Journal of Accounting and Economics 69 (2020) 101281

Contents lists available at ScienceDirect

Journal of Accounting and Economics

journal homepage: www.journals.elsevier.com/journal-of-accounting-and-economics

Using a hidden Markov model to measure earnings quality☆

Kai Du [a, b, *], Steven Huddart [a], Lingzhou Xue [c], Yifan Zhang [a]

[a] Smeal College of Business, Pennsylvania State University, University Park, PA 16802, USA
[b] U.S. Securities and Exchange Commission, 100 F Street NE, Washington, DC 20549, USA
[c] Eberly College of Science, Pennsylvania State University, University Park, PA 16802, USA

ARTICLE INFO

Article history:
Received 21 April 2018
Received in revised form 4 November 2019
Accepted 19 November 2019
Available online 27 November 2019

JEL classification:
C11
C13
M41
M43

Keywords:
Hidden Markov model
Bayesian hierarchical framework
MCMC methods
Earnings quality
Earnings fidelity

ABSTRACT

We propose and validate a new measure of earnings quality based on a hidden Markov model. This measure, termed earnings fidelity, captures how faithful earnings signals are in revealing the true economic state of the firm. We estimate the measure using a Markov chain Monte Carlo procedure in a Bayesian hierarchical framework that accommodates cross-sectional heterogeneity. Earnings fidelity is positively associated with the forward earnings response coefficient. It significantly outperforms existing measures of quality in predicting two external indicators of low-quality accounting: restatements and Securities and Exchange Commission comment letters.

Published by Elsevier B.V.

1. Introduction

Earnings reports are affected by both a firm's fundamental performance and the measurement process as governed by reporting standards, auditing technology, and managerial discretion (e.g., Dechow et al., 1998; Nikolaev, 2017). Prior research has used the statistical properties of earnings (e.g., smoothness, kinks in earnings, and target beating) and regression-based abnormal accrual models to measure the quality of earnings (for a review, see Dechow et al., 2010). In regression-based models, researchers separate the “abnormal” portion of accruals from the “normal” portion related to fundamental

☆ We received many helpful comments from Wayne Guay (editor), Frank Zhou (referee), and Dan Givoly. We thank Orie Barron, Sudipta Basu, Jeremy Bertomeu, Dirk Black, Lei Dong, Guojin Gong, Daniel Jiang, Chen Li (discussant), Jia Li, Pierre Liang, Yi Liang, John Liechty, Henock Louis, Jed Neilson, Lyndon Orton, Hong Qu, Jack Stecher, Shuyang Wang, Biqin Xie, Luo Zuo, and workshop participants at Penn State University, Michigan State University, Carnegie Mellon University, and the 2018 American Accounting Association Annual Meeting for comments and suggestions. Kai Du and Steven Huddart acknowledge financial support from the Penn State Smeal College of Business. Lingzhou Xue acknowledges the financial support in part from the National Science Foundation. The Securities and Exchange Commission (SEC) disclaims responsibility for any private publication or statement of any SEC employee or Commissioner. This paper expresses the authors' views and does not necessarily reflect those of the Commission, the Commissioners, or members of the staff. An earlier version of this paper was titled "Using a hidden Markov model to measure reporting systems." Data on earnings fidelity are publicly available at https://sites.psu.edu/fidelity.

* Corresponding author. Smeal College of Business, Pennsylvania State University, University Park, PA 16802, USA.
E-mail addresses: kxd30@psu.edu (K. Du), huddart@psu.edu (S. Huddart), lzxue@psu.edu (L. Xue), yxz192@psu.edu (Y. Zhang).

https://doi.org/10.1016/j.jacceco.2019.101281
0165-4101/Published by Elsevier B.V.

performance and define accruals quality based on abnormal accruals (e.g., Dechow and Dichev, 2002). Yet recent studies point out that an important issue with this approach is that proxies for unobservable "true" earnings may confound performance shocks with reporting discretion (e.g., Guay et al., 1996; Ball, 2013; Dichev et al., 2013; Owens et al., 2017).

Our approach uses a structural model to separate accounting quality from the process of true earnings, thereby significantly alleviating the concern of having to rely on the noisy proxies. We assume that each firm transitions among states according to a Markov process. The fundamental performance of the firm is its state, which is either *Low* (*L*) or *High* (*H*). The firm's state is unobservable; however, in each period, the firm issues an earnings signal, which is either *low* (*l*) or *high* (*h*). The probability that the firm issues a particular signal depends on the unobservable state, so inferences about the firm's state can be made from its earnings signals. Therefore, each firm in each period is characterized by transition probabilities (i.e., the probabilities that the firm will remain in the present state or transition to the other state) and emission probabilities (i.e., the probabilities, conditional on the state, that the firm will emit a given earnings signal). The structure of this hidden Markov model (HMM) allows us to estimate how faithful earnings signals are in revealing the true state of the firm. We call this measure *Fidelity*.[1]

We apply the model to public firms in the U.S. with a minimum of 20 consecutive quarters of valid data for the period of 1980–2015. We reduce the earnings history of each firm to the series of signs of its quarterly earnings surprises. When earnings in the current quarter are lower than earnings in the same quarter of the prior year, the signal takes the value *l*; otherwise, the signal takes the value *h*.

The history of states, transition probabilities, and emission probabilities for each firm are estimated from a panel consisting of these firms' earnings signals and characteristics, specifically leverage, size, market-to-book ratio, and cash flow volatility. Our Bayesian hierarchical framework assumes that the transition and emission probabilities for a particular firm-period are determined by (i) the firm's characteristics in that period and (ii) firm-specific intercepts.[2] Firms with similar characteristics will have similar values of *Fidelity*. The intercepts capture firm-level departures from the general relationship. In turn, the intercepts and characteristics parameterize the Bayesian hierarchical model, which gives rise to the states, signals, and transition and emission probabilities. In the Bayesian paradigm, parameters are random variables. The parameters selected by our method are the most plausible given the data, i.e., the panel of firms' earnings signals and characteristics.

Based on parameter estimates, we compute two firm-year measures: economic persistence (i.e., the stickiness of the states of the firm) and *Fidelity* (i.e., how faithful earnings signals are in revealing the state). *Fidelity* is computed from the emission probabilities: *Fidelity* increases in both (i) the probability that the earnings signal is *l* (i.e., an earnings decrease) when the state of the firm is *L*, and (ii) the probability that the earnings signal is *h* (i.e., an earnings increase) when the state of the firm is *H*.

We take several steps to validate *Fidelity*. We first show that the observed transition rates between earnings signals are very close to the probabilities of these transitions implied by the model. This test confirms that the HMM structure applied to our data yields reasonable parameter estimates and establishes internal validity.

We then contrast the information contained in *Fidelity* with the information in three existing measures of quality, namely: (i) earnings smoothness; (ii) accruals quality (Dechow and Dichev, 2002); and (iii) unexplained audit fees (Hribar et al., 2014). *Fidelity* is negatively correlated with the notions of quality underlying earnings smoothness and accruals quality. It is positively correlated with the notion of quality captured by unexplained audit fees.

We show that high-*Fidelity* firms have a higher forward earnings response coefficient (FERC), implying that investors incorporate more information about future earnings into stock returns for these firms. Relatedly, the model implies a hump-shaped relationship between economic persistence and FERC, which we confirm in the data.

We also demonstrate that low *Fidelity* is associated with a higher incidence of restatements and Securities and Exchange Commission (SEC) comment letters. The predictive power of our measure is not subsumed by existing measures of quality. *Fidelity* possesses incremental explanatory power that is statistically significant and large in magnitude. It is notable that *Fidelity* predicts comment letters, which indicate accounting problems less severe than those leading to restatements, but accruals quality does not.

Moreover, the relationship between *Fidelity* and the incidence of restatements remains strong when observations are partitioned into firm-size quintiles. The existing measures, in contrast, do not load consistently in firm-size quintile regressions. Notably, *Fidelity* possesses strong explanatory power among small- and mid-cap firms. We suspect that this finding results from two advantages our method offers over conventional ones. First, our method does not rely on proxies for underlying constructs. Second, data on other firms contributes to the estimate of *Fidelity* for a given firm. This feature improves the accuracy of *Fidelity*, especially for firms having limited data. These advantages may be more pronounced for small firms, for which less data is available and proxies are noisier.

We conduct a battery of additional tests to establish the robustness of earnings fidelity. In our main analysis, we estimate the model and test *Fidelity* using data from the period 1980–2015. To remove any potential look-ahead bias, we calculate an alternative fidelity measure for the years of 2012–2015 based on the coefficients estimated for the sub-period before 2012.

[1] There are decision contexts in which the change in accounting earnings, by itself, is not the main informational input (e.g., debt contracts based on non-GAAP performance metrics). In those contexts, *Fidelity* may not be a good measure of quality.

[2] The method is hierarchical because some parameters govern other parameters. In particular, the mean and variance of the common distribution from which the intercepts are drawn govern those intercepts.

The new measure, dubbed *Fidelity*^{p}, possesses the same predictive ability. We also define an alternative measure of earnings fidelity—weighted fidelity, *Fidelity*^{w}—that combines information from the reporting system with information about the likelihood of actual states. This measure exhibits similar properties to the basic fidelity measure. In addition, we illustrate that the measures are not driven by particular covariates and confirm that using matched samples for restatement predictive tests yields qualitatively the same results.

A consensus regarding the best measure for earnings quality has yet to emerge (Dechow et al., 2010). The Bayesian hierarchical method is new to empirical research that proposes measures of accounting quality based on distributional properties of earnings (e.g., smoothness) and reduced-form regression models (e.g., accruals quality).[3] Our structurally-estimated measure of earnings quality is theoretically motivated and offers at least two advantages over existing quality measures. First, the Bayesian hierarchical method treats firm-specific parameters as random variables drawn from a distribution. As a consequence, the method combines firm-specific data with important prior information (i.e., the distribution of all firms' parameters). Second, connections among firms are unconstrained under this method. For example, firms in the same industry or reports in the same year need not be connected in any specific way. This assumption has the potential to reduce measurement error.

A divergence exists between theories of and empirical research on accounting quality. Whereas the analytical literature has focused on static models of financial reporting (see Christensen and Feltham, 2003), the empirical literature has proposed measures of quality based on accounting numbers from multiple periods. The hidden Markov model allows us to extend the static reporting systems to one that is dynamic.[4] By estimating properties of the reporting system in a multi-period setting, we bring theories closer to empirical research on earnings quality.

Bayesian hierarchical modeling is a way to account for cross-firm heterogeneity, which is not addressed in most structural estimations in accounting (e.g., Bertomeu et al., 2018; Zakolyukina, 2018; Du, 2019). Bayesian hierarchical modeling and the estimation method, Markov chain Monte Carlo (MCMC), are standard in many academic disciplines, but have been applied only recently in accounting research: Bernhardt et al. (2016) estimate the stickiness of analyst recommendations in a cross-section of analysts, while Zhou, 2017 estimates the impact of investor learning on cross-sectional variations in the manager's voluntary disclosure. Our study is the first to use the Bayesian approach to estimate firm- and time-varying earnings quality.

In Section 2, we define earnings fidelity in a hidden Markov model and introduce the estimation method. Section 3 describes the data and estimation results. In Section 4, we conduct empirical tests to validate the measure. Section 5 concludes.

2. The model and estimation method

2.1. Earnings fidelity in a hidden Markov model

We conceptualize our measure of earnings quality—earnings fidelity—as the extent to which an earnings signal reveals the underlying state of the firm, where the state refers to the economic conditions that affect the fundamental performance of the firm. An earnings signal with greater fidelity enables the recipients to be more confident in their inference about the state of the firm. To formalize earnings fidelity, we need a theoretical model in which earnings provide a noisy signal of the underlying state of the firm. We adopt a highly stylized model that serves this purpose.

Suppose there are a large number of idiosyncratic firms indexed by i. Firm i's underlying state at time t, x_{it}, can take on one of two values: L or H. The initial state probabilities are $\boldsymbol{\pi}_0 = (\pi_{0L}, \pi_{0H})$,[5] where

$$\pi_{0s} = \Pr(x_{i1} = s), \text{ for } s = L, H. \tag{1}$$

The law of motion of x_{it} is described by a Markov transition matrix influenced by time-varying firm-specific characteristics:

$$\boldsymbol{Q}_{it} = \begin{bmatrix} \Pr(x_{it} = L | x_{it-1} = L) & \Pr(x_{it} = H | x_{it-1} = L) \\ \Pr(x_{it} = L | x_{it-1} = H) & \Pr(x_{it} = H | x_{it-1} = H) \end{bmatrix} = \begin{bmatrix} a_{it} & 1 - a_{it} \\ 1 - b_{it} & b_{it} \end{bmatrix}, \tag{2}$$

where

[3] Dechow et al. (2012) show that exploiting the inherent property of accrual accounting can improve the power and specification of tests for earnings management. This finding is in line with the notion that incorporating important prior information can improve statistical inferences (Berger, 1985). However, Dechow et al. do not use Bayesian hierarchical modeling.

[4] An HMM is a model of a stochastic process that can only be observed through a noisy signal released each period. HMMs have been applied to many fields of scientific inquiry, including speech processing (Rabiner and Juang, 1986) and biological sequence analysis (Durbin et al., 1998). In the social sciences, HMMs have been used to study non-stationary time series in finance and economics (e.g., Hamilton, 1989; Gray, 1996) and consumer behavior in marketing research (e.g., Netzer et al., 2008). For an overview of the applications of HMM, see MacDonald and Zucchini (1997).

[5] We have implicitly assumed that the probability of the initial state is the same for every firm. We examine the impact of this assumption on model estimates in the simulation study (see Appendix B) and find this assumption has no significant impact.

$$a_{it} = 1 - \Phi(\alpha^*_{iL} + \boldsymbol{z}'_{it}\boldsymbol{\alpha}_L), \quad b_{it} = \Phi(\alpha^*_{iH} + \boldsymbol{z}'_{it}\boldsymbol{\alpha}_H). \tag{3}$$

$a_{it}(b_{it})$ is the conditional probability of firm i remaining in state $L\,(H)$ at time t, given it is in state $L\,(H)$ at time $t-1$; $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution; α^*_{is} is the firm-level intercept that captures firm i's tendency to transition to state H from state $s \in \{L,H\}$; $\boldsymbol{z}_{it}$ is the vector of the time-varying characteristics (known as covariates) for firm i at time t that influences the transitions of the latent states; $\boldsymbol{\alpha}_s$ captures the impact of the covariates on the state transition in state s. The assumption about the law of motion of the true earnings rests on the observation that a firm's true profitability exhibits both persistence and mean-reversion (e.g., Fairfield et al., 2009; Fama and French, 2000).

The underlying state is not directly observed by the public, an assumption also adopted by extant accounting theories (e.g., Dye, 1985; Verrecchia, 2001; Gao, 2013). Given the underlying state, x_{it}, firm i in period t issues an earnings signal y_{it} that can take on one of two values: l or h. The earnings signal is a noisy indicator of the firm's underlying state.

The probability of issuing a particular earnings signal depends on the current unobservable state of the firm. The mapping from the state to the earnings signal can be described by the matrix of state-contingent probabilities (i.e., "emission probabilities" in the HMM terminology) $\boldsymbol{\eta}_{it}$,

$$\boldsymbol{\eta}_{it} = \begin{bmatrix} \Pr(y_{it} = l|x_{it} = L) & \Pr(y_{it} = h|x_{it} = L) \\ \Pr(y_{it} = l|x_{it} = H) & \Pr(y_{it} = h|x_{it} = H) \end{bmatrix} = \begin{bmatrix} c_{it} & 1 - c_{it} \\ 1 - d_{it} & d_{it} \end{bmatrix}, \tag{4}$$

where

$$c_{it} = 1 - \Phi(\beta^*_{iL} + \boldsymbol{u}'_{it}\boldsymbol{\beta}_L), \quad d_{it} = \Phi(\beta^*_{iH} + \boldsymbol{u}'_{it}\boldsymbol{\beta}_H). \tag{5}$$

$c_{it}(d_{it})$ is the conditional probability that firm i issues an earnings signal $l\,(h)$ at time t, given that it is currently in state L (H). β^*_{is} is the firm-specific intercept that captures firm i's idiosyncracy in earnings signals when the state is $s \in \{L,H\}$; $\boldsymbol{u}_{it}$ is the vector of the time-varying coefficients, and $\boldsymbol{\beta}_s$ captures the impact of these covariates on the signal y_{it} in state s. To ensure the identification of the states, we require $\beta^*_{iL} < \beta^*_{iH}$, such that the probability of generating signal h is higher when the firm's underlying state is H and all covariates equal to zero. The structure of the model is illustrated in Fig. 1, Panel A.

Covariates $\boldsymbol{z}_{it}$ and $\boldsymbol{u}_{it}$ represent firm characteristics that drive transitions between states and determine the faithfulness of the earnings signals. The covariates help us identify how the transition and emission probabilities change over time.

In both transition and emission probabilities, we have used firm-level intercepts (i.e., $\boldsymbol{\alpha}^*_i = (\alpha^*_{iL}, \alpha^*_{iH})$ and $\boldsymbol{\beta}^*_i = (\beta^*_{iL}, \beta^*_{iH})$, respectively) to capture firm idiosyncrasy. Ideally, each firm-level intercept would be treated as a free parameter to be estimated from the data; however, this approach would require a larger number of observations for each firm than is available and so proves infeasible. Instead, we use Bayesian hierarchical modeling and assume that firm-level parameters α^*_{is} and β^*_{is} follow a common distribution across firms. Specifically, we assume α^*_{is} follows a normal distribution with mean $\overline{\alpha}_s$ and variance σ^2_s:

$$\alpha^*_{is} \sim N\left(\overline{\alpha}_s, \sigma^2_s\right), \quad s = L, H, \tag{6}$$

where $\overline{\alpha}_s$ is the mean of intercepts in the transition model across all firms, and σ^2_s captures the heterogeneity among firms. Let $\overline{\boldsymbol{\alpha}}_0 = (\overline{\alpha}_L, \overline{\alpha}_H)$ and $\boldsymbol{\sigma}^2_\alpha = (\sigma^2_L, \sigma^2_H)$. When $\boldsymbol{\sigma}^2_\alpha$ is large, firm-level intercepts are more likely to deviate from the mean; when $\boldsymbol{\sigma}^2_\alpha$ is small, firm-level intercepts cluster around the mean. We adopt a similar modeling strategy for β^*_{is} with one modification. To enforce the constraint $\beta^*_{iL} < \beta^*_{iH}$ and facilitate parameter estimation, we introduce an additional variable δ_i such that

$$\begin{aligned} &\beta^*_{iH} = \beta^*_{iL} + \exp(\delta_i), \\ &\text{where} \quad \beta^*_{iL} \sim N\left(\overline{\beta}_L, \omega^2_L\right), \quad \delta_i \sim N\left(\overline{\delta}_H, \omega^2_H\right). \end{aligned} \tag{7}$$

Similarly, $\overline{\boldsymbol{\beta}}_0 = (\overline{\beta}_L, \overline{\delta}_H)$ determines the mean level of intercepts in the emission model across all firms, and $\boldsymbol{\omega}^2_\beta = (\omega^2_L, \omega^2_H)$ captures the strength of heterogeneity.

There are two distinct constructs featured in our model: economic persistence (i.e., the stickiness of the economic states of the firm) and earnings fidelity (i.e., the faithfulness of earnings signals as indicators of the underlying state of the firm). We operationalize economic persistence (*Pers*) as the average of the state-dependent transition probabilities, a_{it} and b_{it}:[6]

$$Pers_{it} = \frac{a_{it} + b_{it}}{2}. \tag{8}$$

We operationalize earnings fidelity (*Fidelity*) as the average of the conditional probabilities, c_{it} and d_{it}:

[6] It is important to note that economic persistence (*Pers*) differs from "earnings persistence" used as a proxy of earnings quality in the literature (Dechow et al., 2010). Earnings persistence combines economic persistence and earnings fidelity. Therefore, *Pers* is related to but distinct from earnings persistence.

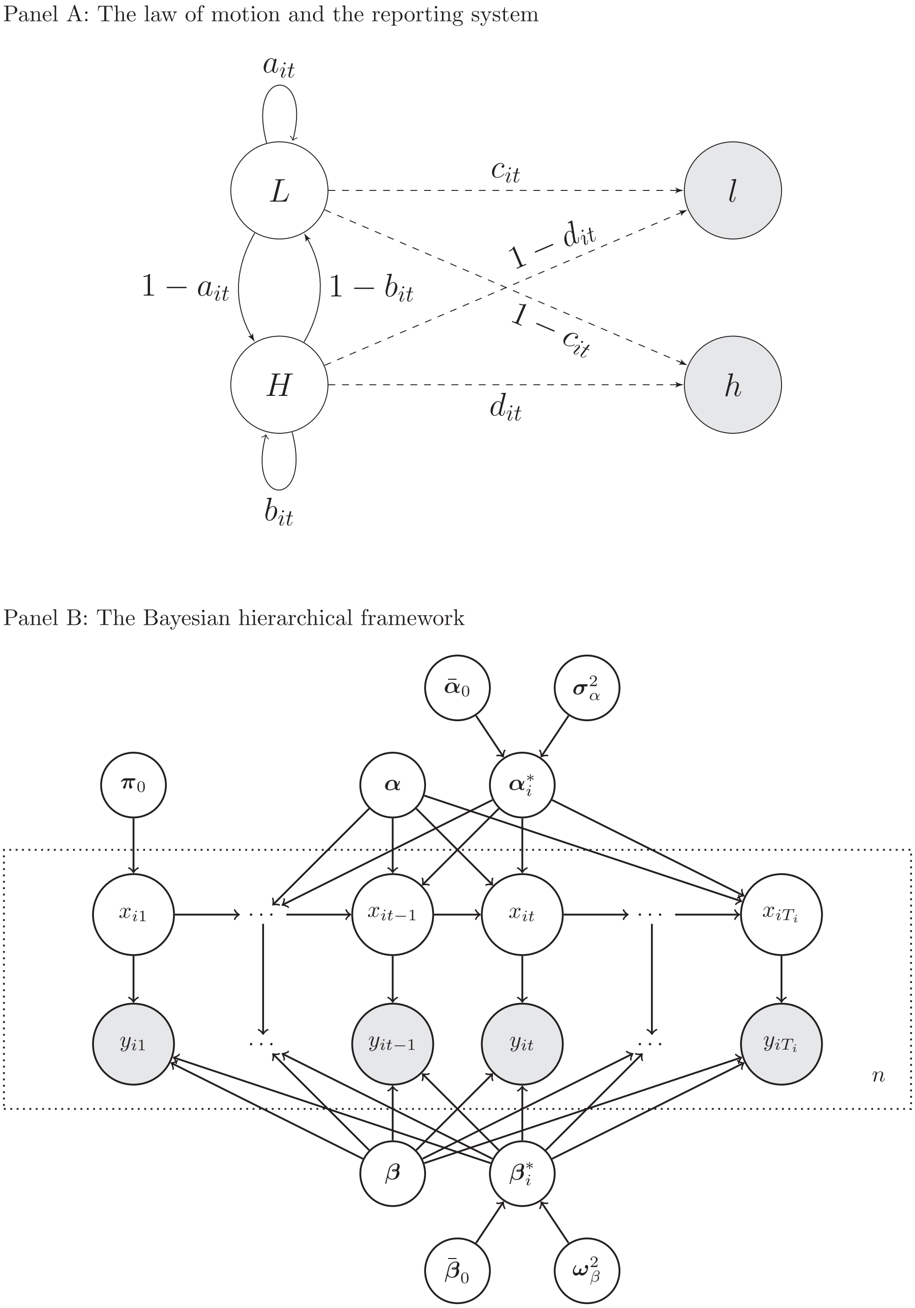

Fig. 1. Illustration of the model and estimation method. Panel A illustrates the transition of the latent state $x_{it} \in \{L, H\}$ and the mapping from the state to an earnings signal $y_{it} \in \{l, h\}$ in the hidden Markov model. Panel B illustrates how the model parameters influence the states and signals in the Bayesian hierarchical framework. Each firm i has T_i periods. There are n firms in the sample.

$$Fidelity_{it} = \frac{c_{it} + d_{it}}{2}. \tag{9}$$

In other words, *Fidelity* is the average of the probability of issuing signal l conditional on state L and the probability of issuing signal h conditional on state H.[7] It is intuitive that *Fidelity* is greater when either c_{it} or d_{it} is greater. *Fidelity* is affected

[7] Earnings fidelity, as a one-dimensional measure of the faithfulness of earnings signals, simplifies the informativeness criterion studied in the statistical decision making literature. Strictly speaking, if we view an earnings signal as a statistical experiment, the “informativeness” of a signal represents a partial ordering of all information systems (Blackwell, 1951). Although this partial ordering cannot typically be represented by a scalar, reducing the dimensionality of the reporting system facilitates empirical analysis and enhances comparability with prior research.

by a range of factors, including accounting rules, economic uncertainty, and the reporting decisions of managers and accountants. First, there is intrinsic randomness in the accounting rules. Second, low *Fidelity* may be an unavoidable consequence of management not having precise knowledge of the state of the firm. Third, managers may intentionally manipulate earnings to meet expectations. We do not model any of these factors explicitly. Instead, we estimate their combined effect on the quality of the earnings signals.

Fidelity defined in Eq. (9) is the simple average of the two conditional probabilities, without considering the probability of each state. We also define an alternative measure of earnings fidelity, based on a notion of "unconditional correctness" (i.e., how likely the earnings signal is a faithful representation of the underlying state, regardless of which state the firm is in).[8] The weighted fidelity measure, or $Fidelity^{w}$, for firm i at time t is

$$Fidelity_{it}^{w} = p_{itL}c_{it} + p_{itH}d_{it}, \tag{10}$$

where p_{its} is the probability that firm i's state at t is $s \in \{L, H\}$ after observing the earnings history.

2.2. Overview of the model estimation and parameter identification

In this section, we outline the estimation procedure and the sources of parameter identification. Appendix A provides detailed information about our estimation algorithm.

2.2.1. Bayesian estimation and the MCMC algorithm

Let $\boldsymbol{\theta}$ be the vector of all unknown model parameters. Denote by $L(\boldsymbol{\theta}|\boldsymbol{y})$ the joint likelihood of the observed signals and $\pi(\boldsymbol{\theta})$ the prior distribution of the unknown parameters. Following MacDonald and Zucchini (1997), the posterior distribution of $\boldsymbol{\theta}$ given the observed data, written in compact matrix form, is given by

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) \propto L(\boldsymbol{\theta}|\boldsymbol{y})\pi(\boldsymbol{\theta}) \propto \left[\prod_{i=1}^{n} \boldsymbol{\pi}_0 \tilde{\boldsymbol{\eta}}_{i1}\mathbf{Q}_{i2}\tilde{\boldsymbol{\eta}}_{i2}\cdots\mathbf{Q}_{iT_i}\tilde{\boldsymbol{\eta}}_{iT_i}\mathbf{1}'\right] \cdot \pi(\boldsymbol{\theta}), \tag{11}$$

where $\mathbf{Q}_{it}$ is the Markov transition matrix for x_{it} defined in Eq. (2), and $\tilde{\boldsymbol{\eta}}_{it}$ is the 2×2 diagonal matrix:

$$\tilde{\boldsymbol{\eta}}_{it} = \begin{bmatrix} c_{it} & 0 \\ 0 & 1-d_{it} \end{bmatrix} \text{if } y_{it} = l, \quad \tilde{\boldsymbol{\eta}}_{it} = \begin{bmatrix} 1-c_{it} & 0 \\ 0 & d_{it} \end{bmatrix} \text{if } y_{it} = h, \tag{12}$$

for $t = 1, 2, \ldots, T_i$. Intuitively, the joint likelihood $L(\boldsymbol{\theta}|\boldsymbol{y})$ is equal to the probability that we observe the panel of signals $\boldsymbol{y}$ conditional on the parameters $\boldsymbol{\theta}$. To reduce the impact of prior distributions $\pi(\boldsymbol{\theta})$ on the estimation results, we use uninformative prior distributions.[9]

From a Bayesian perspective, the goal of our analysis is first to learn about the posterior distribution $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ and then make inferences. Unfortunately, for many Bayesian hierarchical models, including ours, the posterior distribution of the parameters is not analytically tractable. Such simulation-based methods as the MCMC algorithm are commonly used to draw samples of parameters that approximate the posterior distribution.

The main idea of the MCMC method is to construct a Markov chain with the desired posterior distribution $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ as its equilibrium distribution. In each iteration of the MCMC method, we generate samples of the parameters based on their values from the previous iteration and their posterior distributions. After many iterations, the generated samples have converged to the equilibrium distribution. We use these samples to approximate the posterior distributions and estimate the mean value, standard deviation, and the 95% credible interval of the parameters.[10] For example, we calculate the estimated mean values of the parameters by averaging the generated samples. Although the draws from a Markov chain are not independent and identically distributed (i.i.d.), the law of large numbers applies.[11]

To illustrate this process, consider an example model wherein the parameters can be grouped into two blocks: $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$.[12] The MCMC sampler starts with an initial value of $\boldsymbol{\theta}^{(0)} = (\boldsymbol{\theta}_1^{(0)}, \boldsymbol{\theta}_2^{(0)})$. In the first iteration, we sample $\boldsymbol{\theta}_1^{(1)}$ from the

[8] This notion is the dynamic counterpart of the "overall probability of correctness" in the one-period model of accounting information systems proposed by Antle and Lambert (1988). They define overall correctness as the weighted average of c and d (using our notation), where the weights are the *a priori* probability of the states.

[9] For instance, we assume the prior distribution for the mean of firm-level intercept $\overline{\alpha}_s$ to be $N(0, 10)$. The large variance indicates that we have almost no prior knowledge of the value of $\overline{\alpha}_s$ and the posterior distribution of $\overline{\alpha}_s$ will be determined mostly by the data. More details on the prior distributions are provided in Appendix A.

[10] The 95% credible interval is the Bayesian counterpart of the 95% confidence interval.

[11] The computation of various integrals of functions such as posterior means, posterior standard deviation, and predictive distributions poses computational challenges. A natural solution is to make i.i.d. draws from the posterior distributions and then approximate the integrals by the sample mean. Drawing from a multivariate posterior distribution, however, often involves a high-dimensional integration, making it computationally infeasible. Instead of using i.i.d. draws, researchers have constructed the Markov chain with the posterior distribution as its stationary or equilibrium distribution, and then drawn samples from the chain to approximate the integrals. See Robert and Casella (2013) for more discussion.

[12] In our study, there are more than two blocks of parameters. See Appendix A for more details.

conditional distribution $\pi(\boldsymbol{\theta}_1|\boldsymbol{y},\theta_2^{(0)})$. Here, $\pi(\boldsymbol{\theta}_1|\boldsymbol{y},\boldsymbol{\theta}_2^{(0)})$ is the conditional distribution of $\boldsymbol{\theta_1}$ given the current value of $\boldsymbol{\theta_2}$ and the data. In the Bayesian literature, this conditional distribution is often referred to as the full conditional distribution. Once we update the value for $\boldsymbol{\theta}_1$, we update $\boldsymbol{\theta}_2$ by sampling a value from the full conditional distribution $\pi(\boldsymbol{\theta}_2|\boldsymbol{y},\boldsymbol{\theta}_1^{(1)})$. Notice the full conditional distribution uses the updated value of $\boldsymbol{\theta}_1$.

The MCMC sampler continues by iteratively taking draws of $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ from their full conditional distributions until a pre-specified number of draws is collected.[13] For models with more than two blocks of parameters, the MCMC sampler works similarly.

After the MCMC sampler is stopped, we discard the first M draws of the parameters and use the subsequent draws to estimate the parameters of interest. The practice of discarding an initial portion of a Markov chain sample, known as "burn-in," ensures that the effect of initial values on the posterior inference is minimized. In practice, the value of M is chosen such that the MCMC sampler converges to the posterior distribution after M iterations. In our empirical analysis, we run the MCMC sampler for 300,000 iterations. Then, we discard the initial $M = 250,000$ draws and use the subsequent 50,000 draws to estimate the parameters. We also check the trace plots of the parameters to ensure our MCMC sampler has converged to the posterior distribution after the burn-in period. Note that the convergence of hyperparameters such as $\overline{\alpha}_0$ and $\overline{\beta}_0$ often depends on the convergence of firm-level parameters. As a result, the convergence of hyperparameters in MCMC sampling may often take longer than other parameters.

Implementation of the MCMC sampler only requires the full conditional distribution of the model parameters, which is often much easier to derive than the posterior distribution. Thus, the MCMC method is particularly useful for complex latent state models like ours. Detailed derivations of the full conditional distributions of the model parameters are provided in Appendix A.

2.2.2. Parameter identification

Our model consists of the following types of parameters to be estimated: (i) the initial state distribution $\boldsymbol{\pi}_0$; (ii) coefficients $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ on the covariates in the transition and emission matrices; (iii) firm-level intercepts $\boldsymbol{\alpha}_i^*$ and $\boldsymbol{\beta}_i^*$; (iv) hyperparameters $\overline{\boldsymbol{\alpha}}_0$ and $\overline{\boldsymbol{\beta}}_0$ that govern the mean of firm-level intercept distribution; and (v) hyperparameters $\boldsymbol{\sigma}_\alpha^2$ and ω_β^2 that govern the variance of the distribution of firm-level intercepts. In addition, we also need to infer the underlying states, $\boldsymbol{x}$. The relationship between the unknown parameters and the observed signals $\boldsymbol{y}$ is summarized in Fig. 1, Panel B.

Identification of $\boldsymbol{\pi}_0$ primarily comes from the initial state x_{i1} for each firm, which corresponds to the first quarter of each firm's data. Note that $\boldsymbol{\pi}_0$ only influences the distribution of the underlying state for the first observed signal. Identification of $\boldsymbol{\alpha}$ comes from variations in the covariates $\boldsymbol{z}_{it}$ and transitions of the underlying states across all firms. For example, consider a particular covariate: firm size. If larger firms are more likely to stay in state H compared to smaller firms, the coefficient estimate for size in state H will be positive; if firms of different sizes have the same tendency to stay in state H, then the coefficient estimate for size in state H will be close to zero. Similarly, identification of $\boldsymbol{\beta}$ comes mostly from variations in the covariates $\boldsymbol{u}_{it}$ and the mapping between the underlying state and observed signal.

Identification of the firm-level parameter $\boldsymbol{\alpha}_i^*$ comes from firm i's underlying state transitions. If firm i has a higher tendency to transition to state H than an average firm in the sample, α_i^* will be larger than $\overline{\alpha}_0$; if firm i has a lower tendency to transition to state H than an average firm, $\boldsymbol{\alpha}_i^*$ will be smaller than $\overline{\alpha}_0$. Similarly, identification of β_i^* derives from the mapping between firm i's underlying states and observed signals. If firm i is more likely to produce signal h than an average firm, then $\boldsymbol{\beta}_i^*$ will be larger than $\overline{\boldsymbol{\beta}}_0$. The hyperparameters are determined by the cross-sectional distribution of firm-level parameters: $\overline{\boldsymbol{\alpha}}_0$ and $\overline{\boldsymbol{\beta}}_0$ ($\boldsymbol{\sigma}_\alpha^2$ and ω_β^2) are the mean (variance) parameters of $\boldsymbol{\alpha}_i^*$ and $\boldsymbol{\beta}_i^*$.

Note that unlike a regression-based approach, the parameters cannot be estimated separately. We jointly estimate all unknown parameters using the MCMC method outlined in Appendix A. In Appendix B, we conduct a simulation analysis to confirm that the proposed MCMC method can recover the true model parameters.

2.2.3. Benefits and costs of the method

Two main elements define our method: the hidden Markov model and the Bayesian hierarchical framework. In our view, both likely improve the measurement of earnings quality. The hidden Markov model does not require the observability of proxies for underlying states. As a result, our method of decoding earnings signals does not rely on empirical proxies of the underlying states; therefore, it is less susceptible to misspecification and variable omission.

To understand the comparative advantages of the Bayesian hierarchical framework, consider two common approaches used in the accounting literature to develop firm-level empirical measures. First, a researcher may treat unknown parameters as the same across firms in the same group and estimate the parameters using cross-sectional regressions or a panel regression (e.g., Dechow and Dichev, 2002; Hribar et al., 2014).[14] This approach may be inappropriate when parameters vary

[13] For our model estimation, when the full conditional distribution of a parameter is a standard probability distribution (e.g., normal distribution), we update the parameter by directly sampling from that distribution. Such samplers are known as Gibbs samplers. When the full conditional distribution of a parameter is not a standard probability distribution, we update the parameter by using the Metropolis-Hasting algorithm (Hastings, 1970).

[14] Dechow and Dichev (2002) conduct a cross-sectional regression for each industry-year to obtain firm-year measures of abnormal accruals. In other words, firms in the same industry are required to have the same coefficients. Hribar et al. (2014) estimate unexplained audit fees from a cross-sectional regression for each size decile in each fiscal year.

cross-sectionally. Second, a researcher may estimate parameters for each firm using firm-specific time-series regressions (e.g., Lipe et al., 1998).[15] Firm-specific analysis relaxes the assumption that the model parameters are the same for all firms, but it requires a long sequence of observations to reliably estimate firm-level parameters, which is often infeasible. Even when the time series is sufficiently long, it may not be stationary as assumed. A better approach would "shrink" firm-specific estimates towards population means when firm-specific data is scarce.[16]

The Bayesian hierarchical framework assumes that firm-specific parameters are drawn from a distribution.[17] Firm-specific estimates are based on a mixture of the data available for that firm and the population distribution. The relative weight placed on the former is a function of the amount of firm-specific information. When the available time series for a given firm is long, these data are the main determinants of the firm-specific estimates. When the time series is limited, the estimates shrink towards the prior. Studies have found that the accuracy gain of Bayesian hierarchical estimates over conventional estimates is usually large (e.g., Greenland, 2000). Thus, the Bayesian hierarchical framework can significantly improve firm-specific estimates for firms whose time series are limited. We refer readers to the Bayesian analysis literature prior research For more rigorous discussions of the technical advantages of the Bayesian hierarchical framework (Gelman, 2006; Gelman et al., 2013).

The validity of our measure rests on the assumed hidden Markov structure of the financial reporting process. Measurement error could arise if this structure is not descriptive; however, there is no reason to believe that this structure is less plausible than the ones underlying reduced-form accounting quality measures. For instance, some measures based on reduced-form relationships are derived, implicitly, from models that suppose accounting signals are a linear function of proxies for underlying states even though there is no *a priori* reason to expect a linear relationship.

There are three potential costs to the Bayesian approach. First, even though estimates from a Bayesian hierarchical model often have a lower variance relative to firm-specific estimates, they may have a greater bias (Greenland, 2000).[18] Second, any misspecification in the prior distributions of unknown parameters may lead to biased parameter estimates.[19] The impact of priors can be nonnegligible when the data-based information to "parameters" ratio is low (Rossi et al., 2005). Lastly, the estimation algorithm with MCMC simulation is often computationally challenging. This is especially true for firm-level parameters for which implementing the Metropolis-Hasting algorithm involves calculating the likelihood function at each iteration. Despite these potential costs, we believe the Bayesian hierarchical model can be a useful tool for accounting researchers interested in estimating firm-level measures.

3. Data, variables, and estimation results

3.1. Data and variable measurement

We obtain data on quarterly earnings from the Compustat quarterly database and other financial data from the Compustat annual database. Starting with all firm-quarters from January 1980 to December 2015, we exclude firms with a negative book value of equity and firms with fewer than 20 consecutive quarters of valid data.

For the choice of the binary earnings signal y_{it}, we use an important metric from quarterly financial statements: whether earnings are above or below the earnings expectation (e.g., Dichev et al., 2013). We operationalize y_{it} as an indicator that takes the value h if firm i's earnings surprise in quarter t is positive, and l otherwise. Earnings surprise, the basis for the signal y_t, is measured based on a seasonal random walk model (i.e., the expectation equals the earnings from the same quarter of the previous year). This measure of earnings surprise does not rely on data external to the firm's financial reporting process (e.g., analyst forecasts) and allows for a large sample.[20]

For the covariates, we use size (*Size*), leverage (*Lev*), market-to-book ratio (*MB*), and cash flow volatility (*Cvol*). *Size* is the natural log of the market value of equity at year end. *Lev* is total debt divided by total assets at year end. *MB* is market value of equity divided by book value of equity at year end. *Cvol* is the standard deviation of operating cash flow from operations scaled by lagged total assets over the five-year window ending with the current year, with a minimum of three years. The first three variables have been used in prior studies as determinants of reporting systems (e.g., Khan and Watts, 2009). *Cvol* captures the variability of the firm's economic operations. The four covariates cover the basic attributes of the financial

[15] Lipe et al. (1998) estimate firm-specific earnings-response coefficients based on firm-specific time-series regressions.

[16] The term "shrinkage" means averaging with the population distribution (Greenland, 2000).

[17] This assumption is in line with two observations. First, as noted above, there may be substantial cross-sectional heterogeneity in firm-level parameters. Second, an empirical researcher may face parameter uncertainty when estimating models that describe financial reporting practices (e.g., Ball, 2013).

[18] Estimation inaccuracy is typically measured by mean squared error, which is the sum of variance and squared bias. The variance of an estimator measures the scatter around its expectation value across random samples. Statistical bias is the difference between the expectation of the estimator and the true parameter value. See Greenland (2000) for a discussion of the tradeoff between variance and statistical bias.

[19] For example, the normal distribution is commonly used to model firm-level parameters since it usually leads to the closed-form posterior probabilities and efficient computation. However, the normal distribution may introduce bias when the distribution of firm-level parameters is not bell-shaped (e.g., has heavy tails or is bi-modal).

[20] Using the most recent consensus analyst forecasts to calculate earnings surprises would substantially limit the sample of firms.

reporting environment, including the stability of operations, debt-contracting considerations, growth options, and external monitoring.

3.2. Estimation results

The final sample for the estimation consists of 488,533 firm-quarters, representing 9,407 unique firms. The average (median) firm has 51.93 (42) valid quarters. Table 1, Panel A presents the descriptive statistics for the variables used in the estimation. For 54.2% of the firm-quarters, y_{it} takes on a value of h.

Table 1, Panel B summarizes the estimates of the coefficients. For transition probabilities, a positive coefficient on a covariate means that firms with a higher value of the covariate are more likely to move to state H in the next period. There is a positive association between firm size and economic persistence, consistent with prior findings (e.g., Lev, 1983). The relationship between the other covariates (i.e., *Lev*, *MB*, and *Cvol*) and economic persistence is contingent on the state. For example, firms in state H with higher values of *MB* or *Cvol* are less likely to stay in H; however, firms in state L with higher values of *MB* are more likely to stay in L.

For the emission probabilities underlying the reporting system, a positive coefficient on a covariate means that firms with a higher value of the covariate are more likely to generate signal h, regardless of the underlying state. Firm size is negatively associated with the faithfulness of signals emanating from either state. The relationships between the other three covariates and earnings faithfulness are sensitive to the state. *MB* and *Cvol* are positively associated with the faithfulness of signals from state H, but negatively associated with the faithfulness of signals from state L. *Lev* makes signals from state H less faithful, but does not affect the faithfulness of signals from state L. Although some of the findings (e.g., the role of size) are in line with prior studies that employ proxies for the latent state (e.g., Khan and Watts, 2009), others are not. The discrepancies may be due to

Table 1
MCMC estimation.

Panel A: Descriptive statistics of input data for MCMC estimation

Variable	Mean	SD	25 pctl	Median	75 pctl
$1\{y_{it} = h\}$	0.542	0.498	0	1	1
$Size_{it}$	5.296	2.293	3.646	5.228	6.895
Lev_{it}	0.763	1.647	0.040	0.260	0.784
MB_{it}	2.537	3.877	1.029	1.628	2.717
$Cvol_{it}$	0.089	0.093	0.034	0.060	0.108

Panel B: Estimated coefficients on covariates

Coefficient	Mean	SD	95% credible interval
Transition matrix:			
$\widehat{\alpha}_L(Size)$	−0.032*	0.004	(−0.039, −0.026)
$\widehat{\alpha}_H(Size)$	0.024*	0.003	(0.019, 0.029)
$\widehat{\alpha}_L(Lev)$	0.029*	0.003	(0.022, 0.036)
$\widehat{\alpha}_H(Lev)$	0.023*	0.003	(0.017, 0.030)
$\widehat{\alpha}_L(MB)$	−0.005*	0.001	(−0.007, −0.002)
$\widehat{\alpha}_H(MB)$	−0.007*	0.001	(−0.009, −0.004)
$\widehat{\alpha}_L(Cvol)$	−0.019	0.073	(−0.160, 0.128)
$\widehat{\alpha}_H(Cvol)$	−0.242*	0.054	(−0.347, −0.137)
Reporting system:			
$\widehat{\beta}_L(Size)$	0.032*	0.007	(0.019, 0.046)
$\widehat{\beta}_H(Size)$	−0.083*	0.009	(−0.100, −0.065)
$\widehat{\beta}_L(Lev)$	−0.011	0.008	(−0.027, 0.004)
$\widehat{\beta}_H(Lev)$	−0.032*	0.004	(−0.040, −0.024)
$\widehat{\beta}_L(MB)$	0.011*	0.002	(0.007, 0.014)
$\widehat{\beta}_H(MB)$	0.263*	0.015	(0.233, 0.293)
$\widehat{\beta}_L(Cvol)$	0.465*	0.136	(0.181, 0.710)
$\widehat{\beta}_H(Cvol)$	0.409*	0.190	(0.039, 0.791)

Panel C: Estimated values of main parameters

Parameters	Mean	SD	2.5 pctl	97.5 pctl
$\widehat{a}_{it}$	0.726	0.033	0.661	0.780
$\widehat{b}_{it}$	0.761	0.027	0.709	0.807
$\widehat{c}_{it}$	0.902	0.019	0.864	0.929
$\widehat{d}_{it}$	0.941	0.047	0.822	1.000

Panel A presents the descriptive statistics of the inputs for the MCMC estimation. $1\{y_{it} = h\}$ is an indicator variable that is equal to 1 if $y_{it} = h$, i.e., firm i's earnings surprise in quarter t is positive, and 0 otherwise. Firms are required to have positive common equity and at least 20 consecutive quarters of valid data. Panel B describes the estimated coefficients on the covariates. Panel C describes the main parameters that define the transition matrix and the emission matrix (reporting system). The estimation sample consists of 488,533 firm-quarters for January 1980 through December 2015. * indicates that the 95% credible interval of the coefficient does not contain 0.

the noise in the proxies or the methodological differences between the regression approach used in prior studies and our structural approach.

After estimating the coefficients for each firm i, we can derive the main parameters of the model. Table 1, Panel C shows the summary statistics for the parameter estimates $\widehat{a}_{it}$, $\widehat{b}_{it}$, $\widehat{c}_{it}$, and $\widehat{d}_{it}$, along with the measures of economic persistence and earnings fidelity. On average, state L is less persistent than state H: The mean of $\widehat{a}_{it}$ is 0.726, and the mean of $\widehat{b}_{it}$ is 0.761. This is consistent with the notion that shareholders have a liquidation option and losses are not expected to perpetuate (Hayn, 1995). The average probability of issuing signal l given state L, or $\widehat{c}_{it}$, is 0.902. The average probability of issuing signal h given state H, or $\widehat{d}_{it}$, is 0.941. Therefore, the signal y_{it} correctly reflects the latent economic state x_{it} in more than 90% of the cases, irrespective of the state.

We then use Eqs. (8) and (9) to compute *Pers* and *Fidelity* for each firm-quarter. The formula for $Fidelity^W$ given by Eq. (10) also requires p_{its}, which is approximated by the proportion of MCMC draws of x_{it} in state s for a given pair of i and t. To be consistent with prior measures of quality, which are measured at the yearly frequency (e.g., accruals quality), we convert the quarterly measures to annual measures by taking the annual average of the respective quarterly measures.

To understand the determinants of *Fidelity*, note that it is a combination of c (the reliability of the signal in state L) and d (the reliability of the signal in state H). As discussed above, factors (e.g., *MB* and *Cvol*) that enhance the reliability in state L do not necessarily increase the signal reliability in state H. The net impact of a factor on *Fidelity* depends on the relative importance of that factor in explaining c and d. It is a design choice to construct *Fidelity* as a one-dimensional measure. For that reason, characterizing the unidirectional determinants of *Fidelity* is difficult. Whether a multi-dimensional relative of *Fidelity* can be useful is an interesting avenue for future research.

4. Empirical analysis

To assess the validity of a new earnings quality measure, one needs to show that (i) the theoretical model possesses internal validity, and (ii) the estimated measure varies with economic consequences of earnings quality in a predicted way. We assess the validity of our model and measure through a battery of tests. Our empirical analysis includes firm-year observations from 1980 to 2015 with available data for fidelity measures, accounting variables from Compustat, and stock returns from CRSP. For Section 4.4, we further require available data for restatements and comment letters from Audit Analytics.

4.1. Internal validity of the hidden Markov model

To demonstrate that our hidden Markov model aligns well with the true data-generating process of earnings surprises, we examine whether the observed transition patterns of earnings surprises are consistent with the predictions of the model.

For each firm i, we define the persistence of signal l and signal h, respectively, as

$$\rho_{il} = \Pr(y_{it} = l | y_{it-1} = l), \quad \rho_{ih} = \Pr(y_{it} = h | y_{it-1} = h). \tag{13}$$

Thus, the persistence of signal l (h) is the conditional probability of seeing another signal l (h) in the next quarter given that the current earnings signal is l (h). If our theoretical model captures the true data generating process of earnings surprises, we expect the predicted signal persistence estimated from our model to be close to the actual values of signal persistence shown in the data.

We infer the actual values of signal persistence from the frequency of signal reversals and signal continuations in the data. Specifically, for each firm i, the persistence of signal l in the actual data, ρ^A_{il}, is the fraction of quarters with signal l for which the next quarter's signal is also l; the persistence of signal h in the actual data, ρ^A_{ih}, is the fraction of quarters with signal h for which the next quarter's signal is also h.

To derive analytically tractable expressions of earnings signal persistence from our theoretical model, we assume each firm's underlying state is at its stationary distribution. That is, the probability of state transitions is stable and does not change over time:

$$\Pr(x_{it} = L) = \frac{1 - b_i}{2 - a_i - b_i}, \quad \Pr(x_{it} = H) = \frac{1 - a_i}{2 - a_i - b_i}. \tag{14}$$

With the assumption of stationarity, we show in Appendix C that our model implies the following formulas for signal persistence:

$$\rho_{il} = \frac{(1 - b_i)c_i[a_i c_i + (1 - a_i)(1 - d_i)] + (1 - a_i)(1 - d_i)[(1 - b_i)c_i + b_i(1 - d_i)]}{(1 - b_i)c_i + (1 - a_i)(1 - d_i)}, \tag{15}$$

$$\rho_{ih} = \frac{(1 - b_i)(1 - c_i)[a_i(1 - c_i) + (1 - a_i)d_i] + (1 - a_i)d_i[b_i d_i + (1 - b_i)(1 - c_i)]}{(1 - b_i)(1 - c_i) + (1 - a_i)d_i}. \tag{16}$$

Table 2
The persistence of earnings signals.

Panel A: Actual and implied persistence in earnings signals

Variable	Mean	SD	25 pctl	Median	75 pctl
$\widehat{a}_i$	0.723	0.027	0.707	0.725	0.740
$\widehat{b}_i$	0.758	0.023	0.744	0.759	0.773
$\widehat{c}_i$	0.902	0.016	0.894	0.904	0.913
$\widehat{d}_i$	0.942	0.040	0.925	0.951	0.970
$\widehat{\rho}_{il}$	0.641	0.017	0.634	0.643	0.652
$\widehat{\rho}_{ih}$	0.703	0.028	0.689	0.706	0.721
ρ_{il}^A	0.632	0.126	0.563	0.643	0.714
ρ_{ih}^A	0.676	0.126	0.600	0.688	0.763
$\rho_{il}^A - \widehat{\rho}_{il}$	−0.009	0.118	−0.076	0.002	0.069
$\rho_{ih}^A - \widehat{\rho}_{ih}$	−0.027	0.103	−0.082	−0.016	0.043
$\lvert\rho_{il}^A - \widehat{\rho}_{il}\rvert$	0.092	0.077	0.035	0.073	0.129
$\lvert\rho_{ih}^A - \widehat{\rho}_{ih}\rvert$	0.080	0.071	0.028	0.060	0.111

Panel B: Prediction errors by availability of firm data

Variable	T_i quintile				
	1 (low)	2	3	4	5 (high)
T_i	23.394	32.184	44.553	64.005	104.886
$\rho_{il}^A - \widehat{\rho}_{il}$	−0.009	−0.015	−0.002	−0.007	−0.011
$\rho_{ih}^A - \widehat{\rho}_{ih}$	−0.050	−0.038	−0.029	−0.015	−0.003
$\lvert\rho_{il}^A - \widehat{\rho}_{il}\rvert$	0.124	0.109	0.089	0.073	0.062
$\lvert\rho_{ih}^A - \widehat{\rho}_{ih}\rvert$	0.116	0.095	0.080	0.063	0.046

Panel A presents the descriptive statistics for the actual and model-predicted persistence of earnings signals. Panel B presents the means of the variables within quintiles of firms ranked on the number of available quarters of data. For each firm i, $\widehat{a}_i$, $\widehat{b}_i$, $\widehat{c}_i$, $\widehat{d}_i$ are the time-series means of $\widehat{a}_{it}$, $\widehat{b}_{it}$, $\widehat{c}_{it}$, and $\widehat{d}_{it}$, respectively. $\widehat{\rho}_{il}$ and $\widehat{\rho}_{ih}$ are calculated based on Eqs. (15) and (16). For each firm i, the persistence of signal l in the actual data, ρ_{il}^A, is the fraction of quarters with signal l for which the next quarter's signal is also l; the persistence of signal h in the actual data, ρ_{ih}^A, is the fraction of quarters with signal h for which the next quarter's signal is also h. $\rho_{is}^A - \widehat{\rho}_{is}$ is the prediction error. $\lvert\rho_{is}^A - \widehat{\rho}_{is}\rvert$ is the absolute value of the prediction error. T_i is the number of quarters in the estimation sample for firm i. $\widehat{\rho}_{is}$, ρ_{is}^A, and $\lvert\rho_{is}^A - \widehat{\rho}_{is}\rvert$ are winsorized at 1% and 99%. The sample consists of 8,578 firms.

We then calculate the predicted values of signal persistence $\widehat{\rho}_{il}$ and $\widehat{\rho}_{ih}$ by plugging estimated firm-specific parameters into Eqs. (15) and (16). For each firm i, the estimated parameters $\widehat{a}_i$, $\widehat{b}_i$, $\widehat{c}_i$, and $\widehat{d}_i$ are the time-series means of $\widehat{a}_{it}$, $\widehat{b}_{it}$, $\widehat{c}_{it}$, and $\widehat{d}_{it}$, respectively.

Table 2, Panel A presents the descriptive statistics for the actual and model-predicted persistence of earnings signals, for a sample of 8578 firms. The predicted values are fairly close to the actual values of signal persistence. The mean prediction error for the persistence of signal l, $\rho_{il}^A - \widehat{\rho}_{il}$, is −0.009; the mean prediction error for the persistence of signal h, $\rho_{ih}^A - \widehat{\rho}_{ih}$, is −0.027. The interquartile range of the prediction error is 0.145 for signal l and 0.125 for signal h. The absolute values of the prediction errors are also reasonably small. The mean of $\lvert\rho_{il}^A - \widehat{\rho}_{il}\rvert$ is 0.092 and the mean of $\lvert\rho_{ih}^A - \widehat{\rho}_{ih}\rvert$, 0.080.

The prediction error has two possible sources: (i) the discrepancy between our theoretical model and the true data generating process, and (ii) the discrepancy between the actual state transition probability and our simplifying assumption of stationarity. Although the discrepancy from the true data generating process is unlikely to vary with the length of the time-series for each firm, the stationarity assumption is more likely to be descriptive of a longer time series than a shorter time series. We can thus attribute the part of the prediction error that varies with the length of history to deviations from the stationarity assumption.

In Panel B, we sort all firms into quintiles based on T_i, i.e., the number of quarters in the estimation sample for firm i. The absolute value of the errors decreases substantially from the lowest T_i quintile (0.124 and 0.116 for signal l and signal h, respectively) to the highest T_i quintile (0.062 and 0.046). This result suggests that the prediction error arising from the discrepancy between our model and the true data generating process is reasonably small after removing the prediction error due to the stationarity assumption.

Overall, the results of Table 2 suggest that our hidden Markov model possesses internal validity. Combining the estimates with the assumed theoretical structure, we recover the observed earnings surprise patterns in the input data.

4.2. Relationship with other measures of quality

We compare *Fidelity* to three existing measures of earnings quality or more broadly, accounting quality. The first measure is earnings smoothness (*Smoothness*), calculated as the ratio of the standard deviation of net income before extraordinary items divided by the standard deviation of cash flow from operations over year t-4 through t. Even though a smaller value of *Smoothness* intuitively indicates higher earnings quality, it may also be a result of earnings management (Dechow et al., 2010).

The second measure is accruals quality (*AQ*), defined as the standard deviation of firm-level residuals from the McNichols (2002) modification of the Dechow and Dichev (2002) model over the past five years. *AQ* captures the quality of the mapping between accounting earnings and cash flows. A higher value of *AQ* indicates a weaker mapping and can be interpreted as lower accounting quality.

The third measure is unexplained audit fees (*UAF*) proposed by Hribar et al. (2014). *UAF* is the portion of audit fees not explained by resources required to complete the audit and thus unrelated to audit effort. *UAF* should be negatively related to accounting quality because auditors charge higher fees to firms with lower quality accounting. More details on the measures are included in Appendix D.

Our method estimates parameters for two constructs: economic persistence and the representational faithfulness of earnings signals. *Smoothness* conflates these two constructs. *AQ* measures representational faithfulness as the standard deviation of the regression residuals unexplained by economic factors. *UAF* is premised on the notion that unexpectedly high audit fees telegraph low-quality accounting. Which measure best captures the notion of representational faithfulness is an empirical issue.

Table 3, Panel A presents the descriptive statistics of economic persistence and earnings fidelity, selected firm characteristics, and prior measures of quality. The average value of *Fidelity* for our sample is 0.921. As reported in Panel A, the middle 50% of firm-years have *Fidelity* ranging from 0.911 to 0.937. Therefore, earnings signals, though generally faithful, are also heterogeneous.

Panel B of Table 3 reports the correlations. *Fidelity* is negatively correlated with firm size, return on assets (ROA), and economic persistence (*Pers*). *Fidelity* is positively correlated with *Smoothness* (Pearson corr.: 0.086; Spearman corr.: 0.095) and *AQ* (Pearson corr.: 0.125; Spearman corr.: 0.195). Given that higher values of *Smoothness* and *AQ* both suggest low accounting quality, the positive correlations suggest that *Fidelity* captures a notion of quality distinct from *Smoothness* and *AQ*, which are based on the lack of variability in earnings or discretionary accruals. *Fidelity* is negatively correlated with *UAF* (Pearson corr.:

Table 3
Descriptive statistics and correlations.

Panel A: Descriptive statistics

Variable	*N*	Mean	SD	p25	Median	p75
Size	111,180	5.505	2.235	3.878	5.450	7.065
ROA	111,161	0.016	0.153	0.002	0.033	0.077
Smoothness	110,528	0.784	0.761	0.297	0.634	1.038
AQ	94,119	0.046	0.041	0.019	0.033	0.058
UAF	39,579	0.007	0.444	−0.267	0.000	0.285
Pers	111,180	0.745	0.022	0.729	0.745	0.761
Fidelity	111,180	0.921	0.023	0.911	0.927	0.937
Fidelity^w	111,180	0.922	0.034	0.904	0.921	0.943
Fidelity^p	12,228	0.919	0.025	0.908	0.924	0.936
Restatement	60,272	0.101	0.302	0	0	0
Comment	36,998	0.398	0.489	0	0	1

Panel B: Correlations

	Size	*ROA*	*Smoothness*	*AQ*	*UAF*	*Pers*	*Fidelity*	*Fidelity^w*	*Fidelity^p*	*Restatement*	*Comment*
Size		0.362	−0.059	−0.394	−0.036	0.939	−0.319	−0.157	0.024	0.025	0.177
ROA	0.422		−0.124	−0.256	−0.075	0.356	−0.026	0.076	0.091	0.007	0.079
Smoothness	−0.075	−0.117		0.129	0.001	−0.063	0.086	0.029	0.069	−0.021	−0.012
AQ	−0.451	−0.224	0.149		0.040	−0.471	0.125	0.014	−0.062	−0.024	−0.069
UAF	−0.031	−0.069	0.012	0.044		−0.044	−0.044	−0.038	−0.056	0.019	0.046
Pers	0.938	0.394	−0.077	−0.491	−0.035		−0.223	−0.069	0.148	0.028	0.163
Fidelity	−0.337	0.127	0.095	0.195	−0.036	−0.276		0.742	0.890	−0.053	−0.061
Fidelity^w	−0.176	0.284	0.006	0.094	−0.037	−0.145	0.647		0.704	−0.046	−0.036
Fidelity^p	0.102	0.342	0.067	−0.030	−0.062	0.180	0.822	0.585		−0.042	−0.003
Restatement	0.037	−0.033	−0.011	−0.030	0.015	0.039	−0.051	−0.052	−0.043		0.078
Comment	0.191	0.074	−0.012	−0.067	0.041	0.174	−0.083	−0.045	0.005	0.078	

Panel C: Variable means by *Fidelity* quintiles

Variable	*Fidelity* quintile				
	1 (low)	2	3	4	5 (high)
Fidelity	0.884	0.915	0.927	0.935	0.944
Size	6.718	5.908	5.620	5.149	4.129
Pers	0.753	0.748	0.746	0.743	0.733
Smoothness	0.649	0.711	0.769	0.810	0.983
AQ	0.033	0.042	0.046	0.049	0.057
UAF	0.030	0.013	0.000	−0.012	0.004
Restatement	0.120	0.108	0.097	0.091	0.083
Comment	0.431	0.416	0.384	0.376	0.347

Panel A presents descriptive statistics for selected variables. Panel B presents the correlation matrix, with Pearson (Spearman) correlations above (below) the diagonal. All correlations larger than 0.03 (0.02) in magnitude are significant at the 0.01 (0.10) level. Panel C presents the means of selected variables by *Fidelity* quintiles. Quintiles are formed on all observations with available data. All variables are defined in Appendix D. All continuous variables (except *Pers* and the fidelity measures, which are already bounded) are winsorized each year at 1% and 99%. The tabulations are based on all firm-years for which *Fidelity* can be calculated, for the period of 1980–2015. *Fidelity^p* is available for 2012–2015. *Restatement* is available for 1999–2015. *Comment* is available for 2005–2015.

−0.044; Spearman corr.: −0.036). To the extent that a higher value of *UAF* indicates lower accounting quality, the quality constructs underlying *Fidelity* and *UAF* are positively correlated.

We report the means of selected variables by quintile of *Fidelity* in Panel C of Table 3. From the lowest to the highest *Fidelity* quintile, firm size and economic persistence monotonically decrease. Both *Smoothness* and *AQ* increase as *Fidelity* increases: *Smoothness* increases from 0.649 for the lowest *Fidelity* quintile to 0.983 for the highest *Fidelity* quintile, while *AQ* increases from 0.033 to 0.057. For *UAF*, there is a "V"-shaped relationship: *UAF* decreases from 0.030 to −0.012 when moving from the lowest *Fidelity* quintile to the fourth quintile but increases to 0.004 for the highest *Fidelity* quintile.

The negative correlation between *Fidelity* and firm size as well as the notion of quality as implied by some other proxies (i.e., *Smoothness*, *AQ*) may appear counterintuitive. Researchers have not reached a consensus on the best measure of accounting quality (Dechow et al., 2010). An advantage of our measure is that it is based on a structural model of earnings quality. Our measure captures the faithfulness of a simple statistic, y_{it}, whereas traditional measures focus on properties of other metrics (e.g., accruals). Also, several previous measures (e.g., *Smoothness*, *AQ*) are easily confounded with earnings persistence, as seen in Table 3. Hence, convergent validity cannot be used as a criterion to gauge the performance of a candidate measure.

4.3. Forward earnings response coefficient

Market participants should incorporate more information about future earnings in current stock prices when earnings signals are more informative about future earnings (Collins et al., 1994; Lundholm and Myers, 2002; Ettredge et al., 2005; Hribar et al., 2014). We estimate the association between current stock price and future earnings through the following regression:

$$R_t = \delta_0 + \delta_1 E_{t-1} + \delta_2 E_t + \delta_3 E_{t+1} + \delta_4 R_{t+1} + \varepsilon_t, \tag{17}$$

where R_t is the cumulative returns over the 12-month period ending three months after the end of fiscal year *t*; and, E_{t-1}, E_t and E_{t+1} are the net incomes for fiscal years $t-1, t, t+1$, respectively, scaled by the market value of equity three months after the beginning of fiscal year *t*. The coefficient on $E_{t+1}(\delta_3)$ is also known as forward earnings response coefficient, or FERC. FERC quantifies the extent to which current stock returns capture information in future earnings. A higher FERC indicates a greater ability of investors to forecast future earnings conditional on the information included in the current earnings.

FERC depends on the extent to which current earnings inform future earnings, i.e., the persistence of earnings signals. But as we show in Section 4.1, signal persistence is a function of both economic persistence and earnings fidelity. We thus examine whether and how economic persistence and earnings fidelity affect the FERC.[21] We expect a positive relationship between earnings fidelity and the FERC, as more faithful earnings signals improve investors' ability to incorporate information about the profitability process into current prices.

On the other hand, we expect a hump-shaped relationship between economic persistence (*Pers*) and the FERC. To see this, consider the following three cases. When $a = b = 0.5$ (i.e., $Pers = 0.5$), even perfectly faithful earnings signals are not useful in predicting the future state, because the probability that the firm will be in state *H* next period is 0.5 independent of the current state and irrespective of the history of earnings signals. Therefore, the FERC is 0. When $a = b = 1$ (i.e., $Pers = 1$), the state of the firm is unchanging, so there is no information that the earnings signals can convey; therefore, the FERC is again 0. When *Pers* is intermediate, the future state of the firm is somewhat predictable, and earnings signals convey information about both the current state and the likely future state. In other words, the FERC will be greater when *Pers* is intermediate than when *Pers* is either very high or very low.

Table 4 reports the regression results. We first establish the basic facts regarding the earnings response coefficients. In column (1) of Panel A, the coefficient on E_t is the contemporaneous earnings response coefficient. It is positive and significant (0.164, $t = 6.16$). The coefficient on E_{t+1} is the FERC, which is also positive and significant (0.309, $t = 19.71$), consistent with the prior literature.

The relationship between *Fidelity* and the FERC is monotonic and can be captured by a linear regression. We augment the FERC regression model with the interactive effects of *Fidelity*. As reported in column (2) of Panel A, the coefficient on $E_{t+1} \times Fidelity_t$ is positive (3.360, $t = 8.23$). We also examine how much variation can be explained by the moderating effects of *Fidelity*. We scale the ΔR^2 for each regression by the R^2 of the respective base model to determine the relative percentage increase in R^2. Vuong's (1989) test is used to assess whether the incremental R^2 is significant. Including the effects of *Fidelity* improves the base model by 14.12%. The improvement in explanatory power is significant at the 0.001 level (Vuong *z*-stat. = 4.512).

[21] In pricing the firm, an investor attempts to infer the underlying state from the history of earnings signals. In this paper, we do not formally study the implications of the model for the earnings-price relationships. Doing so would require assumptions about the price formation process. See Du and Huddart (2019) for such an analysis.

Table 4
Forward earnings response coefficient (FERC).

Panel A: FERC regressions

Fidelity measure:	(1)	(2)	(3)
	Fidelity	*Fidelity*	*Fidelity*^{w}
Intercept	0.185***	0.287***	−1.010***
	(99.75)	(3.43)	(−14.33)
E_{t-1}	−0.368***	2.148***	0.534*
	(−14.93)	(4.55)	(1.92)
E_t	0.164***	−2.940***	−1.480***
	(6.16)	(−5.92)	(−4.25)
E_{t+1}	0.309***	−2.732***	−1.138***
	(19.71)	(−7.48)	(−2.75)
R_{t+1}	−0.113***	0.074	0.103
	(−31.90)	(0.44)	(0.81)
$Fidelity_t$		−0.113	1.293***
		(−1.24)	(16.88)
$E_{t-1} \times Fidelity_t$		−2.814***	−1.017***
		(−5.32)	(−3.24)
$E_t \times Fidelity_t$		3.495***	1.860***
		(6.25)	(4.73)
$E_{t+1} \times Fidelity_t$		3.360***	1.599***
		(8.23)	(3.45)
$R_{t+1} \times Fidelity_t$		−0.205	−0.235*
		(−1.12)	(−1.69)
N	104,676	104,676	104,676
R^2	7.79	8.89	9.01
Adjusted R^2	7.78	8.88	9.00
ΔR^2		1.10	1.22
$\%\Delta R^2$		14.12%	15.66%
Vuong z-statistic		4.512	5.601
p-value		< 0.001	< 0.001

Panel B: FERC regressions by *Fidelity* quintiles

Variable	*Fidelity* quintile				
	1 (low)	2	3	4	5 (high)
Intercept	0.185***	0.182***	0.185***	0.181***	0.174***
	(46.65)	(41.66)	(40.60)	(36.73)	(35.46)
E_{t-1}	−0.268***	−0.341***	−0.456***	−0.575***	−0.683***
	(−8.77)	(−7.18)	(−8.22)	(−9.91)	(−10.54)
E_t	0.047*	0.166***	0.316***	0.401***	0.723***
	(1.78)	(3.26)	(4.45)	(5.86)	(7.78)
E_{t+1}	0.214***	0.313***	0.430***	0.476***	0.551***
	(12.52)	(8.89)	(14.26)	(9.06)	(9.96)
R_{t+1}	−0.106***	−0.111***	−0.138***	−0.118***	−0.117***
	(−13.14)	(−12.76)	(−16.91)	(−13.19)	(−16.60)
N	20,936	20,935	20,935	20,935	20,935
Adjusted R^2	10.93	9.46	9.91	8.55	8.81

Panel C: FERC regressions by *Pers* quintiles

Variable	*Pers* quintile				
	1 (low)	2	3	4	5 (high)
Intercept	0.274***	0.196***	0.179***	0.154***	0.132***
	(47.04)	(38.03)	(36.24)	(32.52)	(43.63)
E_{t-1}	−0.336***	−0.562***	−0.437***	−0.333***	−0.137***
	(−9.13)	(−12.39)	(−9.24)	(−5.89)	(−3.17)
E_t	0.207***	0.417***	0.217***	0.124*	0.060**
	(3.94)	(8.24)	(3.99)	(1.91)	(2.07)
E_{t+1}	0.330***	0.478***	0.378***	0.307***	0.112***
	(9.92)	(13.11)	(10.59)	(10.40)	(5.36)
R_{t+1}	−0.136***	−0.112***	−0.105***	−0.111***	−0.107***
	(−17.42)	(−13.69)	(−13.15)	(−13.67)	(−14.50)

Table 4 (*continued*)

Panel C: FERC regressions by *Pers* quintiles					
Variable	*Pers* quintile				
	1 (low)	2	3	4	5 (high)
N	20,936	20,935	20,935	20,935	20,935
Adjusted R^2	8.95	10.35	9.76	6.72	3.23

This table examines whether measures of earnings fidelity are associated with the FERC, i.e., the market's ability to price future earnings. Panel A reports the basic FERC regression and the regression model augmented by including interaction terms with *Fidelity*. The sample consists of firm-years for the period of 1980–2015. Panel B reports the basic regression by *Fidelity* quintiles. Panel C reports the basic regression by *Pers* quintiles. Quintiles are formed on all observations with available data. All variables are defined in Appendix D. All continuous variables (except *Pers* and the fidelity measures, which are already bounded) are winsorized each year at 1% and 99%. ΔR^2 is $R^2 - R^2_{\text{base}}$, where R^2_{base} is the R^2 of the corresponding base model without regressors of interest (i.e., all regressors with *Fidelity*; the base model regressions are untabulated). R^2, ΔR^2, and adjusted R^2 are in percent. %ΔR^2 is the percentage increase in R^2 relative to the base model, i.e., $(R^2 - R^2_{\text{base}})/R^2_{\text{base}}$. The Vuong *z*-statistic and its *p*-value are based on Vuong (1989) test of whether R^2 change is statistically significant. *t*-statistics, reported in parentheses, are based on standard errors clustered by firm. *, ** and *** indicate two-tailed statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

To gain further insight on the role of *Fidelity* at a more granular level, we conduct the FERC regression by quintiles. Panel B of Table 4 shows that the impact of *Fidelity* on FERC is monotonic: the FERC increases from 0.214 for the lowest *Fidelity* quintile to 0.551 for the highest *Fidelity* quintile.[22]

On the other hand, we expect the relationship between *Pers* and the FERC to be non-monotonic. Panel C of Table 4 reports the results by *Pers* quintiles. The FERC (i.e., the coefficient on E_{t+1}) first increases and then decreases (0.330, 0.478, 0.378, 0.307, and 0.112, for the five *Pers* quintiles, respectively). This hump-shaped relationship is consistent with the model prediction.[23]

Column (3) of Table 4, Panel A reports the FERC regression using the weighted fidelity measure, $Fidelity^w$. The coefficient on $E_{t+1} \times Fidelity^w_t$ is positive and significant (1.599, $t = 3.45$), consistent with the results using *Fidelity*.

4.4. Predicting external indicators of accounting quality

4.4.1. Basic findings

To shed more light on the empirical usefulness of the earnings fidelity measures, we examine whether they are associated with the expected consequences of low-quality accounting. The literature has proposed several external indicators of low-quality accounting, including restatements and SEC comment letters (for a review, see Karpoff et al., 2017). While restatements represent more severe deviations from accounting rules, SEC comment letters are less extreme indicators of poor accounting quality. As shown in Panel A of Table 3, the percentages of observations with restatements and comment letters in our sample are 10.1% and 39.8%, respectively, which is consistent with the view that less serious reporting deficiencies trigger comment letters, while more serious reporting deficiencies may trigger both comment letters and restatements.

To examine whether our measure of informativeness is associated with the expected consequences of low-quality accounting, we estimate the following linear probability model (LPM):[24]

$$Restatement_{it} \text{ or } Comment_{it} = \gamma_0 + \gamma_1 Fidelity_{it} + Controls_{it} + \varepsilon_{it}. \tag{18}$$

Restatement is an indicator variable that equals 1 if a firm's financial statements in fiscal year *t* are subsequently restated, and 0 otherwise. *Comment* is an indicator variable that equals 1 if a firm's 10-K/10-Q filings are commented on by the SEC, and 0 otherwise. Data on restatements and SEC comment letters are obtained from Audit Analytics. We merge the restatement data with Compustat by the fiscal period that is restated. The sample of comment letters is limited to those pertaining to a 10-K or 10-Q filing. These comment letters primarily represent the potential or actual material accounting, auditing, financial reporting, or disclosure deficiencies identified by the SEC. We merge the comment letters with Compustat by requiring the filing date of the form (10-K or 10-Q) commented on by an SEC comment letter to fall within a 365-calendar-day interval ending 100 days after the fiscal year end. This interval approximately covers the filing dates of all 10-K and 10-Q filings that pertain to a fiscal year.[25]

Poor accounting quality may be associated with a battery of financial and non-financial factors. Therefore, we control for a series of firm characteristics as other determinants of external enforcement, closely following the prior literature (e.g.,

[22] The *p*-values for χ^2 tests of whether the coefficients on E_{t+1} in successive quintiles are equal are 0.011, 0.025, 0.437, and 0.227. These are the comparisons of quintiles 1 and 2, 2 and 3, 3 and 4, and 4 and 5, respectively.

[23] The *p*-values for χ^2 tests of whether the coefficients on E_{t+1} in successive quintiles are equal are 0.002, 0.048, 0.115, and <0.001. These are the comparisons of quintiles 1 and 2, 2 and 3, 3 and 4, and 4 and 5, respectively.

[24] In untabulated analysis, we use logit regressions. The conclusions are the same.

[25] Audit Analytics classifies a comment letter as related to as many as 19 different filings whose form types are stored in the data field *list_form_fkey_ed* and filing dates in *list_file_date_ed*. We use these data items to parse the exact filings to which a comment letter is related and the filing dates of these filings. A comment letter enters our sample as long as it is related to at least one 10-K or 10-Q.

Table 5
Predicting indicators of low-quality accounting.

Panel A: Linear probability model (LPM) regressions with *Fidelity*

Dep. var.:	(1)	(2)	(3)	(4)	(5)	(6)
	Restatement	*Restatement*	*Restatement*	*Comment*	*Comment*	*Comment*
Intercept	0.535***	0.604***	0.521***	1.373***	1.232***	1.241***
	(5.09)	(5.26)	(4.10)	(8.06)	(6.66)	(6.30)
Fidelity	−0.462***	−0.542***	−0.447***	−1.037***	−0.879***	−0.878***
	(−4.03)	(−4.33)	(−3.22)	(−5.57)	(−4.33)	(−4.07)
Smoothness		0.003	0.004		−0.005	−0.003
		(1.08)	(1.05)		(−0.79)	(−0.52)
AQ		0.152**	0.142*		−0.249*	−0.428***
		(2.11)	(1.75)		(−1.72)	(−2.79)
UAF			0.016***			0.059***
			(2.76)			(5.90)
Cvol	−0.006	−0.050*	−0.038	−0.215***	−0.137**	−0.129**
	(−0.24)	(−1.73)	(−1.09)	(−4.82)	(−2.38)	(−2.11)
Δ*Rec*	0.020	0.023	−0.011	0.197***	0.167**	0.180**
	(0.72)	(0.81)	(−0.33)	(3.05)	(2.51)	(2.47)
Δ*Inv*	0.054	0.055	0.113**	−0.093	−0.070	−0.084
	(1.43)	(1.41)	(2.38)	(−1.06)	(−0.77)	(−0.86)
Δ*Csale*	0.002	0.004	0.007	0.022***	0.022***	0.028***
	(0.49)	(1.11)	(1.50)	(2.83)	(2.62)	(2.87)
Δ*NI*	−0.006	−0.007	−0.006	0.008	0.009	−0.001
	(−0.70)	(−0.77)	(−0.56)	(0.32)	(0.36)	(−0.05)
Δ*Emp*	−0.011**	−0.011**	−0.014**	−0.031***	−0.034***	−0.042***
	(−2.24)	(−2.19)	(−2.29)	(−2.62)	(−2.78)	(−3.31)
MB	−0.001**	−0.001*	−0.002***	0.002	0.001	0.002
	(−1.97)	(−1.76)	(−2.65)	(1.38)	(1.23)	(1.42)
N	47,051	43,744	35,913	28,848	26,596	23,917
Adjusted R^2	0.15	0.18	0.21	0.45	0.40	0.70
R^2	0.16	0.20	0.24	0.48	0.43	0.75
ΔR^2	0.11	0.14	0.09	0.24	0.16	0.16
$\%\Delta R^2$	220.00%	233.33%	60.00%	100.00%	59.26%	27.12%
Vuong *z*-statistic	3.528	3.795	2.789	4.167	3.304	3.038
p-value	< 0.001	< 0.001	0.005	< 0.001	0.001	0.002

Panel B: LPM regressions by size quintiles

Variable	*Size* quintile				
	1 (low)	2	3	4	5 (high)
Dep. var.: *Restatement*					
Fidelity	−0.704**	−0.791**	−0.842***	−0.499**	−0.142
	(−2.35)	(−2.52)	(−2.98)	(−2.05)	(−0.62)
Smoothness	−0.000	−0.009*	0.006	0.014*	0.010
	(−0.10)	(−1.79)	(1.05)	(1.85)	(1.27)
AQ	0.106	0.277**	0.539***	0.009	−0.052
	(1.10)	(2.14)	(3.18)	(0.04)	(−0.18)
Controls	Yes	Yes	Yes	Yes	Yes
N	8749	8749	8749	8749	8748
Adjusted R^2	0.26	0.35	0.75	0.27	0.25
R^2	0.37	0.46	0.86	0.39	0.37
ΔR^2	0.17	0.20	0.21	0.10	0.02
$\%\Delta R^2$	85.00%	76.92%	32.31%	34.48%	5.71%
Vuong *z*-statistic	1.667	1.992	2.117	1.465	0.519
p-value	0.096	0.046	0.034	0.143	0.604
Dep. var.: *Comment*					
Fidelity	−0.910**	0.346	−0.691	−0.741*	0.977**
	(−2.12)	(0.90)	(−1.64)	(−1.81)	(2.32)
Smoothness	−0.003	0.002	−0.004	0.009	−0.007
	(−0.25)	(0.16)	(−0.39)	(0.79)	(−0.50)
AQ	0.117	0.029	0.326	−0.158	−0.745
	(0.54)	(0.12)	(1.05)	(−0.38)	(−1.14)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	5320	5319	5319	5319	5319
Adjusted R^2	0.19	0.10	0.46	0.23	0.50
R^2	0.38	0.29	0.64	0.42	0.68
ΔR^2	0.11	0.02	0.07	0.11	0.25
$\%\Delta R^2$	40.74%	7.41%	12.28%	35.48%	58.14%
Vuong *z*-statistic	1.160	0.489	0.977	1.190	1.862
p-value	0.246	0.624	0.329	0.234	0.063

Panel C: LPM regressions with *Fidelity*^{p}

Dep. var.:	(1)	(2)	(3)	(4)	(5)	(6)
	Restatement	*Restatement*	*Restatement*	*Comment*	*Comment*	*Comment*
Intercept	0.601***	0.612***	0.641***	1.100***	0.896***	0.953***
	(3.08)	(2.89)	(2.86)	(4.49)	(3.48)	(3.51)
Fidelity^p	−0.536**	−0.537**	−0.571**	−0.852***	−0.620**	−0.673**
	(−2.54)	(−2.34)	(−2.35)	(−3.19)	(−2.21)	(−2.28)
Smoothness		−0.011*	−0.011*		−0.009	−0.006
		(−1.72)	(−1.75)		(−0.92)	(−0.66)
AQ		−0.083	−0.027		−0.370*	−0.631***
		(−0.51)	(−0.15)		(−1.69)	(−2.76)
UAF			0.015			0.069***
			(1.24)			(4.74)
Cvol	−0.067	−0.037	−0.078	−0.371***	−0.276***	−0.264***
	(−1.34)	(−0.58)	(−1.16)	(−5.30)	(−3.05)	(−2.63)
Δ*Rec*	−0.017	−0.034	−0.048	0.354***	0.314**	0.275**
	(−0.21)	(−0.41)	(−0.59)	(2.89)	(2.54)	(2.10)
Δ*Inv*	−0.014	−0.051	−0.028	0.587***	0.652***	0.650***
	(−0.13)	(−0.48)	(−0.25)	(3.64)	(3.94)	(3.68)
Δ*Csale*	0.012	0.021**	0.018	0.033**	0.032**	0.055***
	(1.24)	(2.07)	(1.64)	(2.18)	(2.04)	(2.96)
Δ*NI*	−0.014	−0.013	−0.027	−0.026	−0.044	−0.077
	(−0.54)	(−0.49)	(−0.96)	(−0.57)	(−0.99)	(−1.64)
Δ*Emp*	0.000	−0.005	−0.017	−0.029	−0.027	−0.041*
	(0.03)	(−0.35)	(−1.09)	(−1.37)	(−1.22)	(−1.78)
MB	−0.001*	−0.001	−0.001	0.003*	0.003*	0.003*
	(−1.68)	(−1.30)	(−0.73)	(1.95)	(1.79)	(1.70)
N	9618	8828	7997	9618	8828	7997
Adjusted R^2	0.21	0.25	0.25	0.82	0.86	1.34
R^2	0.30	0.36	0.39	0.90	0.97	1.48
ΔR^2	0.17	0.15	0.18	0.17	0.09	0.10
$\%\Delta R^2$	130.77%	71.43%	85.71%	23.29%	10.23%	7.25%
Vuong *z*-statistic	1.948	1.794	1.771	1.979	1.366	1.376
p-value	0.051	0.073	0.077	0.048	0.172	0.169

This table examines whether measures of earnings fidelity are associated with indicators of low-quality accounting. Panel A (Panel C) reports the LPM regressions with *Fidelity* (*Fidelity*^{p}). Panel B replicates columns (2) and (5) of Panel A with separate regressions for each of the five *Size* quintiles. Quintiles are formed on all observations with available data. All variables are defined in Appendix D. All continuous variables (except the fidelity measures which are already bounded) are winsorized each year at 1% and 99%. ΔR^2 is $R^2 - R^2_{\text{base}}$, where R^2_{base} is the R^2 of the corresponding base model without *Fidelity* (untabulated). R^2, ΔR^2, and adjusted R^2 are in percent. $\%\Delta R^2$ is the percentage increase in R^2 relative to the base model, i.e., $(R^2 - R^2_{\text{base}})/R^2_{\text{base}}$. The Vuong *z*-statistic and its *p*-value are based on Vuong (1989) test of whether R^2 change is statistically significant. *t*-statistics, reported in parentheses, are based on standard errors clustered by firm. The sample for Panels A and B consists of firm-years with valid data for the period of 1999−2015 for restatements and 2005−2015 for comment letters. The sample for Panel C consists of firm-years over 2012−2015. *, ** and *** indicate two-tailed statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

Beneish, 1999; Dechow et al., 2011; Hribar et al., 2014). These characteristics are primarily related to the market pressures on the reporting company to report higher profits. They include cash flow volatility (*Cvol*), change in receivables (Δ*Rec*), change in inventory (Δ*Inv*), change in cash sales (Δ*Csale*), change in net income (Δ*NI*), abnormal change in employees (Δ*Emp*), and market to book ratio (*MB*). We also include the existing measures of quality (i.e., *Smoothness*, *AQ*, and *UAF*) in the regression model to gauge the incremental explanatory power of *Fidelity*. We expect firms with greater informativeness to exhibit fewer incidences of low-quality accounting (i.e., we expect γ_1 to be negative).

Columns (1) through (3) in Table 5, Panel A report the results for *Restatement*. In column (1), after controlling for other determinants of low-quality accounting, the coefficient on *Fidelity* is negative (−0.462, $t = -4.03$). In other words, firms with higher *Fidelity* are less likely to subsequently restate their accounting numbers. Comparing R^2 to a model without *Fidelity*, we find that the relative percentage increase in R^2 is 220.00% (Vuong *z*-stat. = 3.528, $p < 0.001$).

In column (2), we study the predictive power of *Fidelity* conditional on *Smoothness* and *AQ*. The coefficient on *Fidelity* remains negative (−0.542, $t = -4.33$); the incremental explanatory power relative to a model that includes *Smoothness*, *AQ*, and other controls is 233.33% (Vuong *z*-stat. = 3.795, $p < 0.001$). Further including *UAF* substantially reduces the sample size, but yields similar results. The coefficient on *Fidelity* is still significant (−0.447, $t = -3.22$); the incremental explanatory power relative to a model that excludes *Fidelity* is 60.00% (Vuong *z*-stat. = 2.789, $p = 0.005$). The smaller incremental R^2 after conditioning on *UAF* may be due to a partial overlap between the *Fidelity* and *UAF* in terms of the notion of quality captured (see Section 3).

Columns (4) through (6) of Table 5, Panel A report the results for SEC comment letters (*Comment*). We also progressively condition on other proxies of quality, i.e., *Smoothness* and *AQ* in column (5), and all three proxies in column (6). Similar to the results on restatements, the coefficient on *Fidelity* is consistently negative and significant regardless of controls (column (4): −1.037, $t = -5.57$; column (5): −0.879, $t = -4.33$; column (6): −0.878, $t = -4.07$). Furthermore, the incremental explanatory power relative to a model without *Fidelity* is large: 100.00%, 59.26%, and 27.12% (all significant at the 0.01 level based on the Vuong test), for columns (4), (5), and (6), respectively. For *AQ*, the coefficient is negative (e.g., column (5): -0.249, $t = -1.72$). Ceteris paribus, firms with higher *AQ* (lower quality) are less likely to restate earnings, which runs contrary to the economic intuition underlying *AQ*.

Overall, the evidence suggests that *Fidelity* is a valid predictor of low-quality accounting events and contains information incremental to the existing measures of accounting quality. The incremental explanatory power of *Fidelity* is large in magnitude and statistically significant. *Fidelity* also performs better when predicting less severe outcomes of low-quality accounting (i.e., SEC comment letters), relative to *AQ*.

4.4.2. Cross-sectional tests

We also examine whether there is cross-sectional variation in how *Fidelity* is associated with subsequently-issued restatements and comment letters. In particular, we are interested in how the predictive power of *Fidelity* varies with firm size. Smaller firms have more volatile operations, making it difficult to disentangle the effects of reporting from the effects of operations using any regression analysis that does not distinguish between economic persistence and reporting fidelity.

Table 5, Panel B reports the regression results by size quintile. We include *Smoothness* and *AQ* in addition to the control variables. For restatements, the coefficients on *Fidelity* are -0.704 ($t = -2.35$), -0.791 ($t = -2.52$), -0.842 ($t = -2.98$), -0.499 ($t = -2.05$), and -0.142 ($t = -0.62$), respectively, for size quintiles 1 through 5. Therefore, *Fidelity* is highly significantly related to restatements for smaller and mid-sized firms (quintiles 1 through 4). In contrast, *Smoothness* is not related to restatements in a predicted way in any of the size quintiles. *AQ* is only associated with restatements for quintiles 2 and 3. The incremental explanatory power of *Fidelity* is 85.00%, 76.92%, 32.31%, 34.48%, and 5.71%, respectively, for quintiles 1 through 5. The incremental explanatory power is statistically significant at the 0.10 level for the lowest three quintiles.

For comment letters, the coefficients on *Fidelity* are -0.910 ($t = -2.12$), 0.346 ($t = 0.90$), -0.691 ($t = -1.64$), -0.741 ($t = -1.81$), and 0.977 ($t = 2.32$), respectively, for size quintiles 1 through 5. In other words, *Fidelity* is related to comment letters in the predicted direction for quintiles 1, 3, and 4; however, only the coefficient estimates for quintiles 1 and 4 are significant at conventional levels. Neither *Smoothness* nor *AQ* is related to comment letters in any of the quintiles. Relative to Panel A in which all observations are pooled, the inconsistent sign and reduced significance of *Fidelity* may be attributed to the smaller number of observations in each quintile.

Overall, the evidence suggests that (i) *Fidelity* is a superior measure of earnings quality in predicting restatements, especially for smaller firms; and (ii) *Fidelity* performs better than *Smoothness* and *AQ* in predicting comment letters in the full sample, though the analysis of size quintiles yields mixed results. The superiority we document may result from two factors. First, conventional measures such as *AQ* rely on proxies for economic states, but the proxies may be less reliable for smaller firms. The hidden Markov model, in contrast, does not rely on proxies of unobservable states. Second, the Bayesian hierarchial method adapts the weights on firm-specific data and the population distribution to the availability of firm-specific data. No such adaptation is possible in conventional methods.

4.4.3. Look-ahead bias

Estimation of the Bayesian hierarchical model does not require long time-series for individual firms. To improve the accuracy of the estimates, we use all firm-years in our sample to estimate our main measure, *Fidelity*, which introduces a look-ahead bias in the prediction tests. Though the impact of such a bias should be small, we conduct a robustness check by developing an alternative measure free of such look-ahead bias.

We divide the sample period into an estimation period and a validation period. Specifically, we estimate the model based on the period before 2012 and use those parameter estimates to calculate an alternative measure of fidelity, denoted $Fidelity^p$, for years 2012–2015.

In Panel C of Table 5, the dependent variable, either *Restatement* or *Comment*, is regressed on $Fidelity^p$ and other proxies of quality in various specifications with control variables for firm-years in the period of 2012–2015. Despite the shorter sample period for validation, in these prediction tests, we still find a negative and significant coefficient on $Fidelity^p$ in all of the specifications. Focusing on columns (2) and (5), which predict *Restatement* and *Comment*, respectively, using $Fidelity^p$, *Smoothness*, and *AQ*, it is notable that the coefficient on $Fidelity^p$ has the predicted sign and is highly significant (column (2): -0.537, $t = -2.34$; column (5): -0.620, $t = -2.21$). In contrast, the coefficients on *Smoothness* and *AQ* are either not significantly different from 0 or are significant, but with the wrong sign.

In all specifications, the percentage increases in R^2 due to the inclusion of $Fidelity^p$ are somewhat smaller than the analogous increases reported in the baseline analysis using *Fidelity* (Table 5, Panel A). The increase in explanatory power is significant in columns (1) through (4).

4.4.4. Additional tests

The impact of covariates. We conduct two robustness checks to address the possibility that one of the covariates, *Cvol*, correlates with conventional measures of accounting quality and thereby drives the results of the validation tests. First, we reestimate the model after dropping *Cvol*. In the untabulated analysis, we show that the resulting measure yields the same results, suggesting that our measure is not driven by the inclusion of *Cvol* as a covariate. Second, we control for *Size* and *Lev* in the predictive tests. The conclusions remain the same.

Matched samples. There are far fewer restating firms than non-restating: about 10% of firm-years are classified as restatements in our sample. To address the possibility that confounding factors drive our results on restatements, we match each restating firm to a non-restating firm based on a two-digit SIC code and the market capitalization at the beginning of the year. The analysis using matched samples yields qualitatively the same results as the analysis using the full sample.

5. Concluding remarks

We propose, construct, and validate a new measure of earnings quality, *Fidelity*. The structural model underpinning the estimates separates the underlying economic condition of the firm from the reporting system for earnings. The estimation of *Fidelity* does not require more data than existing measures (e.g., smoothness, accruals quality) and is based on a single earnings metric. It proves particularly useful in assessing the earnings quality of smaller firms.

An important conceptual difference between our study and the studies of earnings persistence is that the latter do not distinguish between earnings signals and underlying states. If states are distinct from signals, then valuation is necessarily a filtering problem whereby the history of a firm's earnings is used to infer the firm's current state. Our view is that earnings persistence is related to, but distinct from, both economic persistence and earnings fidelity.

Although we have focused on earnings, the method we deploy can be applied to other research settings. In settings where the latent states cannot be reliably captured by empirical proxies, it may prove particularly effective. For example, one could estimate the faithfulness of other types of accounting information (e.g., the recognition of a loss contingency) in revealing some underlying state of the firm. The key advantage of the method is that it does not rely on empirical proxies of the latent states.

Our measure does not differentiate among causes of less faithful earnings signals. It is intuitive that earnings fidelity is determined partly by the details of a given framework of accounting rules and partly by the managers' reporting decisions within that framework. Future research could seek to distinguish these components and assess the relative importance of each component.

Appendix

Appendix A: The MCMC algorithm for estimating the model

Here we detail the full conditional distributions of our Bayesian hierarchical framework. Let $i = 1, 2, \ldots, n$ index the idiosyncratic firms.

1. We assume the prior distribution for the initial state $\boldsymbol{\pi}_0$ is Dirichlet$(1,1)$. The resulting full conditional distribution of $\boldsymbol{\pi}_0$ is also a Dirichlet distribution. Because

$$\pi(\boldsymbol{\pi}_0|\boldsymbol{x},\boldsymbol{y}) \propto \pi(\boldsymbol{x}|\boldsymbol{\pi}_0)\pi(\boldsymbol{\pi}_0), \tag{A.1}$$

we have

$$\boldsymbol{\pi}_0|\boldsymbol{x},\boldsymbol{y} \sim \text{Dirichlet}(1+n_L, 1+n_H), \tag{A.2}$$

where $n_L = \#\{i : x_{i1} = L\}, n_H = \#\{i : x_{i1} = H\}$.

2. We assume the firm-level intercepts in Eq. (3), α^*_{iL} and α^*_{iH}, follow normal distributions, $N(\overline{\alpha}_L, \sigma_L^2)$ and $N(\overline{\alpha}_H, \sigma_H^2)$, respectively. The resulting full conditional distribution of $\boldsymbol{\alpha}^*_i = (\alpha^*_{iL}, \alpha^*_{iH})$ is not a standard probability distribution. Thus, we use the Metropolis-Hasting algorithm to sample the parameters. First, we generate the candidate values from a normal proposal distribution. Then, we calculate the acceptance probability ξ. With probability ξ we accept the candidate draws. To ensure good convergence of the algorithm, the variance of the proposal distribution Λ_α^2 is set such that 25%–50% of the proposal draws are accepted. The Metropolis-Hasting algorithm is implemented as follows:

(a) Generate the candidate values from the following distributions:

$$\alpha^*_{iL} \sim N\left(\alpha^{*(g-1)}_{iL}, \Lambda_\alpha^2\right), \quad \alpha^*_{iH} \sim N\left(\alpha^{*(g-1)}_{iH}, \Lambda_\alpha^2\right). \tag{A.3}$$

(b) Calculate the acceptance probability ξ:

$$\xi = \min\left\{1, \frac{\Phi\left(\frac{\alpha_{iL}^{*} - \overline{\alpha}_{L}^{(g-1)}}{\sigma_{L}^{(g-1)}}\right)\Phi\left(\frac{\alpha_{iH}^{*} - \overline{\alpha}_{H}^{(g-1)}}{\sigma_{H}^{(t-1)}}\right)}{\Phi\left(\frac{\alpha_{iL}^{*(g-1)} - \overline{\alpha}_{L}^{(g-1)}}{\sigma_{L}^{(g-1)}}\right)\Phi\left(\frac{\alpha_{iH}^{*(g-1)} - \overline{\alpha}_{H}^{(g-1)}}{\sigma_{H}^{(g-1)}}\right)} \prod_{t=2}^{T_i}\left\{\frac{q_{itx_{it-1}x_{it}}\left(\boldsymbol{\alpha}_i^{*}, \boldsymbol{\alpha}^{(g-1)}\right)}{q_{itx_{it-1}x_{it}}\left(\boldsymbol{\alpha}_i^{*(g-1)}, \boldsymbol{\alpha}^{(g-1)}\right)}\right\}\right\} \tag{A.4}$$

where

$$q_{itrs}(\boldsymbol{\alpha}_i^{*}, \boldsymbol{\alpha}) = \Pr(x_{it} = s | x_{it-1} = r) = \begin{cases} 1 - \Phi(\alpha_{ir}^{*} + \boldsymbol{z}_{it}'\boldsymbol{\alpha}_r) & \text{if } s = L \\ \Phi(\alpha_{ir}^{*} + \boldsymbol{z}_{it}'\boldsymbol{\alpha}_r) & \text{if } s = H. \end{cases} \tag{A.5}$$

(c) Generate u from a uniform distribution on $[0,1]$, i.e., $u \sim U(0,1)$. If $u \le \xi$, accept $(\alpha_{iL}^{*(g)}, \alpha_{iH}^{*(g)}) = (\alpha_{iL}^{*}, \alpha_{iH}^{*})$, otherwise $(\alpha_{iL}^{*(g)}, \alpha_{iH}^{*(g)}) = (\alpha_{iL}^{*(g-1)}, \alpha_{iH}^{*(g-1)})$.

3. We assume the variances of firm-level intercepts in Eq. (3), σ_L^2 and σ_H^2, follow the same inverse-gamma distribution, InvGamma$(0.001, 0.001)$. The resulting full conditional distributions of $\boldsymbol{\sigma}_\alpha^2 = (\sigma_L^2, \sigma_H^2)$ are inverse-gamma distributions. Because

$$\pi\left(\sigma_r^2 | \overline{\alpha}_r, \alpha_{ir}^{*}\right) \propto \prod_{i=1}^{n} \pi\left(\alpha_{ir}^{*} | \overline{\alpha}_r, \sigma_r^2\right)\pi\left(\sigma_r^2\right), \tag{A.6}$$

we have

$$\sigma_r^2 \Big| \overline{\alpha}_r, \alpha_{ir}^{*} \sim \text{InvGamma}\left(0.001 + \frac{n}{2}, 0.001 + \sum_{i=1}^{n} \frac{(\alpha_{ir}^{*} - \overline{\alpha}_r)^2}{2}\right), \quad r = L, H. \tag{A.7}$$

4. We assume the means of firm-level intercepts in Eq. (3), $\overline{\alpha}_L$ and $\overline{\alpha}_H$, follow the same normal distribution, $N(0, 10)$. The resulting full conditional distributions of $\overline{\boldsymbol{\alpha}}_0 = (\overline{\alpha}_L, \overline{\alpha}_H)$ are normal distributions. Because

$$\pi\left(\overline{\alpha}_r \Big| \alpha_{ir}^{*}, \sigma_r^2\right) \propto \prod_{i=1}^{n} \pi\left(\alpha_{ir}^{*} | \overline{\alpha}_r, \sigma_r^2\right)\pi(\overline{\alpha}_r), \tag{A.8}$$

we have

$$\overline{\alpha}_r \Big| \alpha_{ir}^{*}, \sigma_r^2 \sim N(\mu_r^\alpha, \Lambda_r^\alpha), \quad r = L, H, \tag{A.9}$$

where

$$\mu_r^\alpha = \frac{10\sum_{i=1}^{n} \alpha_{ir}^{*}}{(\sigma_r^2 + 10n)}, \quad \Lambda_r^\alpha = \frac{10\sigma_r^2}{\sigma_r^2 + 10n}. \tag{A.10}$$

5. For the purpose of model identification, we have assumed that $\beta_{iL}^{*} \le \beta_{iH}^{*}$. To facilitate this constraint in the MCMC sampler, we introduce an additional variable δ_i such that $\beta_{iH}^{*} = \beta_{iL}^{*} + \exp(\delta_i)$. Thus, sampling the firm-level intercepts in Eq. (5) (β_{iL}^{*} and β_{iH}^{*}) is equivalent to sampling β_{iL}^{*} and δ_i. Now, we assume the prior distributions for β_{iL}^{*} and δ_i are normal distributions $N(\overline{\beta}_L, \omega_L^2)$ and $N(\overline{\delta}_H, \omega_H^2)$. The resulting full conditional distribution of $\boldsymbol{\beta}_i^{*} = (\beta_{iL}^{*}, \beta_{iH}^{*})$ is not a standard probability distribution, and we use the Metropolis-Hasting algorithm to sample these parameters:

(a) Generate the candidate values from the following distributions:

$$\beta_{iL}^{*} \sim N\left(\beta_{iL}^{*(g-1)}, \Lambda_\beta^2\right), \quad \delta_i \sim N\left(\delta_i^{(g-1)}, \Lambda_\delta^2\right). \tag{A.11}$$

(b) Calculate $\beta_{iH}^{*} = \beta_{iL}^{*} + \exp(\delta_i)$.

(c) Calculate the acceptance probability ξ :

$$\xi = \min\left\{1, \frac{\Phi\left(\frac{\beta^*_{iL}-\overline{\beta}_L^{(g-1)}}{\omega_L^{(g-1)}}\right)\Phi\left(\frac{\delta_i-\overline{\delta}_H^{(g-1)}}{\omega_H^{(g-1)}}\right)}{\Phi\left(\frac{\beta^{*(g-1)}_{iL}-\overline{\beta}_L^{(g-1)}}{\omega_L^{(g-1)}}\right)\Phi\left(\frac{\delta_i^{(g-1)}-\overline{\delta}_H^{(g-1)}}{\omega_H^{(g-1)}}\right)}\prod_{t=1}^{T_i}\left\{\frac{\eta_{itx_{it}y_{it}}\left(\boldsymbol{\beta}^*_i, \boldsymbol{\beta}^{(g-1)}\right)}{\eta_{itx_{it}y_{it}}\left(\boldsymbol{\beta}^{*(g-1)}_i, \boldsymbol{\beta}^{(g-1)}\right)}\right\}\right\} \tag{A.12}$$

where

$$\eta_{itsj}(\boldsymbol{\beta}^*_i, \boldsymbol{\beta}) = \Pr(y_{it}=j|x_{it}=s) = \begin{cases} 1-\Phi(\beta^*_{is}+\boldsymbol{u}'_{it}\boldsymbol{\beta}_s) & \text{if } j=l \\ \Phi(\beta^*_{is}+\boldsymbol{u}'_{it}\boldsymbol{\beta}_s) & \text{if } j=h. \end{cases} \tag{A.13}$$

(d) Generate $u \sim \text{unif}(0,1)$; if $u \le \xi$, accept $(\beta^{*(g)}_{iL}, \delta_i^{(g)}) = (\beta^*_{iL}, \delta_i)$, otherwise set $(\beta^{*(g)}_{iL}, \delta_i^{(g)}) = (\beta^{(g-1)}_{iL}, \delta_i^{(g-1)})$.

(e) Set $\beta^{*(g)}_{iH} = \beta^{*(g)}_{iL} + \exp(\delta_i^{(g)})$. Λ_β and Λ_δ are set such that 25%–50% of the proposal draws are accepted.

6. We assume the variances of firm-level intercepts in Eq. (5), ω_L^2 and ω_H^2, follow the same inverse-gamma distribution, InvGamma(0.001, 0.001). The resulting full conditional distributions of $\boldsymbol{\omega}^2_\beta = (\omega_L^2, \omega_H^2)$ are inverse-gamma distributions. Because

$$\pi\left(\omega_L^2|\beta^*_{iL}, \overline{\beta}_L\right) \propto \prod_{i=1}^n \pi\left(\beta^*_{iL}|\overline{\beta}_L, \omega_L^2\right)\pi\left(\omega_L^2\right), \tag{A.14}$$

$$\pi\left(\omega_H^2|\delta_i, \overline{\delta}_H\right) \propto \prod_{i=1}^n \pi\left(\delta_i|\overline{\delta}_H, \omega_H^2\right)\pi\left(\omega_H^2\right), \tag{A.15}$$

we have

$$\omega_L^2\Big|\beta^*_{iL}, \overline{\beta}_L \sim \text{InvGamma}\left(0.001+\frac{n}{2}, 0.001+\sum_{i=1}^n \frac{\left(\beta^*_{iL}-\overline{\beta}_L\right)^2}{2}\right), \tag{A.16}$$

$$\omega_H^2\Big|\delta_i, \overline{\delta}_H \sim \text{InvGamma}\left(0.001+\frac{n}{2}, 0.001+\sum_{i=1}^n \frac{\left(\delta_i-\overline{\delta}_H\right)^2}{2}\right). \tag{A.17}$$

7. We assume the means of firm-level intercepts in Eq. (5), $\overline{\beta}_L$ and $\overline{\delta}_H$, follow the normal distribution, $N(0, 10)$. The resulting full conditional distributions of $\overline{\boldsymbol{\beta}}_0 = (\overline{\beta}_L, \overline{\delta}_H)$ are normal distributions. Because

$$\pi\left(\overline{\beta}_L\Big|\beta^*_{iL}, \omega_L^2\right) \propto \prod_{i=1}^n \pi\left(\beta^*_{iL}|\overline{\beta}_L, \omega_L^2\right)\pi(\overline{\beta}_L), \tag{A.18}$$

$$\pi\left(\overline{\delta}_H\Big|\delta_i, \omega_H^2\right) \propto \prod_{i=1}^n \pi\left(\delta_i|\overline{\delta}_H, \omega_H^2\right)\pi(\overline{\delta}_H), \tag{A.19}$$

we have

$$\overline{\beta}_L\Big|\beta^*_{iL}, \omega_L^2 \sim N(\mu_L^\beta, \Lambda_L^\beta), \tag{A.20}$$

$$\overline{\delta}_H\Big|\delta_i, \omega_H^2 \sim N(\mu_H^\beta, \Lambda_H^\beta), \tag{A.21}$$

where

$$\mu_L^\beta = \frac{10\sum_{i=1}^n \beta^*_{iL}}{\omega_L^2+10n}, \quad \mu_H^\beta = \frac{10\sum_{i=1}^n \delta_i}{\omega_H^2+10n}, \quad \Lambda_L^\beta = \frac{10\omega_L^2}{\omega_L^2+10n}, \quad \Lambda_H^\beta = \frac{10\omega_H^2}{\omega_H^2+10n}. \tag{A.22}$$

8. To facilitate sampling $\boldsymbol{\alpha}_r$ and $\boldsymbol{\beta}_s$, we follow Albert and Chib (1993) and introduce two auxiliary variables, k_{it} and r_{it}:

$$k_{it} = \alpha^*_{ix_{it-1}} + \boldsymbol{z}'_{it}\boldsymbol{\alpha}_{x_{it-1}} + \varepsilon^*_{it}, \quad \varepsilon^*_{it} \sim N(0,1), \; r_{it} = \beta_{ix_{it}} + \boldsymbol{u}'_{it}\boldsymbol{\beta}_{x_{it}} + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0,1), \tag{A.23}$$

so that

$$x_{it} = \begin{cases} H & \text{if } k_{it} \geq 0 \\ L & \text{if } k_{it} < 0 \end{cases}, \quad y_{it} = \begin{cases} h & \text{if } r_{it} \geq 0 \\ l & \text{if } r_{it} < 0 \end{cases}. \tag{A.24}$$

These two auxiliary variables are treated as additional unknown parameters, and we analyze them jointly with the other parameters. Given the sampled states $\boldsymbol{x}$ and other parameters, the full conditional distribution of k_{it} is a truncated normal distribution:

$$\pi(k_{it}|\alpha_i^*, \boldsymbol{\alpha}_r, \boldsymbol{x}) = \begin{cases} TN_0^{\infty}\left(\alpha^*_{ix_{it-1}} + \boldsymbol{z}'_{it}\boldsymbol{\alpha}_{x_{it-1}}, 1\right) & \text{if } x_{it} = H \\ TN_{-\infty}^{0}\left(\alpha^*_{ix_{it-1}} + \boldsymbol{z}'_{it}\boldsymbol{\alpha}_{x_{it-1}}, 1\right) & \text{if } x_{it} = L. \end{cases} \tag{A.25}$$

Here, $TN_{\underline{a}}^{\overline{a}}(\mu, \sigma^2)$ is a truncated normal distribution with mean μ and variance σ^2 that is truncated below at $\underline{a}$ and above at $\overline{a}$. Given the sampled underlying states $\boldsymbol{x}$ and other parameters, the full conditional distribution of r_{it} is also a truncated normal distribution:

$$\pi(r_{it}|\beta_i^*, \boldsymbol{\beta}_s, \boldsymbol{x}, \boldsymbol{y}) = \begin{cases} TN_0^{\infty}\left(\beta^*_{ix_{it}} + \boldsymbol{u}'_{it}\boldsymbol{\beta}_{x_{it}}, 1\right) & \text{if } y_{it} = h \\ TN_{-\infty}^{0}\left(\beta^*_{ix_{it}} + \boldsymbol{u}'_{it}\boldsymbol{\beta}_{x_{it}}, 1\right) & \text{if } y_{it} = l. \end{cases} \tag{A.26}$$

9. We assume the regression coefficient $\boldsymbol{\alpha}_r$, $r = L, H$, in Eq. (3) follows a normal distribution, $\pi(\boldsymbol{\alpha}_r) \sim N(0_P, 10 \cdot \boldsymbol{I}_{P\times P})$. The resulting full conditional distribution of $\boldsymbol{\alpha}_r$ is a normal distribution. Because

$$\pi\Big(\boldsymbol{\alpha}_r|\alpha_i^*, \boldsymbol{k}, \boldsymbol{x}, \boldsymbol{y}\Big) \propto \prod_{idr} \pi\Big(k_{it}|\boldsymbol{\alpha}_r, \alpha_i^*, \boldsymbol{x}, \boldsymbol{y}\Big)\pi(\boldsymbol{\alpha}_r), \tag{A.27}$$

we have

$$\boldsymbol{\alpha}_r|\alpha_i^*, \boldsymbol{k}, \boldsymbol{x}, \boldsymbol{y} \sim N(\boldsymbol{m}_\alpha^r, \boldsymbol{V}_\alpha^r), \tag{A.28}$$

where

$$\boldsymbol{V}_\alpha^r = \Big((10 I_{P\times P})^{-1} + \boldsymbol{z}'_{idr}\boldsymbol{z}_{idr}\Big)^{-1}, \quad \boldsymbol{m}_\alpha^r = \boldsymbol{V}_\alpha^r\big(\boldsymbol{z}'_{idr}(\boldsymbol{k}_{idr} - \alpha^*_{idr})\big), idr = \{(i,t) : x_{it-1} = r\} \tag{A.29}$$

10. We assume the regression coefficient $\boldsymbol{\beta}_s$, $s = L, H$, in Eq. (5) follows a normal distribution, $\pi(\boldsymbol{\beta}_s) \sim N(0_L, 10 \cdot \boldsymbol{I}_{L\times L})$. The resulting full conditional distribution of $\boldsymbol{\beta}_s$ is a normal distribution. Because

$$\pi(\boldsymbol{\beta}_s|\beta_i^*, \boldsymbol{r}, \boldsymbol{x}, \boldsymbol{y}) \propto \prod_{ids} \pi(r_{it}|\beta_i^*, \boldsymbol{\beta}_s)\pi(\boldsymbol{\beta}_s), \tag{A.30}$$

we have

$$\boldsymbol{\beta}_s|\beta_i^*, \boldsymbol{r}, \boldsymbol{x}, \boldsymbol{y} \sim N(\boldsymbol{m}_\beta^s, \boldsymbol{V}_\beta^s), \tag{A.31}$$

where

$$\boldsymbol{V}_\beta^s = \Big((10 I_{L\times L})^{-1} + \boldsymbol{u}'_{ids}\boldsymbol{u}_{ids}\Big)^{-1}, \quad \boldsymbol{m}_\beta^s = \boldsymbol{V}_\beta^s\big(\boldsymbol{u}'_{ids}(\boldsymbol{r}_{ids} - \beta^*_{ids})\big), ids = \{(i,t) : x_{it} = s\} \tag{A.32}$$

11. One way to sample the underlying states $\boldsymbol{x}$ is to draw each x_{it} from its full conditional distribution one at a time. Such MCMC sampler suffers from slow mixing, however, due to the high dependency among draws of x_{it}. Our preferred approach is to use the forward and backward (FB) sampling method introduced in Chib (1996) and sample the whole vector of $\boldsymbol{x}_i$ from $\pi(\boldsymbol{x}_i|\boldsymbol{\theta}, \boldsymbol{y}_i)$ directly. For implementation of the FB sampling method, we refer readers to Scott (2002).

Below is the outline of our MCMC algorithm:

1. Draw a starting point θ^0 from the prior distribution, and determine the number of total iterations G.
2. For $g = 1, 2, \ldots, G$, sample from the full conditional distributions as follows:
(a) Sample $\boldsymbol{\pi}_0^{(g)}$ from the Dirichlet distribution given by Eq. (A.1).
(b) Sample $(\alpha_{iL}^{*(g)}, \alpha_{iH}^{*(g)}), i = 1, 2, \cdots, n$, from the non-standard distributions given by Eq. (A.3) using the Metropolis-Hasting algorithm.

(c) Sample $\sigma_r^{2(g)}$, $r = L, H$, from the inverse gamma distributions given by Eq. (A.7).
(d) Sample $\overline{\alpha}_r^{(g)}$, $r = L, H$, from the normal distributions given by Eq. (A.9).
(e) Sample $(\beta_{iL}^{*(g)}, \delta_i^{(g)}), i = 1, 2, \cdots, n$ from the non-standard distributions given by Eq. (A.11) using the Metropolis-Hasting algorithm.
(f) Sample $\omega_r^{(g)}$, $r = L, H$, from the inverse gamma distributions given by Eqs. (A.16) and (A.17).
(g) Sample $(\overline{\beta}_L^{(g)}, \overline{\delta}_H^{(g)})$, $r = L, H$, from the normal distributions given by Eqs. (A.20) and (A.21).
(h) Sample $\boldsymbol{x}^{(g)}$ using the forward and backward (FB) sampling method introduced in Chib (1996) and Scott (2002).
(i) Sample $k_{it}^{(g)}, i = 1, 2, \cdots, n$ and $t = 1, 2, \cdots, T_i$, from the truncated normal distribution given by Eq. (A.25).
(j) Sample $\boldsymbol{\alpha}_r^{(g)}, r = L, H$, from the normal distribution given by Eq. (A.28).
(k) Sample $r_{it}^{(g)}, i = 1, 2, \cdots, n$ and $t = 1, 2, \cdots, T_i$, from a truncated normal distribution given by Eq. (A.26).
(l) Sample $\boldsymbol{\beta}_s^{(g)}, s = L, H$, from a normal distribution given by Eq. (A.31).
3. After G draws of the parameters are collected, we discard the first M draws of the parameters and use the subsequent $G-M$ draws to estimate the parameters of interest.

Appendix B: Simulation analysis

We conduct a simulation analysis to examine how well model parameters can be recovered from our MCMC algorithm. In the first simulation study, we generate synthetic data from a special case of our model with no covariates ($\boldsymbol{z}$ and $\boldsymbol{u}$) in the transition and emission functions; in the second study, we add covariates to the transition and emission functions when generating the data.

B.1 No covariates

To demonstrate that our model identification does not solely rely on covariates, we include no covariates in both transition and emission functions when generating data in the first study. Thus, the transition probabilities a_i and b_i are completely determined by $\boldsymbol{\alpha}_i^* = (\alpha_{iL}^*, \alpha_{iH}^*)$ and the emission probabilities c_i and d_i are completely determined by $\boldsymbol{\beta}_i^* = (\beta_{iL}^*, \beta_{iH}^*)$.

Recall, the distribution of firm-specific parameters (a_i, b_i, c_i, d_i) is jointly determined by the mean parameters $(\overline{\boldsymbol{\alpha}}_0, \overline{\boldsymbol{\beta}}_0)$ and the heterogeneity parameters $(\boldsymbol{\sigma}_\alpha^2, \boldsymbol{\omega}_\beta^2)$. To generate the synthetic data, we first determine the value of the mean parameters $(\overline{\boldsymbol{\alpha}}_0, \overline{\boldsymbol{\beta}}_0)$ and the heterogeneity parameters $(\boldsymbol{\sigma}_\alpha^2, \boldsymbol{\omega}_\beta^2)$. Then, for each firm $\boldsymbol{i}$, we generate the firm-level intercepts, $\boldsymbol{\alpha}_i^*$ and $\boldsymbol{\beta}_i^*$, from their assumed distributions and then calculate the firm-level parameters (a_i, b_i, c_i, d_i). Last, we generate the observed signal $\boldsymbol{y}_i$ according to our model and apply our MCMC algorithm to recover the parameters.

In the first study, we fix all heterogeneity parameters at 0.1 and examine the performance of our estimation algorithm across different mean levels of economic persistence and earnings fidelity. Based on Section 2.1 and Appendix A, denote

$$a = 1 - \Phi(\overline{\alpha}_L), \quad b = \Phi(\overline{\alpha}_H), \quad c = 1 - \Phi(\overline{\beta}_L), \quad d = \Phi(\overline{\beta}_L + \exp(\overline{\delta}_H)). \tag{A.33}$$

The parameters (a, b, c, d) can be viewed as the persistence and fidelity levels for a representative firm in the data. We consider two levels of persistence and two levels of fidelity when generating the synthetic data. In the low persistence condition, we assign low values to the overall transition probabilities ($a = 0.65$, $b = 0.70$);[26] in the high persistence condition, we assign high values to the overall transition probability ($a = 0.80$, $b = 0.85$). Similarly, in the low fidelity condition, we assign low values to the overall emission probability ($c = 0.75$, $d = 0.80$), and in the high fidelity condition, we assign high values to the overall emission probability ($c = 0.90$, $d = 0.95$).

There are four different combinations of persistence and fidelity levels, which are used as four experimental conditions. For each condition, we generate 50 synthetic datasets and apply our model to recover the parameters. To approximate the dimensionality of the actual data used in our empirical analysis, we generate $n = 1000$ hypothetical firms in each data, and the number of signals per firm T_i is randomly drawn from integers between 20 and 100.

To evaluate the performance of our model, we compute the following measures: (i) an estimate of the overall transition probability $\widehat{a}$, $\widehat{b}$; (ii) an estimate of the overall emission probability $\widehat{c}$, $\widehat{d}$; (iii) the mean absolute error (*MAE*) of the firm-level transition probabilities a_i and b_i; and, (iv) the mean absolute error of the firm-level emission probabilities c_i and d_i.

Panel A of Table A.1 presents the means and standard deviations of the four measures across all 50 synthetic datasets in each condition. Overall, the result shows that the MCMC algorithm estimates both aggregate-level parameters and firm-level parameters well. Further, identification of the model does not depend on covariates. It appears that the accuracy of parameter estimation increases when either persistence or fidelity is high.

B.2 With covariates

In the second simulation study, we add covariates into both the transition and emission functions when generating the data and examine how well our MCMC algorithm can recover the main parameters. Because the model implicitly assumes the

[26] In other words, the cross-sectional mean of the parameters is given, but each firm may have a different parameter.

initial state distribution $\boldsymbol{\pi}_0$ is the same across all firms, we also examine the impact of this assumption on the estimation results in this study.

Similarly to the previous simulation setup, we generate data for $n = 1{,}000$ hypothetical firms, and the number of signals per firm T_i is randomly drawn from integers between 20 and 100. We set the means of firm-level intercepts as $\overline{\boldsymbol{\alpha}}_0 = (-0.75, 0.75)$ and $\overline{\boldsymbol{\beta}}_0 = (-1, \log(2))$ and fix all heterogeneity parameters at 0.1. Firm-level intercepts are then independently generated from the assumed normal distributions. We generate five covariates in the transition function and five covariates in the emission function, and use them to generate the observed earnings signals. For each generated data set, we randomly and independently generate each regression coefficient ($\boldsymbol{\alpha}_s$ or $\boldsymbol{\beta}_s$) from a standard normal distribution.

We consider two experimental conditions related to the initial state distribution. In the first condition, we assume the true initial state distribution used to generate the data is drawn from a Dirichlet distribution, and all firms have the same initial state distribution: $\boldsymbol{\pi}_{0i} = \boldsymbol{\pi}_0 \sim \text{Dirichlet}(1, 1)$. In the second condition, we relax the assumption of identical initial state distribution, and assume that the true initial state distribution for each firm is randomly generated from a Dirichlet distribution $\boldsymbol{\pi}_{0i} \sim \text{Dirichlet}(1, 1)$. Recall that our model assumes that the initial state distribution is the same across all firms. Thus, our model only matches the true data generating process in the first condition but not the second. This allows us to examine any impact of the initial state distribution assumption on estimation results. For each experimental condition, we generate 100 synthetic datasets.

To compare how well our MCMC algorithm recovers the regression coefficients, we compute the mean absolute error (*MAE*) of the regression coefficient estimates $\widehat{\boldsymbol{\alpha}}$ and $\widehat{\boldsymbol{\beta}}$, where

$$MAE(\boldsymbol{\alpha}) = \sum_{s\in\{L,H\}} \sum_{p=1}^{5} \left|\alpha_{sp} - \widehat{\alpha}_{sp}\right|/10, \quad MAE(\boldsymbol{\beta}) = \sum_{s\in\{L,H\}} \sum_{l=1}^{5} \left|\beta_{sl} - \widehat{\beta}_{sl}\right|/10. \tag{A.34}$$

As reported in Panel B of Table A.1, across 100 randomly-generated synthetic datasets with identical initial state distribution, the average $MAE(\boldsymbol{\alpha})$ is 0.024 with a standard deviation of 0.012; the average $MAE(\boldsymbol{\beta})$ is 0.016, with a standard deviation of 0.007. Across 100 randomly-generated synthetic datasets with varying initial state distribution, the average $MAE(\boldsymbol{\alpha})$ is 0.023, with a standard deviation of 0.010; the average $MAE(\boldsymbol{\beta})$ is 0.016, with a standard deviation of 0.006. A two-sample t-test suggests that there are no statistically significant differences in $MAE(\boldsymbol{\alpha})$ ($MAE(\boldsymbol{\beta})$) between the two conditions.

We also compare the *MAE* of the firm-level transition probabilities a_{it} and b_{it} and the *MAE* of the firm-level emission probabilities c_{it} and d_{it} in Table A.1, Panel B. Similarly, a series of two-sample t-tests suggests that there is no evidence that any of the four *MAE* measures is statistically different between the two conditions (p-value > 0.10 for the difference in *MAE* of all parameters).

Overall, the results of the simulation analysis suggest: (i) our algorithm performs well in recovering both regression coefficients and firm-level parameters; (ii) the impact of identical initial state distribution on model estimation is not statistically significant.

Table A.1
Simulation analysis

Experimental Condition	$\widehat{a}$	$\widehat{b}$	$\widehat{c}$	$\widehat{d}$	$MAE(a_i)$	$MAE(b_i)$	$MAE(c_i)$	$MAE(d_i)$
Panel A: No covariates: Parameter estimates								
$(a,b,c,d) = (0.65, 0.70, 0.75, 0.80)$	0.620	0.721	0.773	0.787	0.090	0.085	0.079	0.068
	(0.041)	(0.042)	(0.036)	(0.035)	(0.012)	(0.013)	(0.011)	(0.007)
$(a,b,c,d) = (0.65, 0.70, 0.90, 0.95)$	0.644	0.700	0.902	0.956	0.067	0.062	0.044	0.031
	(0.014)	(0.013)	(0.016)	(0.015)	(0.002)	(0.002)	(0.004)	(0.005)
$(a,b,c,d) = (0.80, 0.85, 0.75, 0.80)$	0.768	0.857	0.780	0.784	0.069	0.052	0.077	0.059
	(0.017)	(0.012)	(0.015)	(0.014)	(0.007)	(0.002)	(0.006)	(0.002)
$(a,b,c,d) = (0.80, 0.85, 0.90, 0.95)$	0.794	0.848	0.903	0.952	0.054	0.044	0.041	0.026
	(0.006)	(0.005)	(0.005)	(0.004)	(0.002)	(0.001)	(0.002)	(0.001)

	$MAE(\boldsymbol{\alpha})$	$MAE(\boldsymbol{\beta})$	$MAE(a_{it})$	$MAE(b_{it})$	$MAE(c_{it})$	$MAE(d_{it})$
Panel B: With covariates: Mean absolute errors						
$\boldsymbol{\pi}_0 \sim \text{Dirichlet}(1,1)$	0.024	0.016	0.065	0.064	0.083	0.084
	(0.012)	(0.007)	(0.016)	(0.016)	(0.018)	(0.020)
$\boldsymbol{\pi}_{i0} \sim \text{Dirichlet}(1,1)$	0.023	0.016	0.065	0.068	0.086	0.083
	(0.010)	(0.006)	(0.016)	(0.018)	(0.021)	(0.019)
p-value	0.312	0.541	0.822	0.108	0.278	0.854

Panel A reports the parameter estimates and mean absolute errors for the simulation study with no covariates. Panel B reports the mean absolute errors for the simulation study with covariates. Standard deviations are reported in parentheses.

Appendix C: Proof of Eqs. (15) and (16)

Assume that firm i's states are governed by a homogeneous Markov chain with transition matrix $[a_i, 1-a_i; 1-b_i, b_i]$. Assume the underlying states are at the stationary distribution.

$$\Pr(x_{it}=L)=\frac{1-b_i}{2-a_i-b_i}, \quad \Pr(x_{it}=H)=\frac{1-a_i}{2-a_i-b_i}. \tag{A.35}$$

From the definitions of signal persistence given in Eq. (13), we have:

$$\begin{aligned}
\rho_{il} &= \Pr(y_{it}=l|y_{it-1}=l)=\frac{\Pr(y_{it}=l, y_{it-1}=l)}{\Pr(y_{it-1}=l)}\\
&=\frac{\sum_{x=L}^{H}\Pr(y_{it}=l, y_{it-1}=l|x_{it-1}=x)P(x_{it-1}=x)}{\sum_{x=L}^{H}\Pr(y_{it-1}=l|x_{it-1}=x)\Pr(x_{it-1}=x)}\\
&=\frac{\sum_{x=L}^{H}\left[\sum_{x^*=L}^{H}\Pr(y_{it}=l|x_{it}=x^*)\Pr(x_{it}=x^*|x_{it-1}=x)\right]\Pr(y_{it-1}=l|x_{it-1}=x)\Pr(x_{it-1}=x)}{\sum_{x=L}^{H}\Pr(y_{it-1}=l|x_{it-1}=x)\Pr(x_{it-1}=x)}\\
&=\frac{(1-b_i)c_i[a_ic_i+(1-a_i)(1-d_i)]+(1-a_i)(1-d_i)[(1-b_i)c_i+b_i(1-d_i)]}{(1-b_i)c_i+(1-a_i)(1-d_i)},
\end{aligned} \tag{A.36}$$

and

$$\begin{aligned}
\rho_{ih} &= \Pr(y_{it}=h|y_{it-1}=h)=\frac{\Pr(y_{it}=h, y_{it-1}=h)}{\Pr(y_{it-1}=h)}\\
&=\frac{\sum_{x=L}^{H}\Pr(y_{it}=h, y_{it-1}=h|x_{it-1}=x)\Pr(x_{it-1}=x)}{\sum_{x=L}^{H}\Pr(y_{it-1}=h|x_{it-1}=x)\Pr(x_{it-1}=x)}\\
&=\frac{\sum_{x=L}^{H}\left[\sum_{x^*=L}^{H}\Pr(y_{it}=h|x_{it}=x^*)\Pr(x_{it}=x^*|x_{it-1}=x)\right]\Pr(y_{it-1}=h|x_{it-1}=x)\Pr(x_{it-1}=x)}{\sum_{x=L}^{H}\Pr(y_{it-1}=h|x_{it-1}=x)\Pr(x_{it-1}=x)}\\
&=\frac{(1-b_i)(1-c_i)[a_i(1-c_i)+(1-a_i)d_i]+(1-a_i)d_i[b_id_i+(1-b_i)(1-c_i)]}{(1-b_i)(1-c_i)+(1-a_i)d_i}.
\end{aligned} \tag{A.37}$$

Appendix D: Variable definitions

Variable	Definition
AQ	Accruals quality. For firm j in year t, *AQ* is the standard deviation of abnormal accruals over year $t-4$ through year t, requires at least three years of data. Abnormal accruals are measured as the residuals from the following cross-sectional regression: $TCA_{jt}=\phi_0+\phi_1 CFO_{jt-1}+\phi_2 CFO_{jt}+\phi_3 CFO_{jt+1}+\phi_4 \Delta Rev_{jt}+\phi_5 PPE_{jt}+\varepsilon_{jt}$ where $TCA=\Delta CA-\Delta CL-\Delta Cash+\Delta STD$ is the total current accruals, $CFO=NI-TCA+Dep$ is the cash flow from operations, *NI* is the net income before extraordinary items (Compustat data item *ib*), ΔRev is the change in revenue (*sale*), and *PPE* is the gross value of property, plant, and equipment (*ppegt*). All variables are scaled by average total assets (*at*). We estimate annual cross-sectional regressions for each of the Fama-French 48 industries (excluding four financial industries) with at least 20 firms in year t.
$\Delta Csale$	The percentage change in cash sales where cash sales are calculated as sales minus change in receivables.
ΔEmp	The abnormal change in employees, calculated as the percentage change in the number of employees minus the percentage change in total assets.

(*continued on next page*)

(*continued*)

Variable	Definition
ΔInv	Change in inventory scaled by lagged total assets.
ΔNI	Change in net income before extraordinary items divided by lagged total assets.
ΔRec	Change in receivables scaled by lagged total assets.
Comment	An indicator variable that equals 1 if at least one of a firm's 10-Q and 10-K filings during the fiscal year is the subject or one of the subjects of an SEC comment letter, and 0 otherwise, based on data from Audit Analytics. We merge comment letters with Compustat by requiring the filing date (Audit Analytics data item *list_file_date_ed*) of the form commented by an SEC comment letter to fall within a 365-day interval ending 100 days after the fiscal year end. This interval approximately captures all the filing dates of all 10-K and 10-Q filings that pertain to a fiscal year. A comment letter enters our sample as long as it is related to at least one 10-K or 10-Q.
Cvol	Cash flow volatility, calculated as the standard deviation over the five-year window of $t-4$ through t (with a minimum of three years) of cash flow from operations (*CFO*) scaled by lagged total assets. Cash flow from operations is calculated based on the formula provided in the definition of *AQ*.
E_{t-1}, E_t, and E_{t+1}	Net income for fiscal years $t-1$, t, and $t+1$, respectively, all scaled by the market value of equity at three months after the beginning of fiscal year t.
Fidelity	Earnings fidelity, calculated as $(c_{it}+d_{it})/2$, i.e., the average of the probability of issuing signal l conditional on state L and the probability of issuing signal h conditional on state H.
$Fidelity^p$	A measure of earnings fidelity based on a rolling estimation. For each year, parameters are estimated based on all available data up to the previous year.
$Fidelity^w$	The weighted fidelity measure, calculated for firm i at time t as $Fidelity^w_{it}=p_{itL}c_{it}+p_{itH}d_{it}$, where p_{its} is the probability that firm i's state at t is $s\in\{L,H\}$ after observing the earnings history.
Lev	Total debts divided by total assets at year end.
MB	Market value of equity divided by book value of equity at year end.
Pers	Economic persistence, calculated as $(a_{it}+b_{it})/2$, i.e., the average of the state-dependent transition probabilities.
ROA	Income before extraordinary items divided by average total assets over the fiscal year.
Restatement	An indicator variable that equals 1 for a firm-year of which the financial statements are subsequently restated for accounting issues, and 0 otherwise, based on data from Audit Analytics.
R	The stock return over the 12 months beginning 9 months prior to the end of fiscal year t.
Size	The natural logarithm of market value of equity at year end.
Smoothness	Earnings smoothness, calculated as the ratio of firm i's standard deviation of earnings before extraordinary items (scaled by average total assets) to the standard deviation of cash flows from operations (scaled by average total assets), where standard deviations are calculated based on the five-year window of $t-4$ through t.
UAF	Unexplained audit fees, calculated as the residual term from the following regression, estimated by fiscal year and size decile for the period of 1999–2015: $\log(Fee)_{jt}=\rho_0+\rho\times Determinants_{jt}+Industry\ \ FE+\varepsilon_{jt}$, where $\log(Fee)$ is the natural log of audit fees. The following variables are included as determinants of audit fees: *Big4*, an indicator variable that equals 1 if the current auditor is a Big-4 accounting firm and 0 otherwise; *LnTA*, the natural logarithm of total assets; *#Segments*, the square root of the number of the business segments of the firm; *Foreign*, an indicator variable that equals 1 if the firm pays any foreign income tax, 0 otherwise; *InvRec*, the inventory and receivables divided by total assets; *Current*, the current ratio, calculated as current assets divided by current liabilities; *BM*, the book value of equity divided by market value of equity; *Lev*, the sum of short-term debt and long-term debt scaled by total assets; *Employ*, the square root of the number of employees; *Acquire*, an indicator variable that equals 1 if the dollar amount of acquisition exceeds 5% of lagged total assets; *Dec_YE*, an indicator variable that equals 1 if the fiscal year-end is not December 31, and 0 otherwise; *ROA*, income before extraordinary items divided by average total assets; *Loss*, an indicator variable that equals 1 if income before extraordinary items is negative in the current or two previous years, and 0 otherwise; *Auditor_Opinion*, an indicator variable that equals 1 if the firm receives any audit opinion other than a standard unqualified opinion, and 0 otherwise; *Auditor_Change*, an indicator variable that equals 1 if there is an auditor change during the fiscal year, and 0 otherwise; *Issue*, an indicator variable that equals 1 if the sum of debt or equity issued in the current and two previous years is more than 5% of the total assets, 0 otherwise. Industry is defined by two-digit SIC code. Continuous variables are winsorized at 1% and 99%.

References

Albert, J., Chib, S., 1993. Bayesian analysis of binary and polychotomous response data. J. Am. Stat. Assoc. 88, 669–679.

Antle, R., Lambert, R., 1988. Accountants' loss functions and induced incentives for conservatism. In: Economic Analysis of Information and Contracts: Essays in Honor of John Butterworth. Edited by G. Feltham, A. Amershi, and W. Ziemba. Kluwer Academic Publishers, Boston, MA, pp. 373–408.

Ball, R., 2013. Accounting informs investors and earnings management is rife: two questionable beliefs. Account. Horiz. 27, 847–853.

Beneish, M., 1999. Incentives and penalties related to earnings overstatements that violate GAAP. Account. Rev. 74, 425–457.

Berger, J., 1985. Statistical Decision Theory and Bayesian Analysis. Springer-Verlag, New York.

Bernhardt, D., Wan, C., Xiao, Z., 2016. The reluctant analyst. J. Account. Res. 54, 987–1040.

Bertomeu, J., Ma, P., Marinovic, I., 2018. How Often Do Managers Withhold Information? University of California at San Diego. Working paper.

Blackwell, D., 1951. Comparison of experiments. In: Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, pp. 93–102.

Chib, S., 1996. Calculating posterior distributions and modal estimates in Markov mixture models. J. Econom. 75, 79–97.

Christensen, P., Feltham, G., 2003. Economics of Accounting – Volume I, Information in Markets. Springer Series in Accounting Scholarship. Springer, Heidelberg and New York.

Collins, D., Kothari, S.P., Shanken, J., Sloan, R., 1994. Lack of timeliness and noise as explanations for the low contemporaneous return-earnings association. J. Account. Econ. 18, 289–324.

Dechow, P., Dichev, I., 2002. The quality of accruals and earnings: the role of accrual estimation errors. Account. Rev. 77, 35–59.

Dechow, P., Ge, W., Larson, C., Sloan, S., 2011. Predicting material accounting misstatements. Contemp. Account. Res. 28, 17–82.

Dechow, P., Ge, W., Schrand, C., 2010. Understanding earnings quality: a review of the proxies, their determinants and their consequences. J. Account. Econ. 50, 344–401.
Dechow, P., Hutton, A., Kim, J., Sloan, R., 2012. Detecting earnings management: a new approach. J. Account. Res. 50, 275–334.
Dechow, P., Kothari, S.P., Watts, R., 1998. The relation between earnings and cash flows. J. Account. Econ. 25, 133–168.
Dichev, I., Graham, J., Harvey, C., Rajgopal, S., 2013. Earnings quality: evidence from the field. J. Account. Econ. 56, 1–33.
Du, K., 2019. Investor expectations, earnings management, and asset prices. J. Econ. Dyn. Control 105, 134–157.
Du, K., Huddart, S., 2019. Economic persistence, earnings informativeness, and stock return regularities. Rev. Account. Stud. In press.
Durbin, R., Eddy, S., Krogh, A., Mitchison, G., 1998. Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids. Cambridge University Press.
Dye, R., 1985. Disclosure of nonproprietary information. J. Account. Res. 23, 123–145.
Ettredge, M., Kwon, S., Smith, D., Zarowin, P., 2005. The impact of SFAS No. 131 business segment data on the market's ability to anticipate future earnings. Account. Rev. 80, 773–804.
Fairfield, P., Ramnath, S., Yohn, T., 2009. Do industry-level analyses improve forecasts of financial performance? J. Account. Res. 47, 147–178.
Fama, E., French, K., 2000. Forecasting profitability and earnings. J. Bus. 73, 161–175.
Gao, P., 2013. A measurement approach to conservatism and earnings management. J. Account. Econ. 55, 251–268.
Gelman, A., 2006. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). Bayesian Anal. 1, 515–534.
Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., Rubin, D., 2013. Bayesian Data Analysis, 3rd ed. Chapman and Hall/CRC, Boca Raton, FL.
Gray, S., 1996. Modeling the conditional distribution of interest rates as a regime-switching process. J. Financ. Econ. 42, 27–62.
Greenland, S., 2000. Principles of multilevel modelling. Int. J. Epidemiol. 29, 158–167.
Guay, W., Kothari, S.P., Watts, R., 1996. Market-based evaluation of discretionary accrual models. J. Account. Res. 34, 83–105.
Hamilton, J., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57, 357–384.
Hastings, W., 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57, 97–109.
Hayn, C., 1995. The information content of losses. J. Account. Econ. 20, 125–153.
Hribar, P., Kravet, T., Wilson, R., 2014. A new measure of accounting quality. Rev. Account. Stud. 19, 506–538.
Karpoff, J., Koester, A., Lee, D., Martin, G., 2017. Proxies and databases in financial misconduct research. Account. Rev. 92, 129–163.
Khan, M., Watts, R., 2009. Estimation and empirical properties of a firm-year measure of accounting conservatism. J. Account. Econ. 48, 132–150.
Lev, B., 1983. Some economic determinants of time-series properties of earnings. J. Account. Econ. 5, 31–48.
Lipe, R., Bryant, L., Widener, S., 1998. Do nonlinearity, firm-specific coefficients, and losses represent distinct factors in the relation between stock returns and accounting earnings? J. Account. Econ. 25, 195–214.
Lundholm, R., Myers, L., 2002. Bringing the future forward: the effect of disclosure on the returns-earnings relation. J. Account. Res. 40, 809–839.
MacDonald, I., Zucchini, W., 1997. Hidden Markov and Other Models for Discrete-Valued Time Series. Chapman & Hall, London.
McNichols, M., 2002. Discussion of the quality of accruals and earnings: the role of accrual estimation errors. Account. Rev. 77, 61–69.
Netzer, O., Lattin, J., Srinivasan, V., 2008. A hidden Markov model of customer relationship dynamics. Mark. Sci. 27, 185–204.
Nikolaev, V., 2017. Identifying Accounting quality. University of Chicago. Working Paper.
Owens, E., Wu, J., Zimmerman, J., 2017. Idiosyncratic shocks to firm underlying economics and abnormal accruals. Account. Rev. 92, 183–219.
Rabiner, L., Juang, B., 1986. An introduction to hidden Markov models. IEEE ASSP Mag. 3, 4–16.
Robert, C., Casella, G., 2013. Monte Carlo Statistical Methods. Springer, New York.
Rossi, P., Allenby, G., McCulloch, R., 2005. Bayesian Statistics and Marketing. John Wiley & Sons, New York.
Scott, S., 2002. Bayesian methods for hidden Markov models: recursive computing in the 21st century. J. Am. Stat. Assoc. 97, 337–351.
Verrecchia, R., 2001. Essays on disclosure. J. Account. Econ. 32, 97–180.
Vuong, Q., 1989. Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 57, 307–333.
Zakolyukina, A., 2018. How common are intentional GAAP violations? Estimates from a dynamic model. J. Account. Res. 56, 5–44.
Zhou, F., 2017. Disclosure dynamics and investor learning. University of Pennsylvania. Working Paper.