547	-0.447	0.816	0.453	0.187	0.865
2012	0.616	1.057	0.615	1.084	0.614	0.075	0.612	-0.666	1.088	0.615	0.106	1.038
2013	0.573	0.882	0.572	0.894	0.571	-0.029	0.681	-0.581	0.853	0.519	0.112	1.090
2014	0.630	1.075	0.626	1.010	0.630	0.065	0.648	-0.692	1.068	0.625	0.122	1.000
Average	0.579	0.859	0.575	0.830	0.575	-0.059	0.601	-0.496	0.816	0.559	0.205	0.890

Panel B: M' and M'' for *V2*

V2

	M*		M		M'					M''		
Year	*DynGoF*	*C**	*DynGoF*	*C*	*DynGoF*	*Intercept*	V_{t+1}/P_t	V_t/P_t	*Implied C*	*DynGoF*	P_t	V_t (= *C*)
2003	0.573	0.767	0.565	0.916	0.569	-0.017	0.723	-0.615	0.850	0.523	0.169	1.118
2004	0.698	0.791	0.697	0.796	0.697	-0.155	0.700	-0.556	0.795	0.678	0.167	0.964
2005	0.710	0.872	0.704	0.905	0.703	-0.054	0.768	-0.713	0.929	0.699	0.104	0.992
2006	0.562	0.913	0.559	0.802	0.546	-0.133	0.620	-0.418	0.674	0.559	0.139	0.942
2007	0.731	0.799	0.730	0.789	0.730	-0.204	1.016	-0.806	0.793	0.730	0.148	0.806
2008	0.751	0.626	0.731	0.451	0.749	-0.413	0.706	-0.396	0.561	0.669	0.347	0.249
2009	0.678	0.668	0.678	0.643	0.678	-0.178	0.830	-0.552	0.665	0.572	0.168	1.174
2010	0.578	0.982	0.571	0.870	0.573	0.027	0.599	-0.526	0.878	0.560	0.338	0.787
2011	0.526	1.012	0.519	0.932	0.525	0.040	0.653	-0.644	0.986	0.512	0.144	0.894
2012	0.651	1.044	0.651	1.049	0.647	0.040	0.719	-0.720	1.001	0.647	0.129	1.002
2013	0.658	0.886	0.658	0.891	0.658	-0.033	0.789	-0.696	0.882	0.594	0.068	1.135
2014	0.683	1.099	0.674	1.014	0.680	0.069	0.687	-0.736	1.071	0.675	0.118	1.021
Average	0.650	0.872	0.645	0.838	0.646	-0.084	0.734	-0.615	0.841	0.618	0.170	0.924

Table 5 comprises two panels, Panel A (for valuation heuristic *V1*, based on EPS1) and Panel B (for valuation heuristic *V2*, based on EPS2). The two panels show estimates of the contraction parameter *C* and associated dynamic goodness-of-fit ("DynGoF") for all four estimation methods considered: M*, M, M' and M''. M* gives the *C* that maximizes the rank correlation between realized return and implied return; M gives *C* as the yearly median cross-sectional value of $(P_{t+1} - V_{t+1})/(P_t - V_t)$. M' gives *C* via Theil-Sen estimations of realized return on $k_1 \times V_{t+1}/P_t + k_2 \times V_t/P_t + k_3$. The ratio $\hat{k}_2/\hat{k}_1$ provides an estimate of *C*. M'' gives *C* via Theil-Sen estimations of V_{t+1} on $k_1 \times P_t + k_2 \times V_t$. $\hat{k}_2$ provides an estimate of *C*. Dynamic goodness-of-fit is the rank correlation between realized return and implied return. The number of observations are approximately 500 for each year.

TABLE 6

ARF compared to OLS—dynamic goodness-of-fit

	OLS		**ARF(M)**	
Year	*V1*	*V2*	*V1*	*V2*
2003	0.460	0.480	0.513	0.565
2004	0.595	0.676	0.633	0.697
2005	0.658	0.702	0.657	0.704
2006	0.514	0.557	0.511	0.559
2007	0.648	0.710	0.691	0.730
2008	0.646	0.734	0.669	0.731
2009	-0.428	0.669	0.518	0.678
2010	0.429	0.573	0.445	0.571
2011	0.455	0.524	0.446	0.519
2012	0.615	0.641	0.615	0.651
2013	0.548	0.647	0.572	0.658
2014	0.627	0.673	0.626	0.674
Average	NA*	0.632	0.575	0.645
Pooled	0.386	0.643	0.662	0.731
w. year FE	0.480	0.644	NA	NA
w. firm FE	0.231	0.729	NA	NA
w. year and firm FE	0.513	0.727	NA	NA

Table 6 compares explanatory power, as measured by dynamic goodness-of-fit, of traditional OLS regressions to that of our autoregressive framework when method M is used to estimate the contraction parameter *C* (“ARF(M)”). The table shows results for individual years, averages across individual years, and for the setting in which all yearly data are pooled. The latter includes permutations with year fixed effects, firm fixed effects and year and firm fixed effects. Results are shown for both valuation heuristic *V1* (based on EPS1) and *V2* (based on EPS2). Dynamic goodness-of-fit is the rank correlation between realized return and implied return. The implied returns for OLS result from estimations of realized return$_{t+1}$ on V_{t+1}/P_t and V_t/P_t, along with an intercept, with or without fixed effects. The number of observations are approximately 500 for each year.

TABLE 7

Panel A: Return intervals and robustness—Dynamic goodness-of-fit and *C*

	DynGoF						**C**					
	V1			**V2**			**V1**			**V2**		
Year	*3 mos.*	*12 mos.*	*21 mos.*	*3 mos.*	*12 mos.*	*21 mos.*	*3 mos.*	*12 mos.*	*21 mos.*	*3 mos.*	*12 mos.*	*21 mos.*
2003	0.180	0.513	0.594	0.258	0.565	0.681	0.949	0.922	0.897	0.891	0.916	0.885
2004	0.278	0.633	0.665	0.281	0.697	0.734	1.073	0.827	0.858	1.044	0.796	0.793
2005	0.385	0.657	0.679	0.424	0.704	0.724	1.042	0.927	0.913	0.989	0.905	0.851
2006	0.296	0.511	0.641	0.266	0.559	0.744	0.967	0.720	0.822	0.973	0.802	0.864
2007	0.414	0.691	0.661	0.492	0.730	0.770	1.026	0.793	0.606	0.986	0.789	0.535
2008	0.515	0.669	0.536	0.543	0.731	0.650	0.956	0.460	0.270	0.957	0.451	0.213
2009	0.160	0.518	0.540	0.254	0.678	0.698	0.895	0.696	0.640	0.835	0.643	0.737
2010	0.332	0.445	0.398	0.369	0.571	0.553	0.956	0.805	0.506	0.963	0.870	0.714
2011	0.237	0.446	0.558	0.303	0.519	0.681	0.964	0.822	0.803	1.008	0.932	0.959
2012	0.305	0.615	0.604	0.391	0.651	0.719	1.047	1.084	0.938	1.040	1.049	0.903
2013	0.345	0.572	0.567	0.405	0.658	0.702	0.963	0.894	0.937	0.931	0.891	1.011
2014	0.395	0.626	0.615	0.389	0.674	0.711	0.975	1.010	0.991	1.000	1.014	1.001
Average	0.320	0.575	0.588	0.365	0.645	0.697	0.984	0.830	0.765	0.968	0.838	0.789

Panel B: Return intervals and robustness—Implied *R*

V1

	3 mos.		12 mos.		21 mos.		Implied R 3 mos.	Implied R 12 mos.	Implied R 21 mos.
Year	*Pt*	*Vt*	*Pt*	*Vt*	*Pt*	*Vt*	*sum coeff.*	*sum coeff.*	*sum coeff.*
2003	0.002	1.090	0.248	1.012	0.339	1.102	1.092	1.260	1.441
2004	0.015	1.009	0.175	0.966	0.199	0.989	1.024	1.141	1.187
2005	-0.008	1.043	0.090	1.017	0.105	1.094	1.035	1.107	1.199
2006	0.011	0.959	0.193	0.902	0.281	0.867	0.971	1.095	1.147
2007	0.002	1.064	0.172	0.776	0.199	0.472	1.066	0.948	0.671
2008	0.026	0.984	0.300	0.295	0.312	0.403	1.010	0.595	0.715
2009	0.016	1.066	0.300	1.026	0.428	1.018	1.082	1.325	1.446
2010	0.079	0.938	0.453	0.695	0.479	0.620	1.017	1.148	1.099
2011	0.008	1.009	0.187	0.865	0.237	0.877	1.018	1.052	1.114
2012	0.005	0.977	0.106	1.038	0.153	1.147	0.982	1.144	1.300
2013	0.009	1.027	0.112	1.090	0.152	1.146	1.036	1.202	1.297
2014	0.008	1.029	0.122	1.000	0.128	0.983	1.037	1.123	1.110
Average	0.014	1.016	0.205	0.890	0.251	0.893	1.031	1.095	1.144

V2

	3 mos.		12 mos.		21 mos.		Implied R 3 mos.	Implied R 12 mos.	Implied R 21 mos.
Year	*Pt*	*Vt*	*Pt*	*Vt*	*Pt*	*Vt*	*sum coeff.*	*sum coeff.*	*sum coeff.*
2003	0.019	1.096	0.169	1.118	0.210	1.207	1.115	1.286	1.417
2004	0.011	1.006	0.167	0.964	0.196	0.979	1.017	1.130	1.175
2005	0.004	1.013	0.104	0.992	0.130	1.046	1.017	1.096	1.177
2006	0.007	0.966	0.139	0.942	0.202	0.940	0.973	1.081	1.142
2007	0.004	1.063	0.148	0.806	0.274	0.394	1.067	0.953	0.668
2008	0.035	0.981	0.347	0.249	0.331	0.405	1.016	0.596	0.736
2009	0.054	1.040	0.168	1.174	0.286	1.157	1.094	1.341	1.443
2010	0.071	0.931	0.338	0.787	0.339	0.706	1.001	1.125	1.044
2011	0.011	0.997	0.144	0.894	0.187	0.930	1.008	1.038	1.118
2012	0.015	0.953	0.129	1.002	0.148	1.155	0.968	1.132	1.303
2013	0.015	1.019	0.068	1.135	0.105	1.199	1.035	1.203	1.303
2014	0.010	1.046	0.118	1.021	0.100	1.035	1.057	1.140	1.135
Average	0.021	1.009	0.170	0.924	0.209	0.930	1.031	1.093	1.138

Table 7 comprises two panels, Panel A (for dynamic goodness-of-fit—"DynGoF"—and the contraction parameter *C*) and Panel B (for implied *R*). The two panels show how these three constructs vary with return interval length (3 months, 12, months, and 21 months), and considers both valuation heuristic *V1* (based on EPS1) and *V2* (based on EPS2). In Panel A, estimates of the contraction parameter *C* and the associated dynamic goodness-of-fits derive from method M to estimate *C*. M gives *C* as the yearly median cross-sectional value of $(P_{t+1} - V_{t+1})/(P_t - V_t)$. Panel B shows results pertaining to method M'': Theil-Sen estimations of V_{t+1} on $k_1 \times P_t + k_2 \times V_t$. $\hat{k}_2$. Per theory, the sum of the two coefficients gives an estimate of implied *R*. Dynamic goodness-of-fit is the rank corr elation between realized return and implied return. The number of observations are approximately 500 for each year.

TABLE 8

Median return errors

	StatGoF		Median return error			
			V1		V2	
Year	*V1*	*V2*	*ARF(M)*	*OLS*	*ARF(M)*	*OLS*
2003	0.321	0.276	0.174	0.201	0.162	0.179
2004	0.267	0.222	0.150	0.140	0.126	0.133
2005	0.241	0.186	0.114	0.119	0.104	0.121
2006	0.215	0.190	0.125	0.121	0.114	0.106
2007	0.186	0.176	0.135	0.153	0.121	0.130
2008	0.215	0.204	0.137	0.152	0.118	0.137
2009	0.247	0.237	0.277	0.403	0.204	0.330
2010	0.228	0.192	0.145	0.137	0.129	0.109
2011	0.217	0.175	0.142	0.116	0.128	0.107
2012	0.199	0.210	0.105	0.096	0.096	0.093
2013	0.220	0.213	0.125	0.116	0.111	0.108
2014	0.191	0.170	0.108	0.109	0.108	0.105
Average	0.229	0.204	0.145	0.155	0.127	0.138
Pooled	0.223	0.203	0.143	0.184	0.128	0.161

Table 8 provides results for an alternative dynamic goodness-of-fit measure by calculating, for each year, the median of the absolute value of the relative difference between realized return and model-implied return. The table considers both valuation heuristic *V1* (based on EPS1) and *V2* (based on EPS2), and complements the rank correlation-based measure of dynamic goodness-of-fit in previous tables. In addition to the median return error for our framework ARF, whose returns are calculated using method M (the contraction parameter C is calculated as the yearly median cross-sectional value of $(P_{t+1} - V_{t+1})/(P_t - V_t)$), the table also shows the median return error for traditional OLS regressions. The implied returns for OLS result from estimations of realized return$_{t+1}$ on V_{t+1}/P_t and V_t/P_t, along with an intercept. The static goodness-of-fit ("StatGoF") is defined as the cross-sectional median percentage absolute value deviation of model value from actual value. The number of observations are approximately 500 for each year.

TABLE 9

Panel A: Expanded sample—*V1*

			ARF (M)		OLS					
Year	*Obs*	*StatGoF*	*DynGoF*	*C*	*DynGoF*	*Intercept*	V_{t+1}/P_t	V_t/P_t	*Implied C*	*adj.* R^2
1990	653	0.244	0.689	0.911	0.685	0.148	0.349	-0.391	1.121	0.276
1991	756	0.263	0.719	0.995	0.681	0.009	0.377	-0.249	0.662	0.331
1992	898	0.258	0.707	0.718	0.705	-0.081	0.696	-0.537	0.773	0.414
1993	1060	0.223	0.592	1.083	0.592	0.097	0.475	-0.499	1.050	0.271
1994	1124	0.262	0.487	0.683	0.462	-0.107	0.255	-0.120	0.472	0.106
1995	1179	0.268	0.592	1.014	0.601	0.274	0.457	-0.523	1.145	0.189
1996	1384	0.266	0.687	0.799	0.670	-0.110	0.506	-0.294	0.581	0.359
1997	1498	0.240	0.601	0.972	0.494	0.421	0.389	-0.550	1.415	0.131
1998	1539	0.248	0.480	0.813	0.452	0.041	0.390	-0.387	0.992	0.082
1999	1394	0.344	0.474	1.034	0.460	0.821	0.462	-0.879	1.904	0.113
2000	1229	0.470	0.799	0.555	0.782	-0.332	0.429	0.017	-0.040	0.422
2001	1252	0.357	0.560	0.831	0.608	-0.030	0.272	-0.139	0.512	0.248
2002	1410	0.348	0.627	0.551	0.646	-0.330	0.371	-0.155	0.419	0.280
2003	1569	0.317	0.534	1.054	0.508	0.539	0.048	-0.060	1.262	0.095
2004	1820	0.279	0.695	0.845	0.703	0.007	0.231	-0.150	0.648	0.271
2005	1945	0.238	0.632	0.979	0.634	0.153	0.297	-0.323	1.085	0.301
2006	2043	0.263	0.587	0.778	0.581	0.010	0.305	-0.249	0.818	0.233
2007	2121	0.224	0.657	0.705	0.666	-0.208	0.383	-0.218	0.570	0.292
2008	2108	0.270	0.555	0.501	0.533	-0.333	0.156	-0.124	0.793	0.127
2009	2111	0.354	0.495	0.882	0.476	0.510	0.573	-0.631	1.101	0.627
2010	2186	0.270	0.513	0.832	0.449	0.006	0.229	-0.083	0.364	0.164
2011	2265	0.271	0.512	0.824	0.447	-0.329	0.765	-0.322	0.422	0.762
2012	2279	0.261	0.597	0.969	0.599	0.017	0.651	-0.567	0.871	0.561
2013	2338	0.268	0.509	0.986	0.512	0.247	0.318	-0.339	1.068	0.215
2014	2544	0.251	0.543	0.928	0.546	-0.037	0.787	-0.667	0.847	0.812
2015	2657	0.266	0.491	0.817	0.501	-0.125	0.763	-0.607	0.796	0.749
2016	2615	0.304	0.470	0.929	-0.364	0.394	-0.214	0.058	0.269	0.214
Average	1703	0.282	0.585	0.851	0.542	0.062	0.397	-0.333	0.812	0.320
Pooled	45977	0.276	0.629	0.843	0.469	0.079	0.401	-0.362	0.902	0.289

Panel B: Expanded sample—*V2*

			ARF (M)		OLS					
Year	*Obs*	*StatGoF*	*DynGoF*	*C*	*DynGoF*	*Intercept*	V_{t+1}/P_t	V_t/P_t	*Implied C*	*adj.* R^2
1990	653	0.211	0.715	0.851	0.713	-0.026	0.649	-0.536	0.826	0.400
1991	756	0.225	0.764	0.882	0.742	-0.041	0.520	-0.373	0.717	0.398
1992	898	0.210	0.709	0.707	0.708	-0.169	0.854	-0.653	0.764	0.485
1993	1060	0.189	0.615	1.008	0.612	0.012	0.656	-0.614	0.935	0.354
1994	1124	0.222	0.554	0.718	0.551	-0.105	0.331	-0.207	0.626	0.142
1995	1179	0.246	0.659	0.992	0.658	0.189	0.487	-0.478	0.982	0.250
1996	1384	0.238	0.731	0.708	0.727	-0.205	0.697	-0.416	0.597	0.473
1997	1498	0.208	0.636	0.931	0.636	0.334	0.472	-0.567	1.202	0.181
1998	1539	0.213	0.530	0.823	0.516	0.041	0.504	-0.485	0.963	0.110
1999	1394	0.327	0.530	1.037	0.516	0.768	1.052	-1.363	1.296	0.139
2000	1229	0.461	0.809	0.534	0.807	-0.438	0.653	-0.170	0.260	0.471
2001	1252	0.328	0.667	0.750	0.688	-0.124	0.424	-0.242	0.572	0.357
2002	1410	0.299	0.694	0.565	0.688	-0.376	0.497	-0.221	0.444	0.350
2003	1569	0.284	0.574	1.013	0.565	0.472	0.120	-0.092	0.767	0.133
2004	1820	0.239	0.754	0.804	0.747	-0.075	0.436	-0.298	0.682	0.429
2005	1945	0.195	0.678	0.919	0.671	0.155	0.414	-0.455	1.099	0.373
2006	2043	0.223	0.622	0.810	0.626	-0.070	0.419	-0.290	0.691	0.327
2007	2121	0.201	0.701	0.703	0.705	-0.236	0.573	-0.367	0.639	0.388
2008	2108	0.249	0.607	0.481	0.600	-0.321	0.280	-0.210	0.751	0.187
2009	2111	0.338	0.604	0.857	0.607	0.216	0.745	-0.617	0.828	0.816
2010	2186	0.237	0.603	0.860	0.598	-0.036	0.510	-0.383	0.751	0.352
2011	2265	0.236	0.563	0.876	0.518	-0.202	0.706	-0.446	0.633	0.893
2012	2279	0.253	0.653	0.949	0.631	0.131	0.874	-0.961	1.100	0.607
2013	2338	0.249	0.586	0.985	0.588	0.204	0.498	-0.514	1.033	0.372
2014	2544	0.235	0.610	0.948	0.608	-0.018	0.808	-0.718	0.889	0.868
2015	2657	0.258	0.535	0.831	0.548	-0.138	0.874	-0.691	0.791	0.816
2016	2615	0.295	0.520	0.918	0.473	0.226	0.320	-0.374	1.168	0.308
Average	1703	0.254	0.638	0.832	0.631	0.006	0.569	-0.472	0.815	0.407
Pooled	45977	0.249	0.687	0.834	0.518	-0.015	0.626	-0.511	0.816	0.546

Table 9 comprises two panels, Panel A (for valuation heuristic *V1*, based on EPS1) and Panel B (for valuation heuristic *V2*, based on EPS2), and compares explanatory power in valuation space and return space (measured by static goodness-of-fit—"StatGoF"— and dynamic goodness-of-fit—"DynGoF"—respectively), of traditional OLS regressions to that of our autoregressive framework. The latter uses method M to estimate the contraction parameter *C* ("ARF(M)"). The table shows results for individual years, averages across individual years, and for the setting in which all yearly data are pooled. Dynamic goodness-of-fit is the rank correlation between realized return and implied return. The implied returns for OLS result from estimations of realized return_{t+1} on V_{t+1}/P_t and V_t/P_t, along with an intercept.

Appendix 1. PVED underpinnings of ARF

We address whether the relatively generic approach ARF connects with familiar concepts associated with valuation. In principle, ARF does not by itself imply, for example, that earnings and PVED are of any relevance. But it appeals if one can derive ARF from assumptions consistent with basics of equity valuation, such as PVED. In particular, one wants to show that $V_t = EPS1_t \times$ constant and ARF can hold simultaneously. In this context "basics" refers not only to PVED and a related discount factor, but also to realized earnings, and to a role for dividends and the Modigliani-Miller precept of dividend-policy irrelevance.

Consider the following assumptions:

(A1) P_t = PVED; discount rate $R = 1 + r$

(A2) The information dynamics

$$x_{t+1} = R \times x_t - 1/m \times d_t + v1_t + v2_t + e1_{t+1} \qquad \text{(i)}$$

$$v1_{t+1} = Q \times v1_t + e2_{t+1} \qquad \text{(ii)}$$

$$v2_{t+1} = e3_{t+1} \qquad \text{(iii)}$$

The information at date t is determined by x_t, d_t, $v1_t$, and $v2_t$.

Three parameters, $R > 1$, $m > 0$, and $0 < Q < 1$, identify the equations. $1/m$ has the same order of magnitude as r, but is typically somewhat smaller.

Given A1 and A2, it follows that

$$P_t = q_1 x_t + q_2 d_t + q_3 v1_t + q_4 v2_t \qquad (1)$$

where

$q_1 = R \times m$; $q_2 = -1$; $q_3 = (R \times m)/(R - Q)$; $q_4 = m$

Moreover,

$$P_t - EPS1_t \times m = H \times v1_t;\ H = (R \times m)/(R - Q) - m \qquad (2)$$

It follows that ARF holds by putting $Q = C$ (the contraction parameter in our framework) and assuming that the error terms in (i) and (ii) do not correlate (to meet the ARF requirement that V_{t+1} does to correlate with u_{t+1}).

Proof: Conclusion (1) follows via routine PVED calculations. Its simplicity depends on dividend policy irrelevance, which applies since the first coefficient on the RHS in A2 (i) equals R.

Regarding conclusion (2), note that $E(x_{t+1}; t) = EPS1_t$. Taking expression (1) and deducting the expression for $EPS1_t$ derived from (i) all terms cancel out, except for the one related to $v1_t$.

Remarks:

(a) the two $v(\cdot)$ terms can be thought of as generic "other information".

(b) The valuation gap is fully determined by $v1_t \times$ constant; here the constant = $(R \times m)/(R - Q)$.

(c) The parameter $1/m$ identifies the marginal return affecting $EPS1$ due to new investment (or the foregone expected earnings due to distribution of wealth).

(d) The PVED assumption can be shown to be necessary.

The model can be modified so that ARF holds when $EPS1_t$ is replaced by ($EPS2_t + 1/m \times dps1_t$, where $EPS2_t = E(x_{t+2}; t)$ and $dps1_t = E(d_{t+1}; t)$. The adjustment of $EPS2_t$ due do the dividend term makes economic sense since, in expectation, dividends forego subsequent earnings at an assumed earnings rate of $1/m$. To obtain this result, one replaces (ii) with $v1_{t+1} = Q \times v1_t + v2_t + e1_{t+1}$. $EPS1_t$ requires no dividend term adjustment.

The model anchored in PVED, A1 and A2 is a special case of Ohlson and Johannesson (2016). Specifically,

$$P_t = EPS1_t \times m + (m \times (R - G)) \times (EPS2_t + 1/m \times dps1_t - R \times EPS1_t).$$

This setting can be generalized so that the ARF holds, with V_t now replacing P_t.

Next, consider the properties of Ohlson (1995)'s model. It, too, connects with ARF, again as a special case. To develop this point, express the Ohlson model in terms of a (reduced form) valuation function and the dynamics related to so-called "other information", denoted v_t:

$$P_t = V_t + K \times v_t$$

$$v_{t+1} = g \times v_t + e_{t+1}; \qquad 0 < g < 1.$$

V_t is identified by a linear function in bv_t, earnings$_t$, and dividends$_t$ (since P equals book value of equity plus the present value of future residual income). Thus, ARF follows with $g = G$, so long as the innovation in V_{t+1} does not correlate with e_{t+1}. This mechanistic analysis, however, overlooks that V_t per the Ohlson model will generally materially underestimate P_t due to GAAP's inherent balance sheet conservatism. That is, the BV/P ratio in reality is typically far too low to suggest that the valuation gap, $K \times v_t$, hovers around zero. Accordingly, to apply a "modified Ohlson model" one has to rescale the valuation so that the resulting valuation heuristic divided by P approximates a median of one.