Understanding the "Numbers Game"

Andrew Bird, Stephen A. Karolyi, and Thomas G. Ruchti

Tepper School of Business
Carnegie Mellon University

July 15, 2019

Abstract

Two well-known stylized facts on earnings management are that the earnings surprise distribution has a discontinuity at zero, and that positive earnings surprises are associated with positive abnormal returns. We link these two facts in a model of the earnings management decision in which the manager trades off the capital market benefits of meeting earnings benchmarks against the costs of manipulation. We develop a new structural methodology to estimate the model and uncover the unobserved cost function. The estimated model parameters yield the percentage of manipulating firms, magnitude of manipulation, noise in manipulation, and sufficient statistics to evaluate proxies for identifying firms suspected of manipulation. Finally, we use the Sarbanes–Oxley Act as a policy experiment and find that by increasing costs, the Act reduced equilibrium earnings management by 36%. This reduction occurred despite an increase in benefits, consistent with the market rationally becoming less skeptical of firms that just meet benchmarks.

We thank David Aboody, Cagla Akin, Brian Akins, Phil Berger, Pietro Bonaldi, Jonathan Bonham, Brian Bushee, Judson Caskey, Nicholas Chen, Si Cheng, John Core (editor), Alan Crane, Kevin Crotty, David De Angelis, Karl Diether (discussant), Aytekin Ertan, Paul Fischer, Henry Friedman, John Gallemore, Pingyang Gao, Zhenyu Gao, Joseph Gerakos (discussant, referee), Matthew Gustafson, Luzi Hail, Sam Hartzmark (discussant), Mirko Heinle, Burton Hollifield, Bob Holthausen, Jack Hughes, Peter Iliev, Andy Koch, Kalin Kolev, Jason Kotter, Rick Lambert, Alina Lerman, Christian Leuz, Guoyu Lin, Valeri Nikolaev, Marios Panayides, K. Ramesh, Haresh Sapra, Shiva Sivaramakrishnan, Chester Spatt, Shyam Sunder, Chris Telmer, Jake Thomas, Shawn Thomas, Phong Truong, Yizhou Xiao, Anastasia Zakolyukina, Frank Zhang, and Frank Zhou (discussant) for helpful comments and discussion. We also thank seminar participants at the University of Chicago, UCLA, and Yale University, and conference participants at the JAE Conference, FARS Midyear Meetings, the CMU-Pitt-PSU Finance Conference, the Computing in Economics and Finance Annual Conference, the Northern Finance Association Conference, Stanford Summer Camp, and the University of Washington Summer Finance Conference.

1 Introduction

We develop and estimate a structural model that connects two well-known stylized facts from the earnings management literature. The first is that the earnings surprise distribution is characterized by a jump at zero. The extant literature typically interprets this jump as evidence that managers manage earnings to meet short-term performance benchmarks (Brown and Caylor [2005]; Dechow et al. [2003]; Degeorge et al. [1999]; Burgstahler and Dichev [1997]; Hayn [1995]). The second fact is that meeting short-term earnings benchmarks is associated with positive abnormal returns (Payne and Thomas [2011]; Zang [2011]; Keung et al. [2010]; Cohen et al. [2010]; Caylor [2010]; Bhojraj et al. [2009]; Roychowdhury [2006]; Graham et al. [2005]; Jensen [2005]; Kasznik and McNichols [2002]). At the discontinuity between meeting and missing the market's expectations, we find that just meeting the benchmark is associated with 1.45 percentage points higher cumulative market-adjusted returns. Using this estimate of benefits and the distribution of actual minus forecasted earnings, our structural model recovers a marginal cost of manipulation for the median firm of 1.6% of market value and implies that 2.6% of firms manipulate.[1] Conditional on the choice to manipulate at all, we find that the average manipulation is 1.21 cents per share.

Our structural approach links manipulation behavior to capital market benefits using an economic model of the tradeoff of these observable benefits against some unobservable costs of manipulation. These costs may take many forms, such as litigation costs, reputational costs for the manager, or even the costly effort required to devise and execute the necessary transactions. This economic model incorporates four parameters in the manipulation cost function: the marginal cost of earnings manipulation, the curvature of the manipulation cost function, and two parameters that capture uncertainty in manipulation. This cost function allows for variation in the cost of a single cent of manipulation, variation in the increase in costs for additional earnings management, variation in the degree to which earnings manipu-

[1] We study manipulation that is motivated by meeting the market's expectations, but cannot rule out the presence of manipulation for other purposes. Therefore our estimate of 2.6% should not be interpreted as the unconditional level of manipulation in the economy.

lation is noisy (if it is noisy at all), and variation in how that noise increases with increasing manipulation.

We estimate these parameters using the simulated method of moments in a four-step procedure.[2] First, we choose candidate parameters for the cost function. Second, given these candidate parameters, we model optimal manipulation behavior. Third, starting with the empirical earnings surprise distribution, we invert optimal manipulation behavior to obtain a candidate latent distribution of unmanipulated earnings. Fourth, among all candidate parameters and latent distributions, we identify our parameter estimates using the particular candidate latent distribution that is the smoothest curve approximating the empirical earnings surprise distribution.

Our approach uses statistical and economic assumptions to produce inferences about managers' earnings management decisions. For example, our economic model studies a manager's use of costly manipulation to achieve capital market benefits, but it is agnostic about the specific manipulation tools that are used. Firms can use a variety of tools to meet earnings benchmarks. In line with the substitutability of these tools (Zang [2011]), our modeling approach is consistent with managers using a portfolio of manipulation tools, starting with the least costly. This would naturally lead to a convex cost function.

Our methodology is flexible in modeling choices, but it requires some important assumptions for identification. Most importantly, we must make assumptions about characteristics of the latent distribution of earnings. Rather than making parametric assumptions, we evaluate candidate latent distributions according to a criterion function based on similarity to the empirical distribution and smoothness. This modeling approach is consistent with a world in which the firm's technology produces a smooth distribution of earnings, and does not impose any particular structure on how market expectations are formed. Further, the properties of the latent distribution implied by our criterion function fit the suggestive empirical evidence presented in Donelson et al. [2013], which investigates the properties of earnings

[2] The code used to construct the inputs and estimate the model is posted at `http://www.thomasruchti.com/NumbersGameCode190530.zip`.

after removing the effects of restatements related to securities litigation.[3]

In the earnings management context, our approach can accommodate different models of manipulation because we do not require any specific functional forms for identification. One could take alternative approaches to assessing candidate latent distributions. For example, parametric restrictions on behavior, such as normality, could also be used to achieve identification using our approach. Such parametric assumptions may be appropriate in settings where researchers have strong priors about the underlying functional forms involved. One benefit of our more agnostic approach is its ability to accommodate alternative statistical assumptions and to embed context-specific economic models. This makes our approach well suited for studying other settings in which only the distribution of equilibrium outcomes are observed, along with either the benefits or the costs of strategic actions.

Our economic model abstracts from several potentially interesting aspects of reality. Most importantly, we model a static tradeoff that does not explicitly incorporate dynamics. This limitation, combined with the positive market response to firms that just meet the market's expectations, means that managers are incentivized only to manipulate upward. However, our choice to use a static model does not restrict managers from having dynamic considerations. For example, managers who are more forward looking and who consider the future costs of present manipulation will engage in less manipulation. In our model, the choice to avoid manipulation is implicitly reflected in higher manipulation costs. Another simplifying assumption in our baseline model is that the benefit of meeting the market's expectations is the same for all firms. However, we extend our analysis to allow for variation in the expected benefit across firms, and we also allow for a firm-level correlation between costs and expected benefits that is consistent with investors imperfectly observing manipulation. We find similar results in both extensions.

To demonstrate potential applications of our methodology, we investigate three empirical questions relevant to our approach. Our first application concerns the significant fraction

[3]Specifically, when the effects of restatements are removed, Figure 1 in Donelson et al. [2013] shows that the distribution of earnings surprises indeed becomes smoother, while remaining similar in shape.

of firms, 6.1%, that just miss their earnings benchmarks. Counterfactual simulations of our model suggest a potential explanation for this fact. In the model, this behavior is driven by the estimates of the marginal cost and noise parameters of the cost function. Changing the marginal cost parameter has straightforward consequences: as the marginal cost of manipulation increases, the optimal strategy of managers shifts toward avoiding manipulation.

Increasing the variance of noise changes optimal manipulation in two ways, both of which help explain this puzzling empirical fact. First, the optimal strategy for managers with low costs who expect to significantly miss their benchmark may be to do some manipulation to get closer to the benchmark. They do so in the hope of getting a positive shock large enough that they end up meeting the benchmark. Increasing the variance of noise increases the expected payoff of this strategy, thus resulting in more firms narrowly missing their benchmarks. Second, firms that aim to just meet their benchmarks and then receive negative shocks will inadvertently end up just missing. Again, increasing the variance of noise, by increasing the probability of receiving a shock, makes this more likely.

We also investigate the literature's workhorse empirical proxies for earnings management. Typically, these proxies depend on the relative number of "just-meet" firms and "just-miss" firms (e.g., Bhojraj et al. [2009]). From our structural estimates, we uncover the proportion of firms in each cent bin that are manipulators, providing a continuous proxy for suspect firms. This distribution of manipulation naturally produces a way to evaluate the commonly used "suspect bin" empirical proxies for earnings management, based on type I and type II errors.

Finally, passage of the Sarbanes–Oxley Act (SOX) provides an important shift in regulation and attention paid to accounting information. Whether or not this increased attention led to higher costs of earnings management remains unclear (Cohen et al. [2008]; Bartov and Cohen [2009]). We estimate our model in the pre-SOX (1999–2001) and post-SOX (2002–2004) periods, and we compare these estimates to uncover the effects of the regulation. We find that the regression discontinuity estimates of the equity return benefits to just meeting or beating earnings increased between these periods, but the marginal costs of manipulation (and the costs of incremental earnings management in particular) increased even more. This cost increase had a major impact on the incentives to manage earnings: the fraction of manipulating firms decreased by 36% following SOX.

The approach we take to study managerial manipulation is similar in spirit to two contemporaneous papers: Zakolyukina [2018] and Beyer et al. [2018], though our focus is different. Zakolyukina [2018] uses observed violations of GAAP by managers to study manipulation driven by the possibility that the manager leaves the firm while it is mispriced. In contrast, we infer benchmark beating behavior from observed capital market benefits. Relative to our "direct" approach to benefits, both Zakolyukina [2018] and Beyer et al. [2018] make additional assumptions about investor behavior and the earnings process in order to estimate the benefits to the manager of manipulating. Beyer et al. [2018] use this additional structure to study the consequences of manipulation and noise on the dynamic characteristics of earnings, such as persistence. While their approach to noise is different from ours and also includes noise in fundamentals, they also find that reporting noise plays an important role.

The rest of the paper proceeds as follows: Section 2 presents our economic model of earnings management. Section 3 outlines our estimation inputs, procedure, and baseline estimates. Section 4 describes identification using counterfactuals and provides robustness tests. Section 5 shows three applications of the methodology. Section 6 concludes.

2 The Model

In this section, we develop a simple model of the manager's earnings manipulation decision. Our objective is to explicitly determine how optimal behavior is affected by features of the economic environment, principally related to the benefits of manipulation and the manipulation technology. The model setup does not impose exactly how earnings management works;

instead, the model allows for possibilities. For example, if the economic environment includes convex costs, then the model informs us about how convex costs affect behavior. However, the data and the estimation alone must determine whether that feature is consistent with reality.

In the model, the manager receives an interim signal about earnings, e, with respect to the market's expectation; we call this the *interim* or *latent* earnings surprise. For example, $e = 0$ means that the firm would just meet the market's expectation if the interim number were to be reported. After observing e, the manager can choose whether to manipulate earnings. Note that this setup means that at the time the manager makes her manipulation decision, there is no remaining uncertainty in e. The manager's choice is determined by trading off the benefit of potentially increasing the value of the firm against the costs of manipulation. The result of earnings manipulation leads to the following report:

$$\text{Reported Earnings Surprise} = \text{Latent Earnings Surprise} + \text{Manipulation} + \text{Noise}$$

—or—

$$R = e + m + \varepsilon, \tag{1}$$

where R is the reported earnings surprise, or the difference between reported earnings and the market's expectation; m is the desired manipulation; and ε is a noise term, which we discuss further below.

We assume that e, m, and R are discrete; and, specifically, we assume that they are integer multiples of a cent per share, consistent with the literature. Further, in Appendix A, we find that within-cent differences in reporting earnings do not seem to matter. For example, there does not appear to be a benefit in reporting 1.1 cents versus 1 cent. In principle, instead of assuming discreteness in the components of R, we could allow its components to be continuous and instead round their sum. This choice does not have a first-order impact on our inferences, because we target the aggregate distribution of R and because neither the starting point (e) nor manipulation (m) are observable.

The effect of the alternative assumption would be to add smoothness to the problem faced by the optimizing manager through, for example, a continuous distribution of starting points. As described below, the way that we allow for some smoothness in this optimization is by incorporating heterogeneity in costs of manipulating, which has an economic effect similar to heterogeneity in starting points. One potential microfoundation for such heterogeneity comes from the way that reported earnings surprise is rounded. For example, a manager with an underlying earnings surprise of $0.0049 will likely find it much less costly to manipulate reported earnings surprise upward by a cent than will a manager with an underlying earnings surprise of -$0.0049. This approach allows for some smoothness, or variation, in firm behavior for a given starting point, in a way that is computationally simpler than modeling a continuous distribution of starting points.

Because the capital market observes reported earnings, the benefit is $\mathcal{B}(R)$, where $\mathcal{B}(\cdot)$ is the return for a particular level of reported earnings relative to the market's expectation. Since $\mathcal{B}(\cdot)$ is likely to be at least weakly increasing in earnings surprise, the manager has an incentive to manipulate upward in our static model. This is a limitation in our setting, relative to a dynamic model in which managers may choose to trade off the present costs of missing the market's expectation with the future benefit of meeting it.

We assume that the manager takes the benefit function as given. This assumption is consistent with the relative lack of evidence in the literature that the existence of a benefit to earnings management depends on how those reported earnings were achieved (Bartov et al. [2002]; Bhojraj et al. [2009]). We take the perspective that the problem facing the market in valuation is assessing the *aggregate* likelihood of manipulation for each level of earnings surprise. It is then an empirical question whether and how much firms benefit from just meeting the market's expectation. Presumably, the market trades off a desire to reward fundamental performance against the cost of rewarding manipulated performance, although this approach is consistent with a variety of models of investor behavior, including a naive reward for reported performance or even a reward for the sophistication and ability to manipulate earnings successfully. A key implication of this idea is that no individual firm can affect this aggregate tradeoff and so would rationally take the observed benefit as given.[4] For this reason, only the equilibrium reward offered by the market matters in our empirical setting, rather than the particular theoretical model that generates this market equilibrium.

The cost of manipulation is

$$c(m) = \beta\, m^{\gamma}, \tag{2}$$

where m is the desired manipulation, $\beta \sim U[0, 2\eta]$ is the linear cost of manipulation, and γ is the exponential curvature of the cost of manipulation function. Note that the variation in β means that there will be heterogeneity in optimal manipulation for firms that have the same discrete e.

These costs could result from manipulating accounting information, altering real operating activities, or inducing bias in analyst forecasts.[5] To make estimation feasible, we abstract from the manager's choice between these three tools to meet benchmarks; we do not attempt to explicitly model substitution among these tools. Rather, we take a reduced-form approach consistent with managers using a portfolio of manipulation tools, starting with the least costly, naturally leading to a convex cost function. The main benefit of this simplification is that we do not need to observe the full set of particular earnings management strategies in use. This simplification also means that the costs of all potential strategies are valued on the same scale. The model is static, in the sense that we do not try to explicitly capture dynamic earnings management tradeoffs (e.g., through accrual reversals). However,

[4] A further implication of this framework is that ex ante uncertainty about benefits would not affect the tradeoff faced by a risk-neutral manager. We explore ex post uncertainty in benefits and arbitrarily correlated benefits and costs in Section 4.3.

[5] The simplest interpretation of our model is that managers take the analyst benchmark as given. However, our setup naturally captures potential manipulation of the benchmark, such as through walkdown (Kross et al. [2011]; Chan et al. [2007]; Burgstahler and Eames [2006]; Cotter et al. [2006]), because the manager's payoff depends only on the difference between reported earnings and the market's expectation. Thus, one can equally well think of m as indicating an increase in earnings or a decrease in the benchmark, as long as the cost of both options can be described by $c(\cdot)$.

future costs are implicitly included in our cost function, to the extent that managers are actually forward looking and consider the future costs or constraints imposed by present manipulation.

An important feature of our model setup is that we allow for the possibility of a difference between the strategy of the manager (i.e., the desired manipulation), and the outcome (i.e., the reported earnings surprise).[6]

We use a noise parameter to capture the potential uncertainty of manipulation, and we define noise as a function of the level of desired manipulation. The variable ε follows a discretized normal distribution with mean zero and variance $\mathbb{1}_{m\neq 0} \cdot (1 + \zeta(m-1))\psi^2$, and we denote its probability density function as $\phi_{m,\theta}(\varepsilon)$, as shown below. The ζ parameter captures the possibility that more manipulation leads to more uncertainty in earnings surprise. Further, the indicator function demonstrates that unless the manager decides to manage earnings, $R = e$; without manipulating, there is no uncertainty.[7]

The utility of the manager is thus a function of the state variable (e), the choice variable (m), and a set of parameters ($\theta = \{\eta, \gamma, \psi^2, \zeta\}$) as follows:

$$u(e, m, \theta) = \int_{-\infty}^{\infty} \phi_{m,\theta}(\varepsilon)\mathcal{B}(e + m + \varepsilon)d\varepsilon - \beta\, m^{\gamma}. \tag{3}$$

The manager chooses m to maximize utility,

$$m_e^* = \arg\max_m u(e, m, \theta). \tag{4}$$

The model thus characterizes optimal manipulation behavior as a function of interim

[6]There are a number of reasons this might be the case. For example, the manager may be uncertain about how successful manipulation will be, due to coordination and organizational frictions, or about what precisely the market or analysts will forecast (Agrawal et al. [2006]), or about the outcome of negotiations with the auditor (Gibbins et al. [2001]). Beyer et al. [2018] estimate a model of earnings management incorporating what they call "reporting noise," which plays a similar role as the noise parameter in our setup, and Gao and Jiang [2018] present a model of managerial manipulation where a key feature is uncertainty between the manager's manipulation choice and the observed outcome, also consistent with our setting.

[7]Note that this formulation means that the variance of noise if $m = 1$ is ψ^2 and that ζ being different from zero provides a test for the presence of heteroskedasticity.

earnings relative to the market's expectation, e; the benefit of achieving the benchmark, $\mathcal{B}(\cdot)$; and the four parameters of the model. Given the manager's optimal choice, we can calculate the reported earnings as the sum of the interim earnings and the amount of manipulation, adjusted for noise. More importantly, we can also take a given R and back out the implied e as a function of the parameters, and we use this kind of inversion procedure in the estimation below.

3 Estimation

The goal of our structural model is to rationalize the empirical earnings surprise distribution with the benefits of just meeting or beating this expectation. The necessary first step is to obtain empirical estimates of these two objects. In this section, we discuss our estimation of model inputs, describe how we estimate the model using the simulated method of moments, and present our baseline results.

3.1 Model inputs

Our structural model uses observed market reactions to earnings announcements and the earnings surprise distribution as inputs. To measure these inputs, we follow the earnings management literature and use data from two primary sources for the period 1995 to 2014. We define earnings surprise using analyst forecasts as a proxy for market expectations. Following Bhojraj et al. [2009], we use reported and forecasted values of annual EPS from I/B/E/S, and we define *earnings surprise* as the difference (in cents) between a firm's actual EPS as reported in I/B/E/S and the consensus forecast.[8] In keeping with the literature, we construct the consensus forecast using the latest forecast that preceded the earnings

[8]Since we obtain actual EPS from I/B/E/S, we follow the literature in using reported earnings net of analyst exclusions. Doyle et al. [2013] show evidence that managers are able to get analysts to exclude expenses as a low-cost method of meeting their benchmarks. While our approach does not allow us to speak directly to this strategy, in Appendix A, we show that our model inputs would be similar, albeit estimated with somewhat more noise, if we instead used various GAAP-based definitions of earnings.

announcement for each analyst following the stock during the forecast period.[9] Additionally, we measure EPS surprise in cents, abstracting from potential intracent incentives. This choice is consistent with standard EPS reporting, prior work in the earnings management literature, and evidence from CEO surveys (Graham et al. [2005]).[10]

We measure the market reaction to earnings announcements using three-day cumulative market-adjusted returns (CMAR), which we calculate using daily CRSP data. Again, this definition is consistent with the measurement used in recent papers in the earnings management literature. For example, Bhojraj et al. [2009] define the market reaction to earnings announcements as the five-day cumulative size-adjusted return. Bartov et al. [2002] calculate several alternative measures of earnings announcement returns, including cumulative beta-adjusted returns and cumulative size-adjusted returns over multiple event windows. Although their primary measures are three-day cumulative abnormal returns, they find similar results across all measures.

We present summary statistics for our model inputs in Table 1. In our sample of 49,604 firm–year observations with EPS surprise in the [–20, 20] range, the average EPS surprise is –0.05 cents, and the interquartile range is –3 cents to 3 cents. Of the observations in our sample, 62.43% meet or exceed their analyst benchmarks. Likewise, the average market reaction is positive 0.55%, though the interquartile range varies between −2.79% and 3.81%.

[9]There is some variation in the literature over whether to exclude forecasts that are made close to the earnings announcement. For example, Bhojraj et al. [2009] define earnings surprise as the difference in cents between a firm's reported EPS and the consensus forecast, but exclude forecasts announced subsequent to the second month of the of the firm's fourth quarter of each fiscal year. Similarly, Bartov et al. [2002] impose a restriction that forecasts precede the earnings release by at least three days. We show in Figures A.2 and A.3 in Appendix A that such restrictions are not important for the measurement of either of our model inputs.

[10]Nevertheless, in Appendix A, Figure A.1, we also show that any intracent incentives are at least one order of magnitude smaller than the incentive to meet the analyst benchmark. Across several measures of unrounded EPS, which we construct using data on net income, S&P core earnings, and shares outstanding used for the calculation of EPS from Compustat; the intracent slope is positive, but statistically and economically indistinguishable from zero.

3.1.1 Benefits

A long literature has studied short-term and long-term stock returns around earnings announcements (Bernard and Thomas [1989]; Bernard and Thomas [1990]; Frazzini and Lamont [2007]; Barber et al. [2013]; Foster et al. [1984]; So and Wang [2014]). A growing component of this literature focuses on the role of short-term performance benchmarks (e.g., analyst EPS forecasts, lagged EPS, zero earnings) in the determination of these stock returns (Athanasakou et al. [2011]; Bhojraj et al. [2009]; Bartov et al. [2002]).

Empirically, papers that document the benefits of beating earnings benchmarks typically focus on the well-known difference in earnings announcement returns between firms that just miss and firms that just meet their analysts' consensus EPS forecast. Methodologically, these papers typically compare two subsamples of firms: those that just miss with those that just meet the benchmark.[11] We extend this approach using regression discontinuity tools.

We estimate the conditional expectation function of CMAR given earnings surprise using either (i) global polynomial control functions of earnings surprise (Hahn et al. [2001]), or (ii) local polynomial/linear control functions within some bandwidth of the zero earnings surprise cutoff (Lee and Lemieux [2010]). Because semiparametric control functions are flexible and use the full earnings surprise distribution, this approach is statistically powerful and provides unbiased estimates of the discontinuity in benefits at the zero earnings surprise cutoff (Hahn et al. [2001]; Lee and Lemieux [2010]).

We thus estimate the following:

$$CMAR_{it} = a + \mathcal{B} \cdot MBE_{it} + f^k(Surprise_{it}) + g^j(MBE \times Surprise_{it}) + e_{it}, \tag{5}$$

where $CMAR_{it}$ is the cumulative three-day market-adjusted earnings announcement returns for firm i in year t, $Surprise_{it}$ is the difference between firm i's realized EPS and its analysts'

[11] Keung et al. [2010] show that the market is becoming increasingly skeptical of firms in the [0,1) cent bin, as evidenced by the fact that earnings response coefficients in that bin are much lower than those in adjacent bins. Consistent with Bhojraj et al. [2009], skepticism is warranted because earnings surprises in the [0,1) bin are relatively less predictive of future earnings surprises.

consensus EPS forecast in year t, MBE_{it} is an indicator that equals one if firm i has a nonnegative earnings surprise in year t, $f^k(\cdot)$ and $g^j(\cdot)$ are order-k and order-j flexible polynomial functions of $Surprise_{it}$ on each side of the zero earnings surprise cutoff, and $\mathcal{B}$ represents the discontinuity in capital market benefits of just meeting analysts' consensus EPS forecast at the zero earnings surprise cutoff. We choose k and j in this and in subsequent tests using the Akaike information criterion and the Bayesian information criterion (Hahn et al. [2001]).

We present estimates of $\mathcal{B}$ from equation (5) in Table 2, and we plot the polynomials in Figure 1. We see an estimate of roughly 1.5% in Column (1), and obtain similar results after including firm and year–quarter fixed effects to control for cross-sectional and time series variation in earnings surprise reactions. This means that the effects are not driven by specific firms or outlier years. Our preferred specification is Column (1). In Columns (3) and (4), we instead use local linear control functions and bandwidth restrictions (Lee and Lemieux [2010]), and we find results similar to Columns (1) and (2). Furthermore, it is encouraging that our estimates of $\mathcal{B}$ are qualitatively similar across specifications and are consistent with estimates from the literature (Payne and Thomas [2011]; Keung et al. [2010]; Bhojraj et al. [2009]; Kasznik and McNichols [2002]).

3.1.2 Earnings surprise distribution

We next estimate the earnings surprise distribution semiparametrically. The literature has long noted a jump in the mass of firms moving from just-below to just-above earnings benchmarks and has interpreted this feature of the data as evidence of earnings management (Brown and Caylor [2005]; Dechow et al. [2003]; Degeorge et al. [1999]; Burgstahler and Dichev [1997]; Hayn [1995]; Gilliam et al. [2015]; Brown and Pinello [2007]; Jacob and Jorgensen [2007]). We account for this important feature of the data by allowing the distribution of firms around the benchmark to be discontinuous. As argued in McCrary [2008], the magnitude of this discontinuity provides a natural reduced-form diagnostic test for the likelihood of manipulation, so it also serves as a benchmark for our structural estimates.

We estimate the following:

$$Frequency_b = a + \Delta \cdot MBE_b + f^k(Surprise_b) + g^j(MBE \times Surprise_b) + e_b, \tag{6}$$

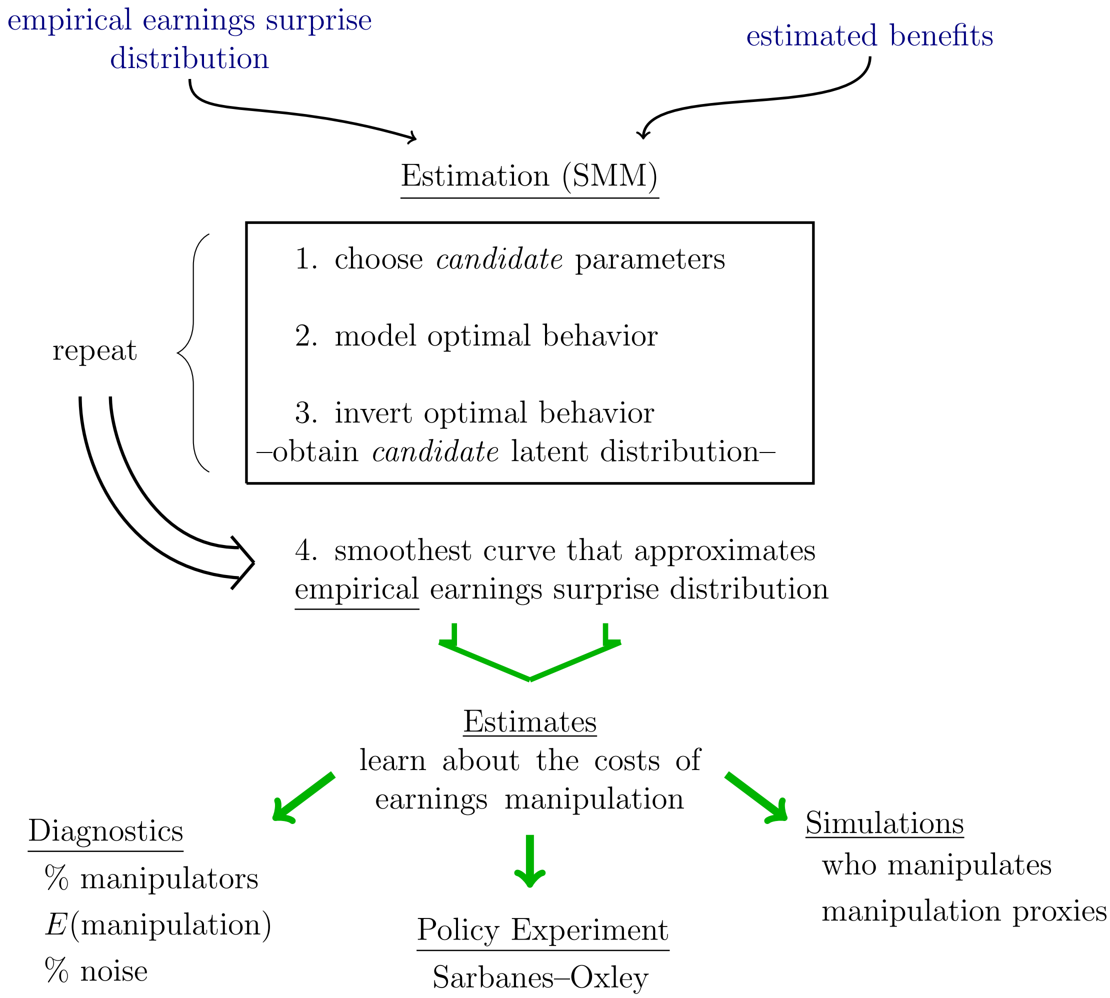
where $Frequency_b$ is the proportion of firm–year observations in earnings surprise bin b, $Surprise_b$ is the earnings surprise in bin b, MBE_b is an indicator that equals one if bin b's earnings surprise is positive, $f^k(\cdot)$ and $g^j(\cdot)$ are order-k and order-j flexible polynomial functions of $Surprise_b$, and Δ represents the discontinuity in frequencies at the zero earnings surprise cutoff.

We present estimates of Δ in Table 3, and we plot the polynomials in Figure 2. Our preferred specification is in Column (1). We estimate a statistically significant discontinuity at zero EPS surprise of 2.63%. This diagnostic evidence is consistent with manipulation. We find similar results using a local linear control and ten cent bandwidths in Column (2). As above, this evidence is broadly in line with differences-in-proportions estimates from the literature.

3.2 Simulated method of moments

Our modeling procedure is illustrated in the figure below. We use as inputs to our model our previously estimated capital market benefits and empirical earning surprise distribution. We estimate the model of Section 2 using the simulated method of moments (McFadden [1989] and Pakes and Pollard [1989]). We choose candidate parameters for the cost function, model optimal behavior given these parameters and estimated benefits, and we invert this optimal behavior to obtain a candidate latent earnings surprise distribution. We evaluate this candidate latent distribution using a criterion function related to its moments, and we repeat these steps until an optimum is found. The following discussion demonstrates the intuition involved in our specific application of the simulated method of moments; however, interested readers can find a more formal discussion with econometric proofs in Appendix B. This overall approach to estimation is similar to that used in recent and contemporaneous papers in the burgeoning structural literature in accounting (Beyer et al. [2018]; Zakolyukina [2018]).

Estimation Method



We begin with the earnings surprise distribution and the estimated capital market benefits from Section 3.1. The first step in the simulated method of moments is to choose candidate parameters for our model of firm behavior. In the second step, we simulate optimal firm behavior, as in equation (4), in which managers take into account the expected benefits and costs of manipulation when deciding how much to manipulate. This second step requires a large number of simulations from each bin in the earnings surprise distribution; in practice, we use 10,000 simulated firms per bin. Therefore, for each cent bin, we draw a cost parameter for each simulated firm, and we calculate its optimal choice and the outcome of this choice. After doing this, we have for each cent bin the proportion of firms that move to a new bin and the proportion of firms that remain in the same bin. This procedure produces the most important object for the estimation: a transition matrix from bin to bin, which essentially models the final position of a firm after starting in a given bin. If we consider $\boldsymbol{\pi}$ as the vector representing the earnings surprise distribution, $\boldsymbol{x}$ as the latent distribution, and $\boldsymbol{P}$ as the transition matrix we can simulate in step two of our procedure, then in matrix notation,

$$\boldsymbol{P} \cdot \boldsymbol{x} = \boldsymbol{\pi} \tag{7}$$

—or—

(Manipulation, bin to bin) $*$ (Latent Earnings Surprise) = Empirical Earnings Surprise.

From equation (7), we can see that the empirical earnings surprise distribution is a function of latent earnings surprise and the transition matrix. Since we observe $\boldsymbol{\pi}$ and we wish to recover $\boldsymbol{x}$, we invert the transition matrix to recover latent earnings in the third step of our application of the simulated method of moments. Using the same notation as above,

$$\boldsymbol{x} = \boldsymbol{P}^{-1} \cdot \boldsymbol{\pi} \tag{8}$$

—or—

(Latent Earnings Surprise) = (*Inverted* Manipulation, bin to bin) $*$ Empirical Earnings Surprise.

In the fourth step, we must choose a latent distribution to identify the parameters of the model. In the simulated method of moments, we do this by evaluating the moments of the latent distribution according to a criterion function, iterating through sets of candidate parameters until we reach an optimum. Our criterion function depends on both the smoothness and the distance between the latent and the empirical distributions.

We face several identification challenges in this setting.[12] To begin with, the transition matrix, $\boldsymbol{P}$ above, must be invertible. That is, to recover the latent earnings distribution, as above, we must be able to find the inverse of this transition matrix. In our application, this is related to having enough cent bins. A second and related concern is the number of relevant moments. In principle, we have 41 moments (drawn from 41 bins), while in practice, managers may manipulate only around the threshold. Given that these methods are only identified through actions taken, this possibility limits the effective moments we can use for identification. We use the available moments to identify four parameters. A third concern in achieving identification is that these moments can separately identify the parameters in the model we choose. This would be an issue if, for example, one modeled the marginal cost of real and accruals earnings management as being linearly additive. In this case, the manipulation share of the two tools would not be separable. Because of this concern, and to lend transparency to our approach, we explicitly discuss the identification of our model parameters in Section 4. In principle, within the limits of these identification constraints, one could identify other interesting economic models using this approach.

With our fitted parameters, we simulate the model to describe the consequences of optimal manipulation: the fraction of firms that manipulate earnings, the fraction of firms that experience noisy outcomes, and the average amount of manipulation for manipulating firms. Further, from the transition matrix, we can determine which bins firms manipulate from and which bins they manipulate to.

As discussed above, our criterion function incorporates two characteristics of candidate latent distributions. The first is smoothness, which we calculate as the sum of the squared differences in the frequencies of adjacent bins. In practical terms, a smoother distribution reduces bunching at the discontinuity, reduces the peak of the discontinuity, and evens out the distribution. The second characteristic is calculated as the sum of the squared differences in the frequencies between candidate latent distributions and the empirical distribution. All

[12]Interested readers can find a technical discussion of these points in Appendix B.

else equal, if two distributions have equal smoothness, our criterion function chooses the distribution that is more similar to the empirical distribution.[13]

This approach is consistent with a world in which the firm's technology produces a smooth latent earnings distribution, and in which the analyst's utility is smooth in forecast error. These conditions do not place restrictions on whether analysts care about forecast accuracy at all, whether they are biased in a positive or negative direction, or whether they can detect manipulation. The conditions require only that small changes in forecast error are associated with small (or no) changes in the analyst's utility. Donelson et al. [2013] develop an alternative approach to study the latent distribution using earnings frauds revealed by securities lawsuits. Although this is a selected setting where the manipulation was particularly egregious, it is reassuring that the pre-manipulation distribution they recover is indeed similar to, but smoother than, the empirical distribution, which is consistent with our criterion function.

An alternative approach could include parametric assumptions about the shape of the latent earnings surprise distribution, such as normality. Similarly, one could take a different approach to the estimation procedure by beginning with a latent distribution of earnings and then simulating optimal behavior by firms starting from that distribution. The problem with each of these approaches is that we do not observe the pre-manipulation earnings surprise distribution, and there is no obvious model to describe it. However, our approach is flexible enough to accommodate other criterion functions that might be more suitable in other contexts.

3.3 Baseline results

Table 4 presents estimates of the model. The parameters we estimate are η, the marginal cost; γ, the curvature of the cost function; ψ^2, the noisiness of earnings manipulation; and ζ, the heteroskedasticity of earnings manipulation noise, which we refer to as the *variance*

[13] In our baseline approach, we give equal weight to each of these criteria. In Section 4.2, we show that shifting the weight toward either criterion yields similar estimates.

multiplier. We find that the marginal cost parameter, η, is 1.61%. We find that the cost function is reasonably convex, with costs increasing with exponent of 2.08. However, we find that earnings management is not certain, and manipulation has a 0.82 cent variance (i.e., outcomes can deviate from planned) that increases nearly fourfold (variance multiplier parameter of 3.71) for each additional cent of manipulation. Figure 3 shows the empirical and latent distributions, as estimated by the model.

From our simulations, we identify the marginal manipulating firm, and we calculate that its marginal cost of manipulation is 1.04%. This firm starts in the [–1,0) bin and chooses to manipulate by a single cent.[14] Notably, the marginal cost of manipulation for this firm is lower than the marginal benefit of meeting the benchmark (1.45%). This apparent inconsistency is explained by the presence of noise. Based on our estimates of the noise parameters, there is a $\sim$28% likelihood that this firm will nonetheless miss its benchmark due to a negative shock. The marginal cost is thus exactly equal to the *expected* marginal benefit.

Given these estimates, we can simulate the model to calculate the proportion of firms that manipulate their earnings. Namely, we find that 2.62% of firms in our sample manipulate earnings. This is remarkably close to 2.63%, the magnitude of the discontinuity in the distribution of firms by earnings surprise estimated in Section 3.1.2. This similarity is notable because our structural approach does not directly target the discontinuity, does not impose any polynomial restrictions on the data, and allows for a much broader range of manipulation strategies than merely moving from just to the left of the threshold to just to the right.

We also show that manipulating firms do so by 1.21 cents on average. Because of stochasticity in manipulation, 59.6% of firms do not reach their intended targets, missing either above or below. The differences in model setups mean that the magnitudes are not directly comparable, but the fact that noise plays a quantitatively important role is consistent with the findings in Beyer et al. [2018]. Figure 4 illustrates the effects of the equilibrium ma-

[14]Because of the convexity of the cost function, the marginal cost of the marginal manipulator in bins farther to the left is necessarily lower.

nipulation strategies on the proportion of firms in each surprise bin. Panel (a) plots the fraction of firms manipulating into and out of each bin. Panel (b) shows the net effect of manipulation strategies on the earnings surprise distribution.

4 Identification and Robustness

4.1 Parameter identification

In this section, we describe the variation used to identify the four structural parameters in the model. While the data identify all parameters simultaneously, it is possible to build some intuition for the features of the data that are associated with each of the parameter estimates. We start with the estimated latent distribution and then calculate optimal manipulation strategies from each bin for counterfactual parameter choices, which yields the counterfactual empirical distributions. Comparing each of these counterfactual empirical distributions with the actual empirical distribution demonstrates the outcomes affected by each parameter.

We examine four counterfactuals, each associated with reducing one of the parameters to 10% of its estimated value and then adding the optimal manipulation to the latent distribution to generate counterfactual empirical distributions. Figure 5 shows these distributions, where the solid lines replicate the actual empirical distribution, and the dashed lines represent the counterfactual distributions. In Table 5, we present the parameters used in each of these simulations as well as statistics to summarize the associated manipulation behavior.

We start with η, the marginal cost parameter, which we estimate to be 1.61%. We reduce this to 0.16% in the counterfactual scenario. The top left panel of Figure 5 shows that this change has a significant effect. Relative to the actual empirical distribution, the counterfactual has much less weight in the bins to the left of the threshold, and much more weight in the bins to the right. This is consistent with a much higher level of manipulation (12.6% versus 2.6%), as can be seen in Table 5. The amount of manipulation (conditional on any manipulation occurring) increases, but by a much smaller amount. This means that, unsurprisingly, reducing the marginal cost parameter leads to more manipulation, especially manipulation of a single cent. Altogether, this tells us that the width of the peak in the bins around the threshold drives the estimation of the marginal cost parameter.

We next turn to the curvature parameter, γ. When this parameter is reduced to 10% of its estimated level, the counterfactual empirical distribution differs significantly across almost the whole support. In this case, unlike with changing the marginal cost parameter, we observe increased manipulation away from the threshold, which results in a significant increase in average manipulation to 10.1 cents (from 1.2 cents), in addition to an increase in the incidence of manipulation. If it is worthwhile to engage in any manipulation at all, then it is worthwhile to engage in extensive manipulation, as the cost function is less steep in this counterfactual. The curvature parameter also interacts with the presence of noise. Some firms manipulate well past the threshold, as insurance against a negative shock, despite the fact that they will enjoy no direct marginal benefits from this action. The top right panel of Figure 5 shows that this parameter is identified by the bins to the right of the threshold, since it would be worthwhile to manipulate past the target only if the cost of large manipulations is low.

Finally, we investigate counterfactual scenarios for the two parameters related to manipulation noise, starting with the variance of noise. Reducing the importance of noise has a similar effect to reducing the marginal cost parameter. The important difference, which can be seen in the bottom left panel of Figure 5, is that changing the noise parameter has a meaningful effect on the optimal strategy only around the threshold. This observation is confirmed by the statistics reported in Table 5, which show that the incidence of manipulation increases, but the level of manipulation is almost unchanged. Reducing noise increases the number of firms that find it worthwhile to manipulate at all, but does not change the optimal strategy, conditional on manipulation occurring.

This argument suggests that the noise parameter is identified by the bin-to-bin difference around just meeting the benchmark, as distinct from the width of the peak, which determines the marginal cost parameter. We further consider the role of the noise parameter in explaining the incidence of just failing to meet the target, in Section 5. Last, the figure and the table suggest that the variance multiplier parameter does not play an important role in determining optimal manipulation behavior. Reducing this parameter to 10% of its estimated level leads to a small decrease in the incidence of manipulation and a small increase in the level of manipulation, neither of which is discernible in the figure.

To provide further intuition for the variation that underlies the estimated parameters, we investigate the sensitivity of the estimates to the features of the empirical distribution by changing the empirical distribution of earnings and re-estimating parameters. First, we reduce the peak found in the [0,1) cent bin, halving its size relative to the [1,2) cent bin. Using this *Reduce peak* empirical distribution of earnings, the implied marginal costs increase by 27%, while the curvature parameter increases by only 5%. If we remove the peak entirely, making the [0,1) cent bin equal to the [1,2) cent bin, then the curvature remains unchanged, but marginal costs increase by 70% relative to the baseline. Further, *Reduce peak* and *Kill peak* reduce the implied number of manipulators to 1.91% and 1.36% of firms, respectively. Consistent with the conclusions drawn from the other counterfactual scenarios discussed above, these alternative estimates demonstrate that our results are sensitive to the size of the earnings surprise discontinuity, and they demonstrate that removing this discontinuity yields much lower levels of manipulation, via higher cost parameters.

4.2 Estimation robustness

In Table 6, we show that our results are robust to differences in input measurement and model specifications. We use bins from [–20,–19) to [20,21) cents in our main estimates, but most of the manipulating firms are in bins much closer to the discontinuity. As such, it is important to show that reducing the number of bins we use to fit our model does not meaningfully change our results. Our results are similar when we estimate our model on data from the [–15,–14) to [15,16) cent bins, or if we further reduce the range of bins to the

[–10,–9) to [10,11) or [–5,–4) to [5,6) bins. These estimates are the most similar for the most important parameters, including η, which varies from 1.54% to 1.67% (versus a baseline of 1.61%), and ψ^2, which varies from 0.76 to 0.79 (versus a baseline of 0.82). In addition, $\%manip$ is relatively unchanged, varying from 2.56% to 2.84% (versus a baseline of 2.62%), as is $\%noise$, varying from 55.7% to 60.0% (versus a baseline of 59.6%).

Our estimation equally weights similarity to the empirical distribution and smoothness as qualities of the latent earnings surprise distribution. To investigate sensitivity to this choice, we re-estimate our model twice, first overweighting empirical similarity by 10%, and then overweighting smoothness by 10%. Our results do not significantly change, and cost parameters and manipulation move in the expected direction. Namely, when empirical sameness is emphasized, costs increase (1.70% versus 1.61%) and manipulation decreases (2.45% versus 2.62% of firms manipulate), but when smoothness is emphasized, costs decrease (1.46%) and manipulation increases (3.11% of firms manipulate). These are intuitive effects because more manipulation smooths the discontinuity. Comparing these three different weighting schemes, we are comforted by the fact that our preferred scheme yields a propensity to manipulate almost exactly equal to that of a simple, theory-free McCrary [2008] test, as can be seen from the estimates in Table 3. This is our preferred weighting because we want to avoid imposing a higher rate of manipulation by oversmoothing.

Because ζ has the greatest standard error and plays the smallest role in our model, we also estimate the model without ζ (setting $\zeta = 0$). Our estimates change the most in this robustness check. While ζ does not have major effects on equilibrium behavior, we can see from its effect on other parameters that it is an important feature of the model.

In untabulated results, we employ the full distribution of capital markets benefits for different levels of earnings relative to the benchmark. As shown in Figure 1, the main difference is the existence of small, but positive, marginal benefits of increasing earnings beyond the benchmark. In this test, we obtain similar parameter estimates for the cost function; however, this yields somewhat higher levels of manipulation due to increased benefits.

4.3 Benefit heterogeneity

So far, we have assumed that the benefits of manipulating earnings are the same for all firms that are in the same earnings surprise bin. In this subsection, we extend our model to consider the possibility of heterogeneity in benefits. Importantly, for this heterogeneity to have an effect on the firm's earnings management problem, it must be known by the firm before the reporting decision is made. It is likely that a significant ex post variation in benefits exists, but this variation is not relevant to manager behavior, since we assume the manager is risk neutral and thus acts only on her ex ante expectation of benefits. To study this issue, we incorporate cross-firm variation in expected benefits, using the dispersion (standard error) in our estimated benefit parameter. This problem is straightforward because once the firm knows its benefit, the structure of the tradeoff that determines the optimal manipulation is the same as in our base model.

We present several parameterizations of benefits heterogeneity in Table 7. In the second row, before making the reporting decision, firms draw a benefit from a normal distribution with mean and standard deviation as in the first column of Table 2, our preferred model for estimating benefits. We instead allow for the distribution of benefits to come from a Bernoulli distribution (with the same underlying mean and standard deviation) in the fourth row, which we interpret as firms either receiving a "large" benefit or no benefit at all. In either case, our estimates are essentially the same as in our baseline model, which lacks this type of heterogeneity. In the third and fifth rows of Table 7, we magnify the variation (i.e., increasing the standard deviations by a factor of five, well beyond any of our estimates in Table 2) as a robustness check and to gain a better understanding of how this new feature affects our estimation. Our estimates are again consistent with the baseline. We conclude from this investigation that, in practice, unconditional heterogeneity in benefits is not an important feature of the earnings management problem.

The economics of heterogeneity becomes much more interesting, at least in theory, if firm-specific benefits are correlated with firm-specific costs. This kind of correlation broadly allows for relaxation of the assumption that manipulation (or firm-side heterogeneity) is not observable by investors who make valuation decisions. With a nonzero correlation, we allow investors to differentially reward firms in the same bin in a way that is correlated with the firm's manipulation costs. To study this issue, and to understand its empirical importance, we augment the above analysis of unconditional benefits heterogeneity to allow for a correlation between the firm's benefits and the marginal cost parameter. In particular, we take the normal distribution, and we incorporate correlations of 0.5 and -0.5. These large correlations imply that the market receives a strong signal about individual firms and their likelihood of engaging in manipulation.

This structure can capture a variety of particular economic models of investor behavior. For example, the positive correlation has a natural interpretation as the market wanting to reward good performance, but not manipulated performance. When investors see high costs (or a signal of high costs), they rationally infer that, all else being equal, firms are less likely to manipulate. In this case, firms that meet expectations are more likely to have done so based on real performance and thus these firms warrant a greater benefit. On the other hand, one would observe a negative correlation from a model in which low costs (and thus the capacity to manipulate earnings) are a positive signal, perhaps indicating financial flexibility or a sophisticated understanding of the business on the part of managers. In this case, investors see low costs and rationally pay a relatively high reward for meeting expectations, since such firms are expected to be of better quality along some relevant dimension. Rather than take a stand on which of these mechanisms dominates, our goal in this analysis is simply to obtain a general understanding of whether these kinds of mechanisms are empirically significant.

Table 7 presents the 0.5 and -0.5 correlation models in row six and row eight, respectively. Our findings are essentially unchanged, even when we also use the exaggerated variation in benefits, as discussed above. We thus conclude that allowing for a relationship between benefits and costs at the firm level does not have a meaningful impact on our understanding of the manager's manipulation decision. This finding has two important implications. First, identifying a correlation parameter is not feasible in our setting because even wide variation in this parameter has a limited effect on the fit of the model. More importantly, we believe that these findings support our approach of focusing on the problem of an individual manager taking the capital market benefits as given, rather than explicitly modeling the source of these benefits. Despite this simplification, the results in Table 7 suggest that our findings are generalizable to a more sophisticated general equilibrium model in which both the market and the manager are strategic players.

5 Applications

An understanding of the costs of earnings management has many implications both for how we interpret firm behavior and for the validity of measures employed in the literature to identify earnings management. In this section, we discuss a number of applications of the results described above.

5.1 Why do so few firms manage earnings?

Given the significant benefit of meeting earnings benchmarks, and given the results of surveying managers themselves (Dichev et al. [2016]), our estimated fraction of firms that manage earnings is surprisingly low. Of particular interest are firms in the [–1,0) earnings surprise bin, since these firms have the most potential to take advantage of the discontinuity in capital market responses to earnings announcements. The counterfactual simulations discussed above can help explain why so many firms remain below, and yet very close to, the benchmark.

Two of the parameters in the cost function appear to drive this behavior: the marginal cost parameter, η, and the noise parameter, ψ^2. In the latent distribution, 6.1% of firms are in the [-1,0) cent bin. Reducing the marginal cost parameter to 10% of its estimated level, as in Table 5, reduces this fraction to 3.0%. This reduction is a straightforward consequence of decreasing the marginal cost of manipulation. If the noise parameter is simultaneously reduced to 10% of its estimated level, only 0.9% of firms remain in the -1 bin. Reducing the magnitude of noise has two key effects. The optimal strategy for managers who face relatively low costs in bins farther away from the benchmark may be to manipulate upward to -1 in the hopes of receiving a shock that helps them meet the benchmark.[15] Reducing noise decreases the expected payoff of this strategy. The second effect is mechanical: negative noise leads some firms that meet the benchmark to fall into the -1 bin.

The results of these counterfactuals show that both the marginal cost effect and the noise effect are economically significant explanations for the relative lack of manipulation out of the -1 bin. Determining the exact contribution of these two forces is complicated because they interact. As marginal costs decrease, more firms aim to barely meet the benchmark, but the mechanical effect of noise means that a higher noise parameter shocks more of these firms back down into the -1 cent bin. However, this interaction accounts for only 9% of the overall effect of the counterfactual.

5.2 Identifying "suspect" firms

Our model estimates reveal the incidence (and amount) of manipulation in each earnings surprise bin, which is of interest given the prevalent use of proxy measures of manipulation based on whether a firm ends up in the [0,1) cent bin (e.g., Bhojraj et al. [2009], Cohen et al. [2008], Roychowdhury [2006]). Table 8 quantifies how well various definitions of "suspect" firms fare in accurately identifying manipulation. Row one shows the percentage of firms in each surprise bin that engaged in manipulation to reach that bin. The second row captures the percentage of all manipulators from the sample that ended up in that surprise bin. Rows three and four replicate the quantification of rows one and two, respectively, while weighting

[15] Although managers are risk neutral (so that noise is not necessarily harmful), there is an asymmetry in the effect of noise on the manager's objective function. From the [0,1) cent bin, which is the most likely target, a negative shock hurts much more than a positive shock helps, because of the nature of the discontinuity in benefits. From the [–1,0) bin, the asymmetry is reversed, with the expected benefit of noise becoming positive.

the manipulating firm observations according to the mean level of manipulation in each surprise bin (e.g., the firm counts twice if it manipulated upward by two cents).

The first column of Table 8 describes the firms captured by the typically used [0,1) "suspect" bin. Of the firms in this bin, 11% manipulated in order to reach this bin, which is perhaps surprisingly low, but in line with our overall estimates of the extent of manipulation in aggregate. This subset covers 53.1% of all manipulators in the sample. On these two metrics, this bin performs the best at identifying manipulation when restricted to single bin measures. The next best bin to add to the measure would be [1,2), which reduces accuracy to 9.9% but also accounts for 80.9% of the total manipulators. Given the salience of the discontinuity in the capital market response to earnings surprise, it is interesting that the bin with the third most manipulators is the [–1,0) bin. As described in the previous subsection, this is in part because the possibility of a positive shock makes manipulating into this bin the optimal strategy for some firms that have an intermediate level of costs. If the priority were to account for nearly all of the manipulators, the [–1,3) range may be desirable in that 97.4% of manipulators fall into this range and the accuracy falls only to 7.3%. Relative to the [0,1) bin, this reflects an 83% increase in the fraction of all manipulators that are included, at the cost of reducing accuracy by 34%.

Turning to rows three and four of Table 8, we find a similar pattern of results, though with a clear tendency toward improvement from including more bins. This is because bins that are farther away from the benchmark have a higher expected rate of manipulation (conditional on manipulation occurring) relative to bins closer to the benchmark. For example, in this case, the [0,2) range performs better than [0,1) on both margins, increasing the weighted fraction of manipulators and significantly increasing the fraction of the total manipulation covered. Extending to [–1,3) yields an even larger increase in coverage and reduces accuracy by only 13% relative to the [0,2) range.

Overall, Table 8 shows how different definitions of "suspect" bins trade off type I and type II errors and can help guide this measurement choice in a variety of contexts. Further, it may be preferred for some applications to use a continuous measure of the probability of manipulation by bin. This approach would naturally place more weight on bins close to the benchmark and less weight on those farther away, rather than making a discrete choice of which bins to include in the measure and which to exclude.

5.3 Sarbanes–Oxley and the cost of earnings management

After a string of accounting scandals in the late 1990s and early 2000s, including Enron and Worldcom, the Sarbanes-Oxley Act (SOX) was enacted in 2002. In line with the SEC's stated objective of protecting investors, SOX initiated broad changes in corporate disclosures, with the goal of improving the reliability of financial information in company reports. Our methodology can be used to uncover the effects of such regulation on earnings management by exploiting time series variation in the empirical earnings surprise distribution and the benefits of meeting earnings benchmarks.

SOX provides an ideal setting for such an analysis because the effects of SOX have been of great interest to both practitioners and academics. Cohen et al. [2008] and Bartov and Cohen [2009] find that earnings management declined after the introduction of SOX, and they attribute this decline specifically to reduced accruals-based earnings management (as well as a decline in downward expectations management using negative managerial guidance). Cohen et al. [2008] note that the three years before SOX were characterized by an abnormal increase in earnings management relative to earlier in the 1990s.

To investigate the effects of SOX using our methodology, we examine subsamples of the three years before SOX (1999–2001) and the three years after it (2002–2004).[16] Table 9 presents estimates from these two subsamples. Consistent with the work discussed above, our diagnostic test statistic of the extent of manipulation (i.e., the discontinuity in the earnings distribution Δ) falls post-SOX. After SOX, the benefit of meeting the benchmark ($\mathcal{B}$) increases by 0.80%. Together, these two findings suggest that SOX had its intended

[16] Our results are not sensitive to these specific definitions of the pre- and post-SOX periods. For example, dropping either 2001 or 2002 yields similar estimates.

effect of improving the quality of financial reporting, and they also suggest that capital markets responded with larger rewards for meeting the benchmark. This is likely due to an increased willingness by markets to attribute the meeting of earnings benchmarks to fundamental performance rather than to manipulation.

Next, we use our structural model to estimate the consequences of these changes for the costs of earnings management as well as the extent of manipulation. As expected, we find significant increases in costs. Table 9 shows that both the marginal cost parameter and the curvature parameter increase significantly. This suggests that both the cost of the first cent of managed earnings and the cost of incremental manipulation both increased, consistent with the goals of the reform.[17] Earnings management also became much less noisy after SOX; we conjecture that this is due to a switch away from riskier, less certain strategies and toward a more conservative use of discretion.[18] In the aggregate, manipulation fell from 4.05% of firms to 2.61%, a 36% reduction. This is evidence that SOX had its desired effect on financial reporting and earnings management.

6 Conclusion

Two prominent stylized facts emerge from the earnings management literature: firms that just meet their consensus analyst earnings forecast benefit from positive abnormal returns, and the earnings surprise distribution exhibits a large discontinuity at this benchmark. In this paper, we develop and estimate an economic model of the earnings management decision that links these stylized facts. Our model and estimation approach are generalizable to other settings where we wish to understand unobserved behavior when only some of the decision-

[17] The costs could have increased through a number of mechanisms. For example, auditors likely increased their scrutiny of reported earnings both because of increased regulation on their behavior, and as a rational response to the increased risk to their survival, made evident by the collapse of Arthur Andersen.

[18] The fact that we estimate a relatively small noise parameter post-SOX is also reassuring with regard to our main results. One might have been concerned that the role of noise was conflated with a model misspecification in some other dimension. The post-SOX partition shows that our approach can indeed produce significant variation in the importance of noise, suggesting that the higher noise parameter we estimate for the full sample is indeed indicative of uncertainty in earnings manipulation.

relevant information can be observed.

In our context, the cost of earnings management is unobserved. As we recover the implied cost function, we learn that costs are convex and subject to noise. Manipulation consistent with this cost function and the observed benefits and manager behavior is infrequent and small. We address several outstanding questions in the earnings management literature using applications of our methodology. Notably, we assess the validity of identifying manipulation using a "suspect" earnings surprise bin, and we estimate the effect of SOX on the financial reporting process.

Our approach is flexible enough to allow for alternative identification assumptions, measurement of empirical inputs, and economic models of behavior. Therefore, we believe our approach can be used both to further our understanding of the tradeoffs that underlie earnings management and also to study manipulation, broadly defined, in other settings. For example, our model could be straightforwardly re-estimated using an alternative benchmark, such as zero earnings. Since this benchmark is fixed, rather than being chosen by potentially strategic, forward-looking, analysts, differences in the estimated cost function and recovered latent distribution could be informative about the objectives and behavior of analysts. Perhaps the most interesting way that our model could be extended would be to make it explicitly dynamic, such that optimal strategies could involve the possibility of downward manipulation in order to ease a firm's ability to surpass future benchmarks.

References

Anup Agrawal, Sahiba Chadha, and Mark A. Chen. Who is afraid of Reg FD? The behavior and performance of sell-side analysts following the SECs fair disclosure rules. *The Journal of Business*, 79(6):2811–2834, 2006.

Vasiliki Athanasakou, Norman C. Strong, and Martin Walker. The market reward for achieving analyst earnings expectations: does managing expectations or earnings matter? *Journal of Business Finance & Accounting*, 38(1-2):58–94, 2011.

Brad M. Barber, Emmanuel T. De George, Reuven Lehavy, and Brett Trueman. The earnings announcement premium around the globe. *Journal of Financial Economics*, 108(1):118–138, 2013.

Eli Bartov and Daniel A. Cohen. The "Numbers Game" in the pre-and post-Sarbanes-Oxley eras. *Journal of Accounting, Auditing & Finance*, 24(4):505–534, 2009.

Eli Bartov, Dan Givoly, and Carla Hayn. The rewards to meeting or beating earnings expectations. *Journal of Accounting and Economics*, 33(2):173–204, 2002.

Victor L. Bernard and Jacob K. Thomas. Post-earnings-announcement drift: Delayed price response or risk premium? *Journal of Accounting Research*, 27:1–36, 1989.

Victor L. Bernard and Jacob K. Thomas. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics*, 13(4):305–340, 1990.

Anne Beyer, Ilan Guttman, and Iván Marinovic. Earnings management and earnings quality: Theory and evidence. *The Accounting Review, forthcoming*, 2018.

Sanjeev Bhojraj, Paul Hribar, Marc Picconi, and John McInnis. Making sense of cents: An examination of firms that marginally miss or beat analyst forecasts. *The Journal of Finance*, 64(5):2361–2388, 2009.

Lawrence D. Brown and Marcus L. Caylor. A temporal analysis of quarterly earnings thresholds: Propensities and valuation consequences. *The Accounting Review*, 80(2):423–440, 2005.

Lawrence D. Brown and Arianna Spina Pinello. To what extent does the financial reporting process curb earnings surprise games? *Journal of Accounting Research*, 45(5):947–981, 2007.

David Burgstahler and Ilia Dichev. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics*, 24(1):99–126, 1997.

David Burgstahler and Michael Eames. Management of earnings and analysts' forecasts to achieve zero and small positive earnings surprises. *Journal of Business Finance & Accounting*, 33(5-6):633–652, 2006.

Marcus L. Caylor. Strategic revenue recognition to achieve earnings benchmarks. *Journal of Accounting and Public Policy*, 29(1):82–95, 2010.

Louis K.C. Chan, Jason Karceski, and Josef Lakonishok. Analysts' conflicts of interest and biases in earnings forecasts. *Journal of Financial and Quantitative Analysis*, 42(04): 893–913, 2007.

Daniel Cohen, Raj Mashruwala, and Tzachi Zach. The use of advertising activities to meet earnings benchmarks: Evidence from monthly data. *Review of Accounting Studies*, 15(4): 808–832, 2010.

Daniel A. Cohen, Aiyesha Dey, and Thomas Z. Lys. Real and accrual-based earnings management in the pre-and post-Sarbanes-Oxley periods. *The Accounting Review*, 83(3):757–787, 2008.

Julie Cotter, Irem Tuna, and Peter D. Wysocki. Expectations management and beatable targets: How do analysts react to explicit earnings guidance? *Contemporary Accounting Research*, 23(3):593–624, 2006.

Patricia M. Dechow, Scott A. Richardson, and Irem Tuna. Why are earnings kinky? An examination of the earnings management explanation. *Review of Accounting Studies*, 8 (2-3):355–384, 2003.

Francois Degeorge, Jayendu Patel, and Richard Zeckhauser. Earnings management to exceed thresholds. *The Journal of Business*, 72(1):1–33, 1999.

Ilia Dichev, John Graham, Campbell R. Harvey, and Shiva Rajgopal. The misrepresentation of earnings. *Financial Analysts Journal*, 72(1):22–35, 2016.

Dain C. Donelson, John M. McInnis, and Richard D. Mergenthaler. Discontinuities and earnings management: Evidence from restatements related to securities litigation. *Contemporary Accounting Research*, 30(1):242–268, 2013.

Jeffrey T. Doyle, Jared N. Jennings, and Mark T. Soliman. Do managers define non-GAAP earnings to meet or beat analyst forecasts? *Journal of Accounting and Economics*, 56(1): 40–56, 2013.

George Foster, Chris Olsen, and Terry Shevlin. Earnings releases, anomalies, and the behavior of security returns. *The Accounting Review*, 59(4):574–603, 1984.

Andrea Frazzini and Owen Lamont. The earnings announcement premium and trading volume. *Working paper*, 2007.

Pingyang Gao and Xu Jiang. The Economic Consequences of Discrete Recognition and Continuous Measurement. *Working paper*, 2018.

Michael Gibbins, Steven Salterio, and Alan Webb. Evidence about auditor–client management negotiation concerning clients financial reporting. *Journal of Accounting Research*, 39(3):535–563, 2001.

Thomas A. Gilliam, Frank Heflin, and Jeffrey S. Paterson. Evidence that the zero-earnings discontinuity has disappeared. *Journal of Accounting and Economics*, 60(1):117–132, 2015.

John R. Graham, Campbell R. Harvey, and Shiva Rajgopal. The economic implications of corporate financial reporting. *Journal of Accounting and Economics*, 40(1):3–73, 2005.

Jinyong Hahn, Petra Todd, and Wilbert Van der Klaauw. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209, 2001.

Carla Hayn. The information content of losses. *Journal of Accounting and Economics*, 20 (2):125–153, 1995.

John Jacob and Bjorn N. Jorgensen. Earnings management and accounting income aggregation. *Journal of Accounting and Economics*, 43(2):369–390, 2007.

Michael C. Jensen. Agency costs of overvalued equity. *Financial Management*, 34(1):5–19, 2005.

Ron Kasznik and Maureen F. McNichols. Does meeting earnings expectations matter? Evidence from analyst forecast revisions and share prices. *Journal of Accounting Research*, 40(3):727–759, 2002.

Edmund Keung, Zhi-Xing Lin, and Michael Shih. Does the stock market see a zero or small positive earnings surprise as a red flag? *Journal of Accounting Research*, 48(1):91–121, 2010.

William J. Kross, Byung T. Ro, and Inho Suk. Consistency in meeting or beating earnings expectations and management earnings forecasts. *Journal of Accounting and Economics*, 51(1):37–57, 2011.

David S. Lee and Thomas Lemieux. Regression discontinuity designs in economics. *Journal of Economic Literature*, 48:281–355, 2010.

Justin McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714, 2008.

Daniel McFadden. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica*, 57(5):995–1026, 1989.

Ariel Pakes and David Pollard. Simulation and the asymptotics of optimization estimators. *Econometrica*, 57(5):1027–1057, 1989.

Jeff L. Payne and Wayne B. Thomas. The torpedo effect: Myth or reality? *Journal of Accounting, Auditing & Finance*, 26(2):255–278, 2011.

Sugata Roychowdhury. Earnings management through real activities manipulation. *Journal of Accounting and Economics*, 42(3):335–370, 2006.

Eric C. So and Sean Wang. News-driven return reversals: Liquidity provision ahead of earnings announcements. *Journal of Financial Economics*, 114(1):20–35, 2014.

Anastasia A. Zakolyukina. How common are intentional GAAP violations? Estimates from a dynamic model. *Journal of Accounting Research*, 56(1):5–44, 2018.

Amy Y. Zang. Evidence on the trade-off between real activities manipulation and accrual-based earnings management. *The Accounting Review*, 87(2):675–703, 2011.

A Measurement Robustness

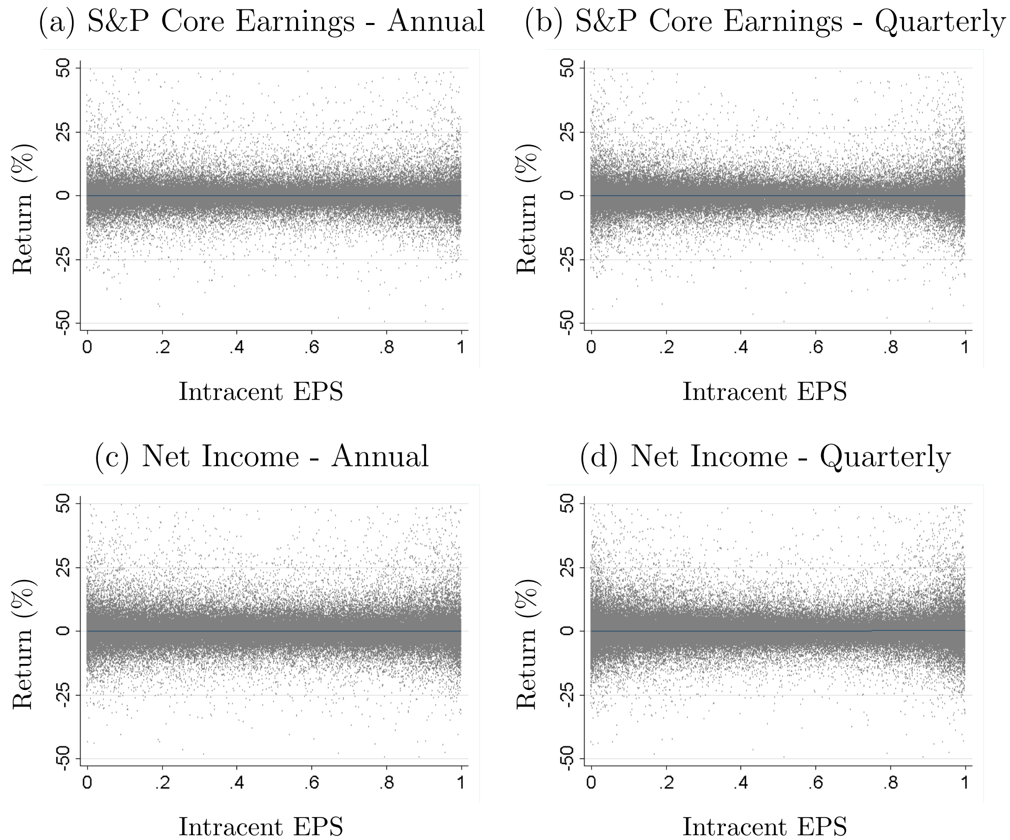
Our estimation relies on two model inputs: the market reaction to earnings announcements and the earnings surprise distribution. As a result, the robustness of our economic inferences depends on our measurement choices for these two inputs. We measure the market reaction to earnings announcements using three-day cumulative market-adjusted returns (CMAR). While there is limited consensus in the earnings management literature on measuring event study returns, our choice on this dimension is standard in event study methodology, and alternative benchmark returns for event studies should have limited effect on inference, since market and factor returns are, on average, small over short windows.

The earnings surprise construct is also not devoid of measurement choices. We measure earnings surprise as the difference in cents between the reported EPS and the consensus analyst forecast, which we define as the average forecast error among analysts who cover the stock during the reporting period. There are two important choices to make regarding this measurement. First, we measure EPS surprise discretely in cents and, therefore, abstract from incentives derived from intracent surprises. Our assumption is that, from the perspective of economic agents in our model, an earnings surprise of 5.1 cents is equivalent to an earnings surprise of 5 cents. Naturally, investors may reward firms for intracent surprises, and, if these incremental rewards are significant, managers may respond to these incremental incentives.

To investigate the importance of this measurement choice, we present evidence of earnings announcement returns as a function of the intracent EPS surprise. In Figure A.1, we measure the remainder of reported EPS after rounding down to the nearest integer. To do so, we use data on net income (i.e., S&P core earnings) and shares outstanding used in EPS calculations from Compustat. We present subfigures for EPS defined using net income or S&P core earnings and, in each figure, we use three-day CMAR as our measure of market reactions to earnings announcements.

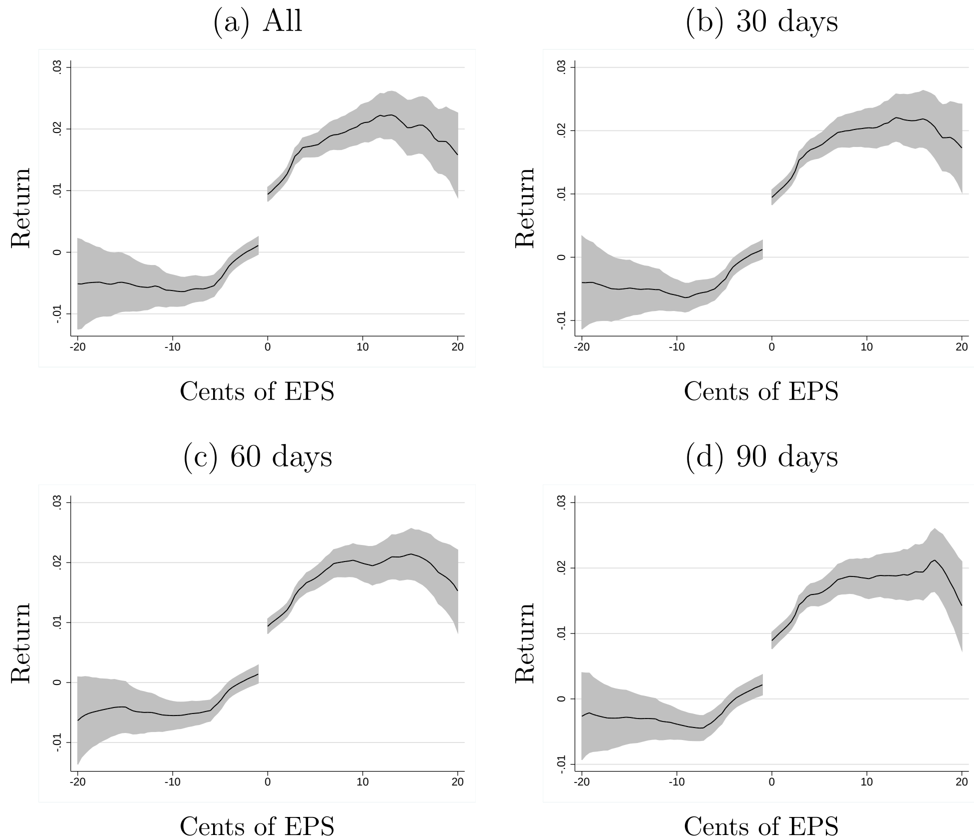
The thin navy blue line in each figure is the fitted line for the conditional expectation function of earnings announcement returns to intracent EPS, and the dots represent individual data points. In each subfigure of Figure A.1, we estimate that the navy blue line has a statistically insignificant and positive slope. The point estimates in corresponding linear regressions are all less than 0.1%, which suggests that any intracent benefits to EPS surprise, if they exist at all, are at least one order of magnitude smaller than the 1.45% return benefit for just meeting the analyst consensus forecast.

Figure A.1: Intracent Benefits



A second measurement choice made in the earnings management literature concerns which analyst forecasts to consider when defining the consensus forecast. For example, Bhojraj et al. [2009] define the consensus forecast using forecasts from the penultimate month of the final fiscal quarter, excluding forecasts that immediately precede the earnings announcement. We make no such restrictions in our baseline measures. To investigate the effect of removing forecasts that occur late in the fiscal year, we construct alternative EPS surprise measures based on forecasts that precede the fiscal year-end by at least 30, 60, or 90 days. In Figure A.2, we plot the market reaction to earnings announcements conditional on EPS surprise using local polynomial functions for each of these definitions of earnings surprise for comparison to our benchmark definition.

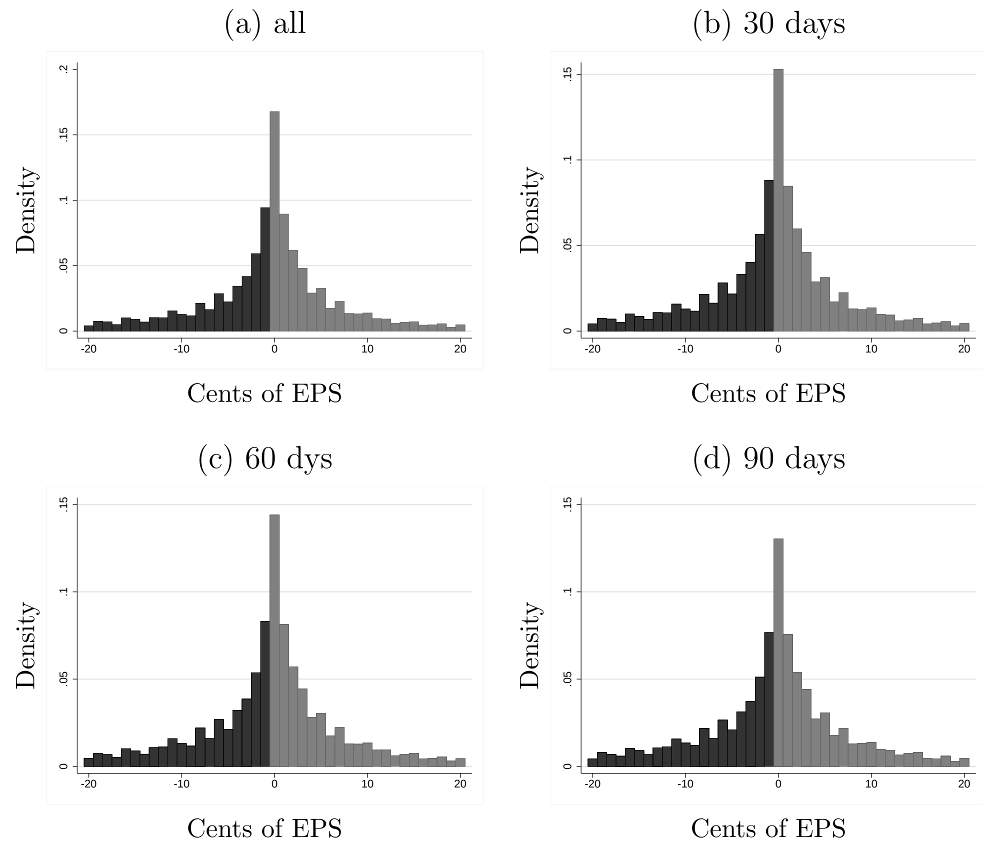
Figure A.2: Forecast Heterogeneity: Benefits



Specifically, Figure A.2 plots the benefits distribution for four different measures of EPS surprise, and it provides visual evidence that this measurement choice is unlikely to have an impact on our estimates. For a quantitative comparison, we estimate the discontinuity in benefits at zero EPS surprise using our preferred specification from equation (5). In untabulated results, we also estimate coefficient estimates that are similar to our preferred estimate. We also plot the EPS surprise distribution for each alternative definition of EPS surprise using bar plots.

Figure A.3 plots the empirical distribution of the EPS surprise using four different measures of EPS surprise, each of which places a different restriction on forecast recency relative to the fiscal year-end. These figures suggest that the forecasts that precede the fiscal year-end by 30, 60, or 90 days have mild effects on the EPS surprise distribution. We estimate the discontinuity in these distributions as in equation (6), and we find estimates that are similar to our preferred estimate of 2.63%. Overall, the figures in this appendix lend credence to the robustness of our estimation to alternative measurement choices.

Figure A.3: Forecasts Heterogeneity: Distribution

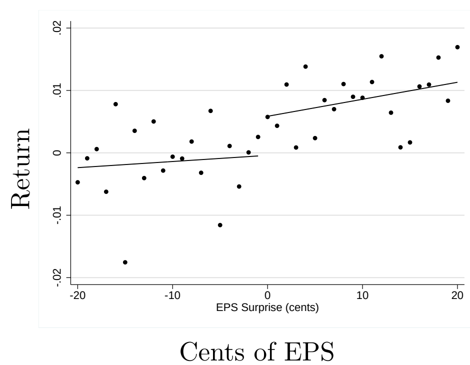


We use actual and forecasted EPS data from I/B/E/S, which matches forecasts to actuals according to the measure forecasted. To corroborate this measurement choice, we investigate the distributional properties of $EPS\,Surprise$ and $CMAR^{3\,day}$ when using alternative measures of actual earnings. In the figures that follow, we measure $EPS\,Surprise$ using analyst forecasts from I/B/E/S as before, but we vary the measure of actual EPS. We construct $CMAR^{3\,day}$ as before. We investigate four alternative measures of actual EPS, each of which we collect from the Compustat Annual data file: (a) basic EPS including extraordinary items, (b) basic EPS excluding extraordinary items, (c) diluted EPS including extraordinary items, and (d) diluted EPS excluding extraordinary items.

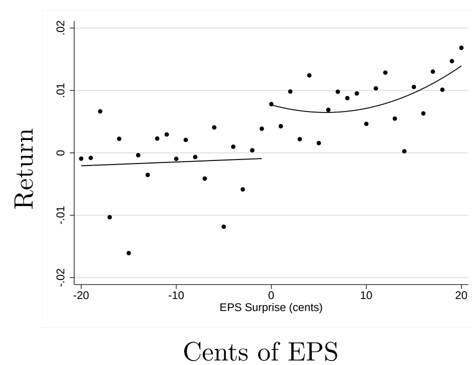
These figures show qualitatively similar evidence of a discontinuity in the capital market benefits of meeting the market's expectations and in the earnings surprise distribution. The economic magnitude of the discontinuity, in particular, is moderately smaller than our baseline estimates that use actual EPS from I/B/E/S. This difference be explained by two alternatives. First, because I/B/E/S matches the actual and forecasted quantities, tests using this data source have less measurement error in earnings surprise. Introducing measurement error to the running variable in a threshold-based empirical design poses a challenge for quantification, since the mass of firms around the threshold will be smoothed by noise from this measurement error. Second, I/B/E/S forecasts incorporate analyst exclusions, which may be chosen strategically (Doyle et al. [2013]). In particular, one inexpensive source of EPS manipulation could be selecting exclusions that analysts will either fail to anticipate or overestimate the effects of the excluded items on EPS. If managers can fool analysts either about which items will be excluded or about the effects of the exclusions, the smaller discontinuity estimates from these measures may reflect a form of earnings manipulation. Because we are unable to disentangle these alternatives, we prefer estimates that use the matched actual-forecast data from I/B/E/S.

Figure A.4: Alternative GAAP Earnings Measures: Benefits

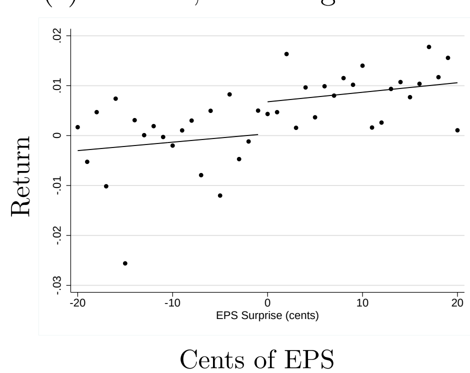
(a) Basic, including ex. items



(b) Basic, excluding ex. items



(c) Diluted, including ex. items



(d) Diluted, excluding ex. items

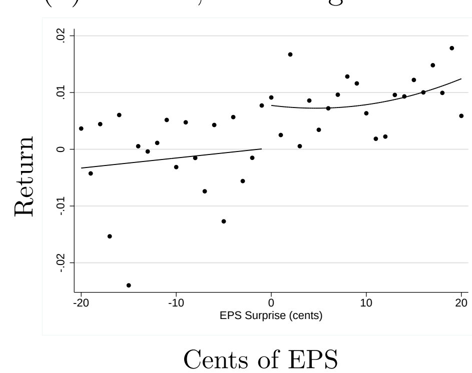
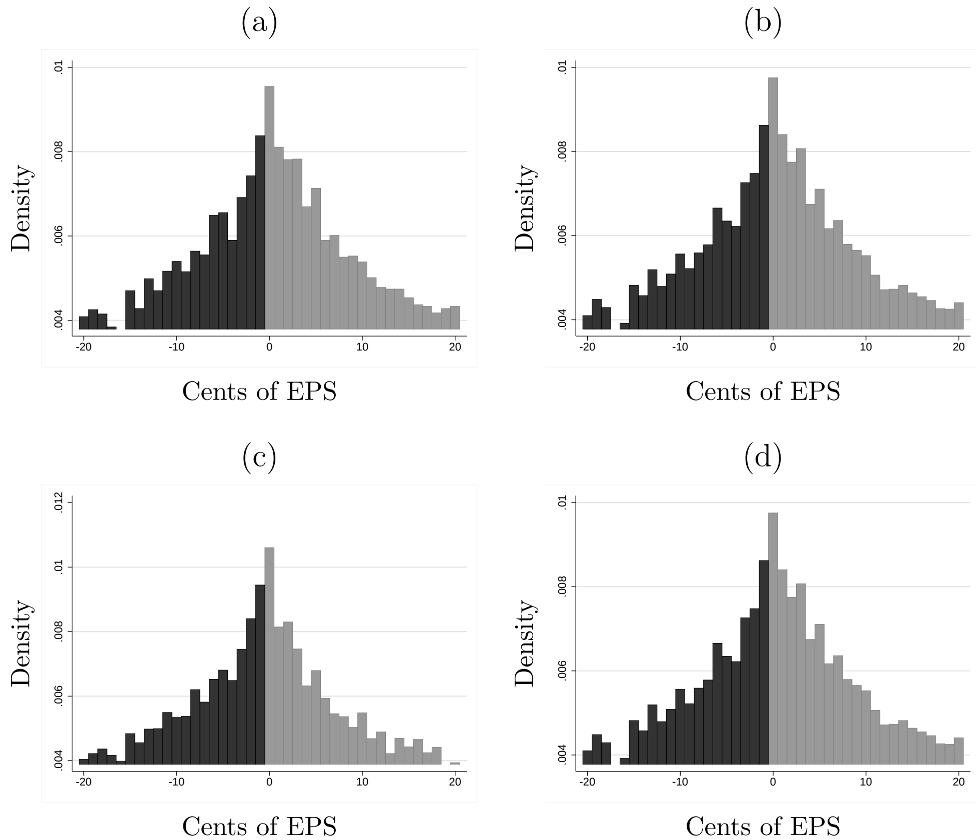


Figure A.5: Alternative GAAP Earnings Measures: Earnings Surprise



B Estimating Equations

There exists an empirical earnings surprise distribution from $[-20, 20]$ cents per share, in one cent increments. Let $\boldsymbol{\pi}$ be the vector representing the number of firms with each per–share earnings surprise. We index π_1 as the number of firms that have -20 cents in per–share earnings surprise, and π_i as the number of firms that have $-21 + i$ cents in per–share earnings surprise, such that π_{41} is the number of firms that have 20 cents in per–share earnings surprise.

We assume that the empirical distribution is the result of a latent distribution of (unmanipulated) earnings surprise plus the effect of manipulation, or earnings management, on the part of each firm. While many firms go without managing earnings, some do so, which results in a higher earnings surprise. The latent distribution represents the number of firms that would fall into each cent bin of earnings surprise before engaging in earnings management. Let $\boldsymbol{x}$ be the model-implied latent distribution, indexed in a similar fashion, $x_1, x_2, \ldots, x_{41}$, to $\boldsymbol{\pi}$. We identify costs by the proportion of firms, in each cent bin, that manipulate earnings, and we tie this to the properties of the latent distribution that we find.

Let us examine the –20 cents bin, which represents an earnings surprise in the [–20,–19) cents interval. Firms in this bin could have only one potential simulated latent earnings surprise, and that is –20 cents. A proportion of those firms chose not to manipulate their earnings numbers, so they landed in the –20 cents bin. We call the transition probability from bin i to bin $j \geq i$, $p_{i,j}$. Therefore, we have $\pi_1 = p_{1,1}x_1$, or the number of -20 cent bin firms that did not manage earnings but rather reported latent earnings. This number is equal to the probability of a firm making the transition from the –20 cents bin to the –20 cents bin multiplied by the latent number of –20 cents bin earners. Similarly, for the number of realized –19 cents bin firms, we have $\pi_2 = p_{1,2}x_1 + p_{2,2}x_2$, equal to the number of firms that manipulate from the latent –20 cents bin, and firms that chose not to manipulate from –19 cents bin. We can see this below:

$$\begin{aligned}
\pi_1 &= p_{1,1}x_1 \\
\pi_2 &= p_{1,2}x_1 + p_{2,2}x_2 \\
\pi_3 &= p_{1,3}x_1 + p_{2,3}x_2 + p_{3,3}x_3 \\
&\vdots
\end{aligned}$$

In matrix form, we obtain

$$\begin{pmatrix} p_{1,1} & p_{2,1} & p_{3,1} & \dots & p_{T,1} \\ p_{1,2} & p_{2,2} & p_{3,2} & \dots & p_{T,2} \\ \vdots & & & \ddots & \\ p_{1,T} & p_{2,T} & p_{3,T} & \dots & p_{T,T} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{pmatrix} = \begin{pmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_T \end{pmatrix}$$

or

$$\boldsymbol{P} \cdot \boldsymbol{x} = \boldsymbol{\pi} \qquad \text{(7 revisited)}$$

—or—

(Manipulation, bin to bin) $*$ (latent earnings) = empirical earnings.

Of course, it is important to note that in our strategy representation, $p_{i,j} = 0,\ \forall j < i$; that is, firms do not manipulate earnings down relative to analysts' consensus forecast, so our matrix of $\boldsymbol{P}$ should be lower triangular:

$$
\boldsymbol{P} = \begin{pmatrix} p_{1,1} & 0 & 0 & \dots & \dots & 0 \\ p_{1,2} & p_{2,2} & 0 & \dots & \dots & 0 \\ \vdots & & & & \ddots & \\ p_{1,T-1} & p_{2,T-1} & p_{3,T-1} & \dots & p_{T-1,T-1} & 0 \\ p_{1,T} & p_{2,T} & p_{3,T} & \dots & p_{T-1,T} & p_{T,T} \end{pmatrix}
$$

To find the latent distribution, we invert the transition matrix and multiply by the empirical distribution, $\boldsymbol{x} = \boldsymbol{P}^{-1} \cdot \boldsymbol{\pi}$. We find the transition matrix $\boldsymbol{P}$ by analyzing optimal firm behavior with respect to the costs and benefits of earnings management. For each bin, b, we calculate the proportion of firms that land in each subsequent bin after managing earnings. That is, for b, we calculate $p_{b,b}$, $p_{b,b+1}$, $\dots$, $p_{b,T}$. Below, we introduce noise into the transition matrix, which complicates notation but illustrates how realized behavior can differ from strategies employed.

We calculate this proportion by simulating firm decisions to manage earnings. For each bin b, we simulate S firms. We take as given the parameters $\theta = \{\eta, \gamma, \psi^2, \zeta\}$. For each simulation, s, we draw a $\beta_{b,s} \sim U[0, 2\eta]$, the marginal cost of manipulation, and can calculate the utility of manipulation as in equation 3:

$$
u_s(b, m, \theta) = \int_{-\infty}^{\infty} \phi_{m,\theta}(\varepsilon)\mathcal{B}(b + m + \varepsilon)d\varepsilon - \beta_{b,s}\, m^{\gamma}, \tag{9}
$$

Since the real–world problem we wish to model is discrete, utility can be rewritten as:

$$
u_{s,\ Discrete}(b, m, \theta) = \sum_{-20}^{20} \phi_{m,\theta}(\varepsilon)\mathcal{B}(b + m + \varepsilon)d\varepsilon - \beta_{b,s}\, m^{\gamma}, \tag{10}
$$

where $\phi_{m,\theta}(\varepsilon)$ provides the discrete approximation of the continuous normal probability density function with mean 0, and variance $\mathbb{1}_{m \neq 0} \cdot (1 + \zeta(m - 1))\psi^2$. The optimal level of earnings management is then

$$
m^*_{b,s}(\theta) = \underset{m \in [0, T-b]}{\arg\max}\ u(b, m, \theta). \tag{11}
$$

From this, we calculate the strategy taken by each firm, from each bin b to a weakly higher bin $j \geq b$:

$$
p^{\text{strategy}}_{b,j,S}(\theta) = \begin{cases} \dfrac{\sum_{s=1}^{S} \mathbb{1}_{m^*_{b,s}(\theta) = j - b}}{S}, & \forall b, \text{ and } \forall j \geq b \\ 0, \text{ otherwise} \end{cases} \cdot
$$

However, because firms may face noise in their decisions, transitions are affected accordingly:

$$p_{b,j,S}(\theta) = \left\{ \frac{\sum_{s=1}^{S} \sum_{k=0}^{T-j} \mathbb{1}_{m^*_{b,s}(\theta)=j+k-b} \cdot \mathbb{1}_{\varepsilon_s(m^*_{b,s}(\theta),\theta)=-k}}{S}, \quad \forall b,\ j \right. .$$

where $\varepsilon_s(m,\theta)$ is a simulated outcome for a random variable following a discrete approximation of the normal distribution, $\phi_{m,\theta}(\cdot)$, where the mean is 0, the variance is $\mathbb{1}_{m\neq 0} \cdot (1+\zeta(m-1))\psi^2$, and, in equilibrium, $\forall s,\ q<0,\ m^*_{b,s}(\theta) \neq q$.

For nonzero ψ^2, the matrix is technically—though, in probability, not effectively—fully populated:

$$\boldsymbol{P}_S(\theta) = \begin{pmatrix} p_{1,1,S}(\theta) & p_{2,1,S}(\theta) & p_{3,1,S}(\theta) & \dots & p_{T,1,S}(\theta) \\ p_{1,2,S}(\theta) & p_{2,2,S}(\theta) & p_{3,2,S}(\theta) & \dots & p_{T,2,S}(\theta) \\ \vdots & & & \ddots & \\ p_{1,T,S}(\theta) & p_{2,T,S}(\theta) & p_{3,T,S}(\theta) & \dots & p_{T,T,S}(\theta) \end{pmatrix}$$

Claim 1. *If $\boldsymbol{K}$ is the true transition matrix from the latent distribution to the realized earnings surprise distribution, then*

$$\operatorname*{plim}_{S\to\infty} \boldsymbol{P}_S(\theta_0) = \boldsymbol{K}. \tag{12}$$

Proof. For each bin, i, we know that firms behave optimally. Therefore, for each $j \geq i$, we can calculate, for S, the number of firms that manipulate to j, and divide that by S to find $p_{i,j,S}(\theta)$, for some θ. If θ_0 is the true set of parameters, then as $S \to \infty$, $p_{i,j,S}(\theta_0)$ should approach $K_{i,j}$ for all j, and then for all i, according to the law of large numbers. ■

We assume this is the unique set of parameters that generates this transition matrix, or

A 1. $\theta_0 \equiv \operatorname*{arg\,min}_{\theta\in\Theta} (\boldsymbol{P}_S(\theta) - \boldsymbol{K})^2$.

We call the true earnings surprise distribution $\boldsymbol{\ell}^* = \ell^*_1, \ell^*_2, \dots, \ell^*_T$, which represents earnings before manipulation. We define the latent earnings surprise distribution as

$$\boldsymbol{K} \cdot \boldsymbol{\ell}^* = \boldsymbol{\pi},$$

thus,

$$\boldsymbol{\ell}^* = \boldsymbol{K}^{-1}\boldsymbol{\pi}. \tag{13}$$

This is the smoothest curve (as measured by the differences between bins) close to the realized empirical earnings distribution, as follows:

A 2. $\boldsymbol{\ell}^* \equiv$

$$\arg\min_{\boldsymbol{\ell} \in \mathbb{R}^+_{>0}} \left[\begin{pmatrix} \boldsymbol{\pi} \\ \boldsymbol{\ell}_{(2,\ldots,T)} \end{pmatrix} - \begin{pmatrix} \boldsymbol{\ell} \\ \boldsymbol{\ell}_{(1,\ldots,T-1)} \end{pmatrix} \right]' \Omega \left[\begin{pmatrix} \boldsymbol{\pi} \\ \boldsymbol{\ell}_{(2,\ldots,T)} \end{pmatrix} - \begin{pmatrix} \boldsymbol{\ell} \\ \boldsymbol{\ell}_{(1,\ldots,T-1)} \end{pmatrix} \right],$$

with the constraint that the mass under $\boldsymbol{\ell}^*$ is equal to the mass under $\boldsymbol{\pi}$.

Further, let

$$\boldsymbol{x}_S(\theta) \equiv \boldsymbol{P}_S(\theta)^{-1} \cdot \boldsymbol{\pi}. \tag{14}$$

Our SMM estimator finds the true parameters when the simulated distribution of latent earnings coincides with the true earnings surprise distribution. From equations 13 and 14, using these T moments, we propose our second stage estimator:

$$\hat{\theta}^S_{T,\text{ 2-stage}} \equiv \arg\min_{\theta \in \Theta} \left[\boldsymbol{\ell}^* - \boldsymbol{x}_S(\theta)\right]' \Omega \left[\boldsymbol{\ell}^* - \boldsymbol{x}_S(\theta)\right]. \tag{15}$$

Proposition 2. $\operatorname*{plim}_{S \to \infty} \hat{\theta}^S_{T,\text{ \textit{2-stage}}} = \theta_0$

Proof. Simply,

$$\begin{aligned} 0 &= \left[\boldsymbol{\ell}^* - \boldsymbol{\ell}^*\right]' \Omega \left[\boldsymbol{\ell}^* - \boldsymbol{\ell}^*\right] \\ &= \left[\boldsymbol{\ell}^* - \boldsymbol{K}^{-1} \cdot \boldsymbol{\pi}\right]' \Omega \left[\boldsymbol{\ell}^* - \boldsymbol{K}^{-1} \cdot \boldsymbol{\pi}\right]. && \text{(equation 13)} \end{aligned}$$

And, from Claim 1, we know that

$$\begin{gathered} \left[\boldsymbol{\ell}^* - \boldsymbol{K}^{-1} \cdot \boldsymbol{\pi}\right]' \Omega \left[\boldsymbol{\ell}^* - \boldsymbol{K}^{-1} \cdot \boldsymbol{\pi}\right] = \\ \operatorname*{plim}_{S \to \infty} \left[\boldsymbol{\ell}^* - \boldsymbol{P}_S(\theta_0)^{-1} \cdot \boldsymbol{\pi}\right]' \Omega \left[\boldsymbol{\ell}^* - \boldsymbol{P}_S(\theta_0)^{-1} \cdot \boldsymbol{\pi}\right], \end{gathered} \tag{16}$$

which, with Assumption 1, implies that

$$\begin{aligned} &\operatorname*{plim}_{S \to \infty} \left[\boldsymbol{\ell}^* - \boldsymbol{P}_S(\theta)^{-1} \cdot \boldsymbol{\pi}\right]' \Omega \left[\boldsymbol{\ell}^* - \boldsymbol{P}_S(\theta)^{-1} \cdot \boldsymbol{\pi}\right] \\ &\quad = \operatorname*{plim}_{S \to \infty} \left[\boldsymbol{\ell}^* - \boldsymbol{x}_S(\theta)\right]' \Omega \left[\boldsymbol{\ell}^* - \boldsymbol{x}_S(\theta)\right] && \text{(equation 14)} \\ &\quad = \operatorname*{plim}_{S \to \infty} \hat{\theta}^S_{T,\text{ 2-stage}} && \text{(equation 15)} \end{aligned}$$

is minimized uniquely at $\theta = \theta_0$. ■

We translate this two-step estimator, which first estimates $\boldsymbol{\ell}^*$, then estimates $\hat{\theta}^S_{T,\,2\mathrm{S}}$ using $\boldsymbol{\ell}^*$, into a one-step estimator. From Assumption 2 and equation 15, we have:

Proposition 3.

$$\hat{\theta}^S_T \equiv \underset{\theta \in \Theta}{\arg\min} \left[\begin{pmatrix} \boldsymbol{\pi} \\ \boldsymbol{x}_S(\theta)_{(2,\ldots,T)} \end{pmatrix} - \begin{pmatrix} \boldsymbol{x}_S(\theta) \\ \boldsymbol{x}_S(\theta)_{(1,\ldots,T-1)} \end{pmatrix} \right]' \Omega \left[\begin{pmatrix} \boldsymbol{\pi} \\ \boldsymbol{x}_S(\theta)_{(2,\ldots,T)} \end{pmatrix} - \begin{pmatrix} \boldsymbol{x}_S(\theta) \\ \boldsymbol{x}_S(\theta)_{(1,\ldots,T-1)} \end{pmatrix} \right] \tag{17}$$

is equivalent to $\hat{\theta}^S_{T,\,2\mathrm{S}}$.

Proof. The intuition is clear, given the definitions of $\boldsymbol{\ell}^*$ and $\hat{\theta}^S_{T,\,2\mathrm{S}}$. The variable $\boldsymbol{\ell}^*$ is defined as the earnings surprise distribution that balances similarity to the empirical distribution, $\boldsymbol{\pi}$ with the most similarity between bins (i.e., smoothness). Defining $\hat{\theta}^S_{T,\,2\mathrm{S}}$, above, as the model-implied latent distribution closest to $\boldsymbol{\ell}^*$ is the same as applying the same criterion to $\boldsymbol{x}_S(\theta)$. So, $\hat{\theta}^S_T$ is defined as the model-implied latent distribution that is closest to the empirical distribution while also having the smallest differences between bins. ■

Figure 1: Benefits of Earnings Surprise

This figure presents a regression discontinuity plot of cumulative earnings announcement returns around EPS surprise, focusing on the -10 to 10 cent window, where *EPS surprise* is defined as the realized earnings per share minus analysts' consensus forecast. The dependent variable is $CMAR^{3\,day}$, which is defined as three-day cumulative market-adjusted returns. The scatterplot presents gray dots that correspond to the average $CMAR^{3\,day}$ within each $EPS\,Surprise$ bin, and two fitted polynomials in black, which represent the best-fitting (both second degree) polynomials on either side of the EPS surprise cutoff of zero cents. The intersection of the fitted polynomials at $EPS\,Surprise = 0$ represents our preferred discontinuity estimate of the benefits of meeting analysts' consensus EPS forecast. Note that the estimation uses the full -20 to 20 cent window.

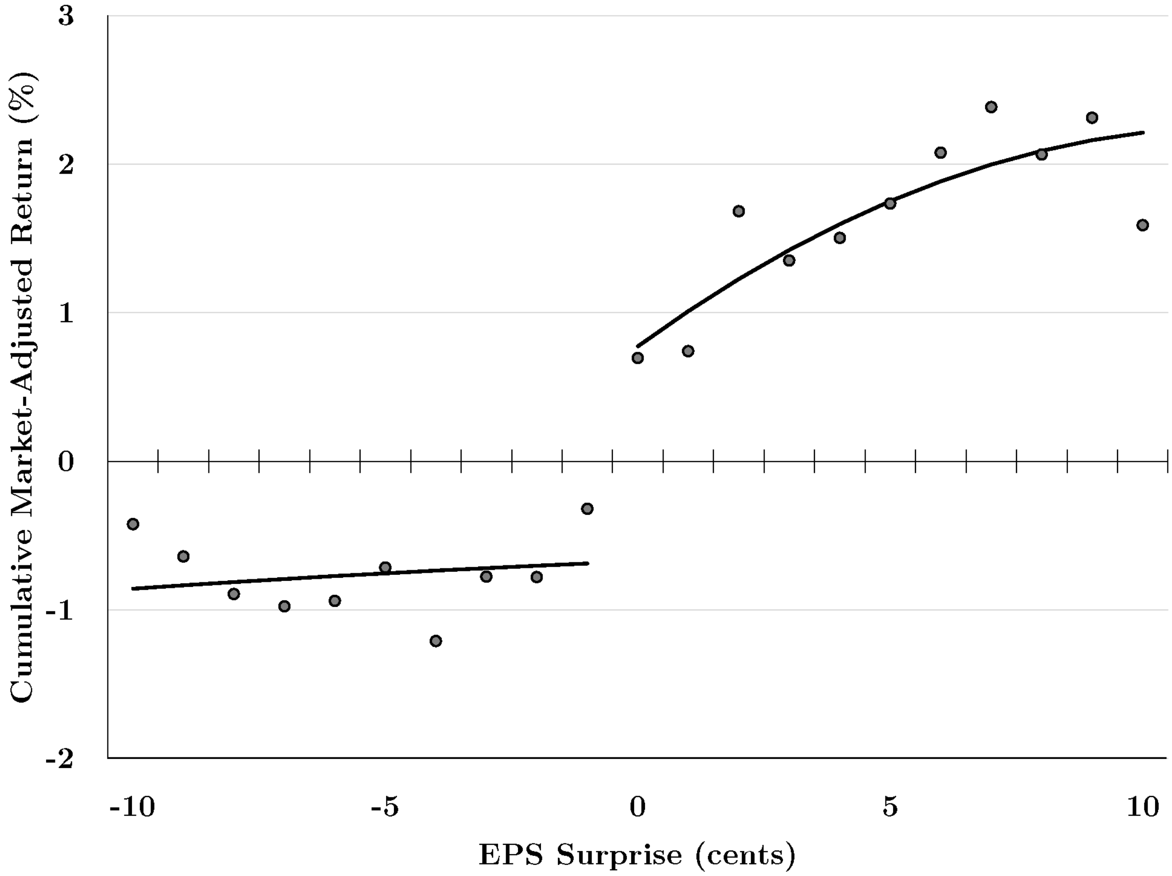


Figure 2: Fitted EPS Surprise Frequencies

This figure presents a bar plot of the fitted distribution of EPS surprise, focusing on the −10 to 10 cent window, where *EPS surprise* is defined as the realized earnings per share minus the analysts' consensus forecast. The distribution is fitted using a McCrary [2008] test with best-fitting polynomial functions of the density on either side of zero EPS surprise, which are presented in gray. The difference between the bar at $EPS\,Surprise = -1$ and the bar at $EPS\,Surprise = 0$ represents our preferred diagnostic test statistic of the prevalence of earnings management around the zero earnings surprise cutoff. Note that the estimation uses the full −20 to 20 cent window.

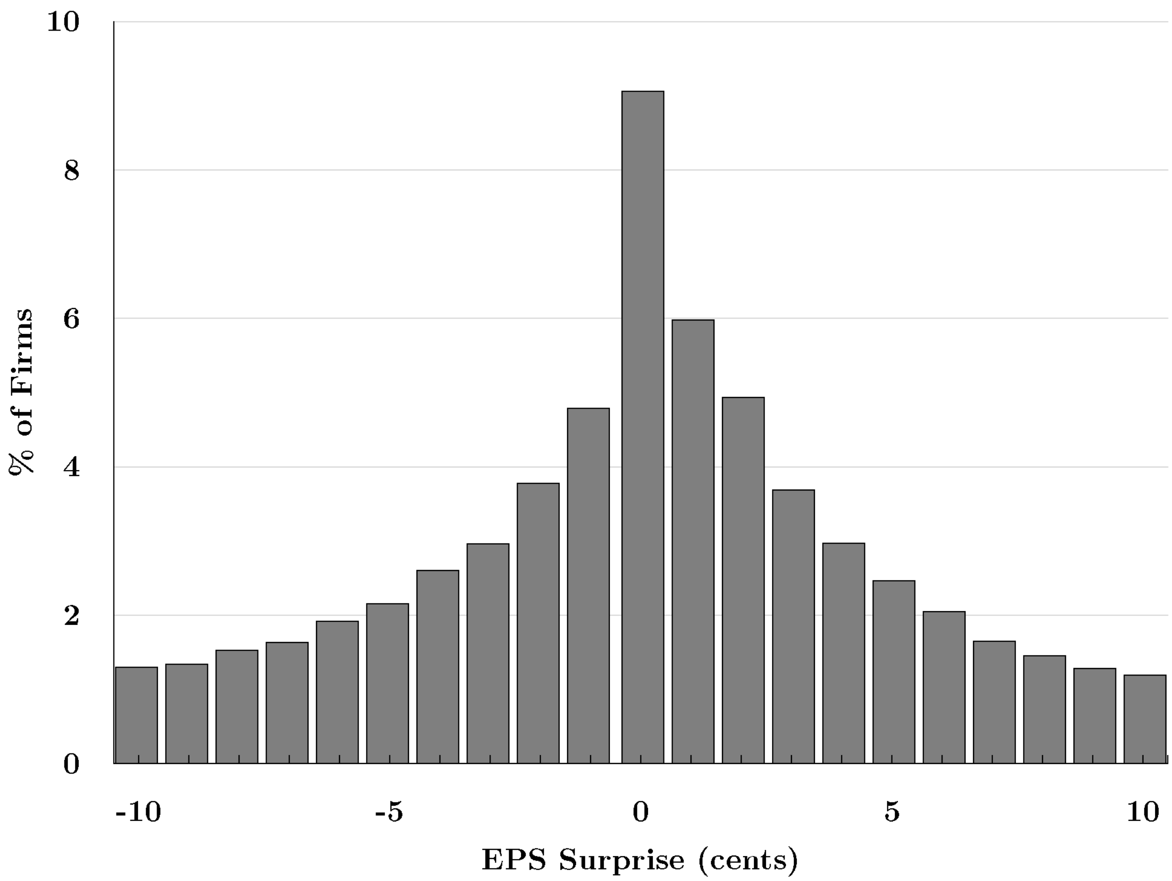


Figure 3: Latent and Empirical Distributions

This figure presents the empirical and latent distributions of EPS surprise, focusing on the -5 to 5 cent window, where *EPS surprise* is defined as the realized earnings per share minus the analysts' consensus forecast. The empirical distribution (dark gray) is estimated using a regression discontinuity approach (McCrary [2008]), which allows for a discontinuity in the frequencies of observations around the zero earnings surprise cutoff. The latent distribution (light gray) is an output of our structural estimation that accounts for managers' optimal manipulation decisions and trades off distance from the empirical distribution for smoothness. Note that the estimation uses the full -20 to 20 cent window.

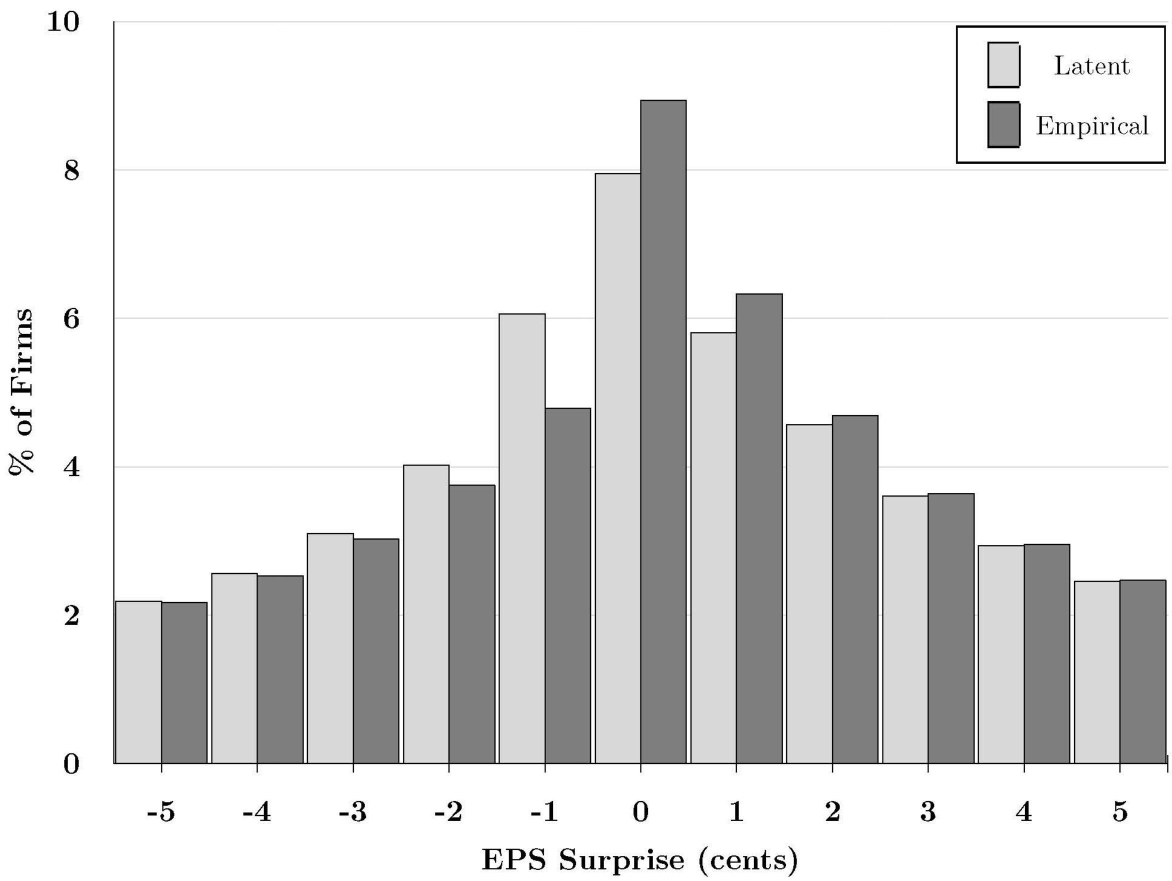
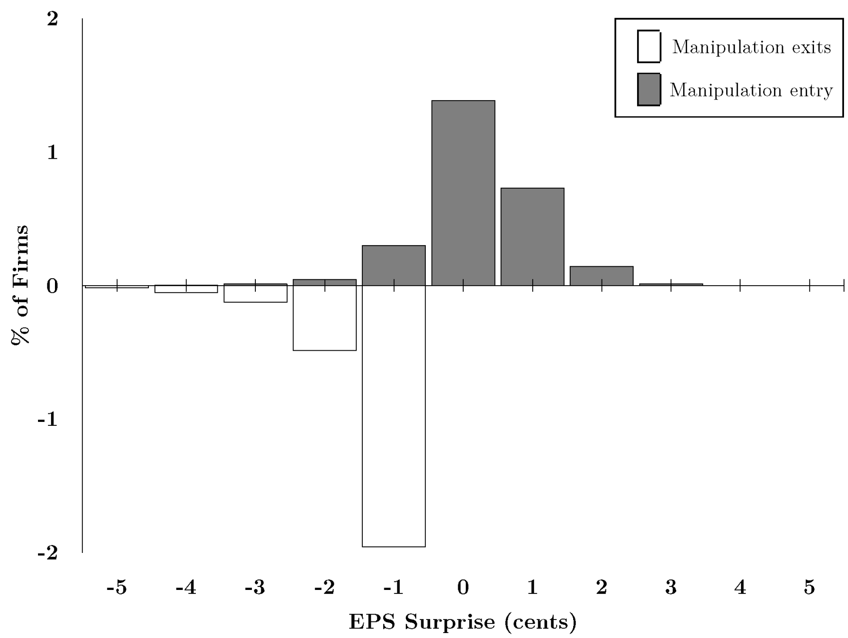


Figure 4: Cent Manipulation

This figure illustrates the effects of the equilibrium manipulation strategies on the proportion of firms in each EPS surprise bin, focusing on the -5 to 5 cent window, where *EPS surprise* is defined as the realized earnings per share minus the analysts' consensus forecast. Panel (a) plots the fraction of firms manipulating into and out of each bin. Panel (b) shows the net effect of manipulation strategies on the earnings surprise distribution. Panel (b) also highlights how much manipulation was required for firms entering each bin. Note that the estimation uses the full -20 to 20 cent window.

(a) Manipulation exits and entries



(b) Manipulation strategies

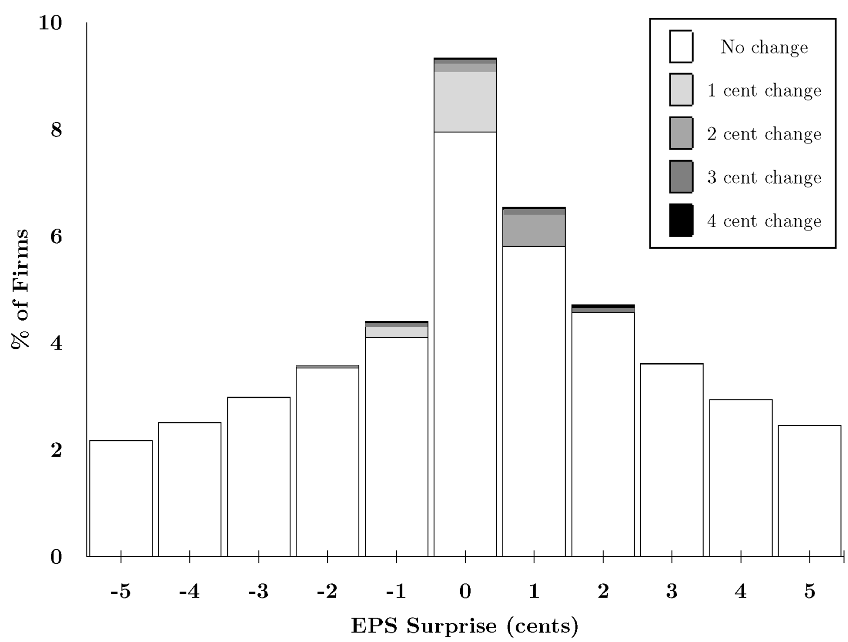


Figure 5: Counterfactual Parameters

This figure contrasts the empirical EPS surprise distribution (where *EPS surprise* is defined as the realized earnings per share minus the analysts' consensus forecast) with counterfactual distributions of EPS surprise that arise by reducing each estimated structural parameter to 10% of its estimated value and then recalculating optimal manipulation. The dashed lines represent counterfactual empirical distributions, and the solid lines represent the empirical distribution. These counterfactuals provide intuition for the role of each cost function parameter in shaping earnings manipulation behavior. Panels (a), (b), (c), and (d) present the counterfactuals of setting η, γ, ψ^2, and ζ to 10% of estimated values, respectively.

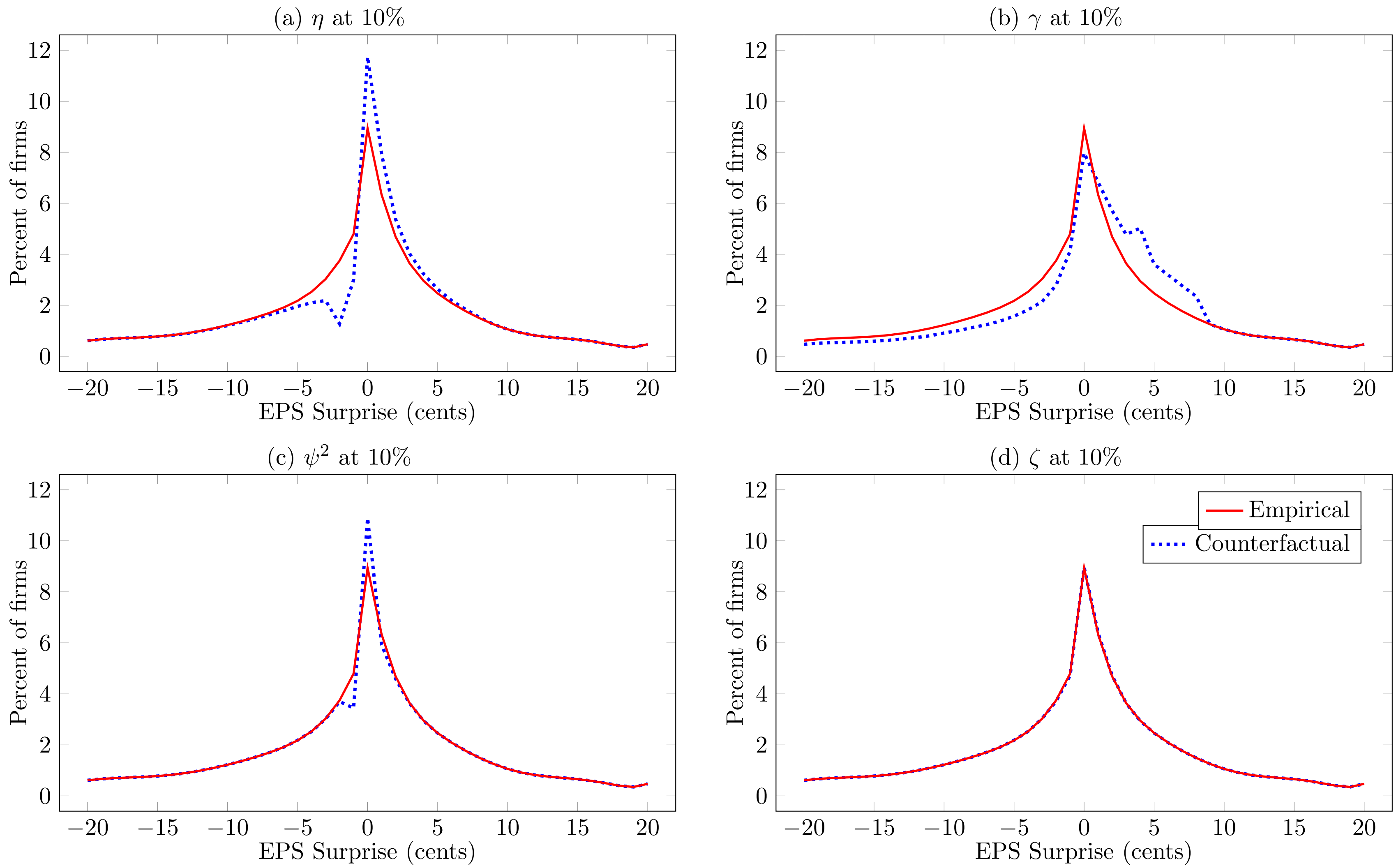


Table 1: Summary Statistics

This table presents summary statistics (i.e., mean, standard deviation, 25th percentile, median, and 75th percentile) of key characteristics used in our regression discontinuity tests to follow. We define $EPS\,Surprise$ as the realized earnings per share minus the analysts' consensus forecast, in cents. For cases in which an analyst issues multiple EPS forecasts during the forecast period, we include the latest forecast that precedes the earnings announcement so that our average incorporates one forecast for each analyst. $CMAR^{3\,day}$ is the three-day cumulative market-adjusted return around the earnings announcement. The underlying data come from I/B/E/S and CRSP.

	Mean	SD	25%	Median	75%
$EPS\,Surprise$	-0.05	7.29	-3	1	3
$EPS\,Surprise \geq 0$	62.43%	-	-	-	-
$CMAR^{3\,day}$	0.55	8.33	-2.80	0.46	3.81

Table 2: The Benefits of Beating Earnings Benchmarks

This table presents regression discontinuity tests of earnings announcement returns around earnings benchmarks. The running variable is $EPS\,Surprise$, which is defined as realized earnings-per-share (EPS) minus the analysts' consensus EPS forecast, and the dependent variable is $CMAR^{3\,day}$, which is defined as three-day cumulative market-adjusted returns. We estimate the discontinuity in $CMAR^{3\,day}$ at the cutoff $EPS\,Surprise = 0$. Columns (1) and (2) use global polynomial controls. Columns (3) and (4) use local linear controls with bandwidth restrictions of 10 bins. Robust standard errors are clustered at the firm level and are reported in parentheses. ***,**, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable: $CMAR^{3\,day}$				
	Global Polynomial		Local Linear	
	(1)	(2)	(3)	(4)
$EPS\,Surprise \geq 0$	1.447**	1.514***	1.620***	1.733***
	(0.168)	(0.198)	(0.148)	(0.179)
Fixed effects:				
Firm		X		X
Year-quarter		X		X
Polynomial degree	2	2	1	1
Bandwidth	—	—	10	10
R^2	0.0241	0.0779	0.0209	0.0773
Observations	49,604	44,535	41,428	36,680
Bins	41			

Table 3: Distributional Evidence of Manipulation Around Earnings Benchmarks

This table presents McCrary [2008] tests of manipulation of EPS around analysts' consensus earnings forecasts. The running variable is $EPS\,Surprise$, which is defined as reported earnings minus analysts' consensus earnings forecast, and the dependent variable is the frequency of observations in each earnings surprise bin. We estimate the discontinuity in frequencies at the cutoff $EPS\,Surprise = 0$. Column (1) uses global polynomial controls, and column (2) uses local linear controls with bandwidth restrictions. Robust standard errors are presented in parentheses. ***,**, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable: π_i		
	Global Polynomial	Local Linear
	(1)	(2)
$EPS\,Surprise \geq 0$	2.631***	2.322***
	(0.555)	(0.992)
Polynomial degree	6	1
Bandwidth	—	10
R^2	0.9899	0.8378
$Bins$	41	

Table 4: Baseline Estimates

This table presents our structural estimates as well as statistics regarding the percentage of manipulators ($\%manip$), the expected amount of manipulation by manipulators ($\overline{manip}$), and the fraction of manipulators shocked out of their target bin by noise ($\%noise$). Bootstrapped standard errors are presented below each parameter estimate.

Marginal cost η	Curvature γ	Variance ψ^2	Variance multiplier ζ	$\%manip$	$\overline{manip}$	$\%noise$
0.0161	2.0817	0.8201	3.7105	2.62%	1.21	59.6%
(0.0004)	(0.0517)	(0.0236)	(0.0941)			

Table 5: Counterfactuals

This table shows parameter estimates and manipulation statistics based on counterfactual simulations or alternative empirical distributions.The first row replicates the baseline results from Table 4. The next four rows present the results of simulations that involve, one at a time, reducing the value of each parameter to 10% of its estimated value, then recalculating the optimal manipulation behavior. The last two rows use alternative empirical distributions: *Reduce peak* takes the actual empirical distribution and reduces the peak found in the [0,1) cent bin by half relative to the [1,2) cent bin. *Kill peak* brings the frequency in the [0,1) bin down to the frequency in the [1,2) cent bin.

	Marginal cost η	Curvature γ	Variance ψ^2	Variance multiplier ζ	$\%manip$	$\overline{manip}$	$\%noise$
Baseline	0.0161	2.0817	0.8201	3.7105	2.62%	1.21	59.6%
η at 10%	*0.0016*	2.0817	0.8201	3.7105	12.6%	1.92	64.3%
γ at 10%	0.0161	*0.2082*	0.8201	3.7105	9.45%	10.16	78.4%
ψ^2 at 10%	0.0161	2.0817	*0.0820*	3.7105	3.22%	1.23	6.9%
ζ at 10%	0.0161	2.0817	0.8201	*0.3711*	2.56%	1.29	57.8%
Reduce peak	0.0205	2.1867	0.8178	2.2439	1.91%	1.17	73.5%
Kill peak	0.0273	2.0523	1.1699	0.9738	1.36%	1.20	64.3%

Table 6: Estimation Robustness

This table shows parameter estimates and manipulation statistics for a range of alternative samples and estimation approaches. The first row replicates the baseline results from Table 4. The next three rows use narrower windows of earnings surprise when estimating the model inputs. The bottom three rows all use the baseline model inputs, but they vary the features of the estimation procedure. First, we increase the weight on the closeness to the empirical distribution criterion by 10%. Then, we increase the weight on the smoothness criterion by 10%. The last row forces the variance multiplier parameter (ζ) to be zero.

	Marginal cost η	Curvature γ	Variance ψ^2	Variance multiplier ζ	$\%manip$	$\overline{manip}$	$\%noise$
Baseline	0.0161	2.0817	0.8201	3.7105	2.62%	1.21	59.6%
-15 to 15 cents	0.0157	2.2398	0.7761	3.8515	2.71%	1.15	55.7%
-10 to 10 cents	0.0154	1.9381	0.7942	3.9199	2.84%	1.27	60.0%
-5 to 5 cents	0.0167	2.0950	0.7599	4.1904	2.56%	1.19	57.3%
Overweight empirical	0.0170	2.2847	0.7992	2.6322	2.45%	1.15	57.3%
Overweight smoothness	0.0146	1.8871	0.7901	2.5761	3.11%	1.31	58.9%
$\zeta = 0$	0.0211	2.0907	0.2983	0	2.29%	1.29	26.4%

Table 7: Benefits Heterogeneity

This table presents parameter estimates and manipulation statistics for a range of methods for capturing heterogeneity and correlation with costs in the benefits function. The first row replicates the baseline results from Table 4. In the first set of alternative tests, we allow for heterogeneity in benefits. For the rows labeled "Normal," we assume that firms draw a random benefit that is normally distributed according to the mean and standard deviation of our preferred benefit estimate (Table 2, Column 1). For the rows labeled "Bernoulli," we assume that firms draw a random benefit that is either "high" or zero, with standard error as before, where "high" is assigned to maintain the same mean benefit as in the baseline. In the second row for each distribution, we multiply standard errors on benefits by a factor of five. In the last four rows, we use the normally distributed benefits, and we allow for a correlation with costs of either 0.5 or −0.5 (and again with alternate specifications using inflated standard deviations).

	Marginal cost η	Curvature γ	Variance ψ^2	Variance multiplier ζ	$\%manip$	$\overline{manip}$	$\%noise$
Baseline	0.0161	2.0817	0.8201	3.7105	2.62%	1.21	59.6%
Normal	0.0148	2.1297	0.7600	4.5100	2.88%	1.19	57.5%
Normal (x5 s.e.)	0.0152	1.9273	0.8068	4.4838	2.85%	1.31	59.7%
Bernoulli	0.0153	1.9731	0.8045	3.4927	2.81%	1.27	60.1%
Bernoulli (x5 s.e.)	0.0151	2.0110	0.8004	3.0636	2.83%	1.28	59.6%
0.5 correlation	0.0153	2.0135	0.7930	4.4206	2.84%	1.23	59.3%
0.5 correlation (x5 s.e.)	0.0141	1.9505	0.8178	4.6580	3.20%	1.31	59.7%
–0.5 correlation	0.0160	2.2237	0.7945	3.5933	2.64%	1.16	58.9%
–0.5 correlation (x5 s.e.)	0.0173	2.1715	0.7854	4.2884	2.45%	1.19	58.8%

Table 8: Identifying Manipulators by Earnings Surprise Bins

This table illustrates the performance of various definitions of "suspect" bins on identifying manipulation. The second set of percentages weights manipulation by the amount of manipulation, i.e. a firm manipulating two cents receives twice the weight of a firm that only manipulates by one cent.

	$[\mathbf{0}, \mathbf{1})$	$[0, 2)$	$[0, 3)$	$[-1, 1)$	$[-1, 2)$	$[-1, 3)$
% Manipulating	**11.0%**	9.9%	8.1%	8.7%	8.5%	7.3%
% of Manipulators	**53.1%**	80.9%	86.4%	64.2%	92.0%	97.4%
% Manipulating (weighted)	**12.8%**	14.2%	12.4%	10.2%	12.1%	11.1%
% of Manipulators (weighted)	**39.6%**	77.3%	88.3%	48.3%	86.0%	97.0%

Table 9: SOX and the Cost of Earnings Management

This table shows how the costs of earnings management changed around SOX. The pre-SOX period is 1999–2001, and the post-SOX period is 2002–2004. The variable Δ is the estimated discontinuity in the earnings distribution, estimated as in Table 3. The variable $\mathcal{B}$ is the benefit of meeting the analyst benchmark, estimated as in Table 2. η, γ, ψ^2, and ζ are the estimated parameters of the cost function. The final three rows describe the estimated equilibrium manipulation in each period.

	Pre-SOX	Post-SOX	Diff.	p-value
Δ	4.196	3.012	–1.184	0.008
$\mathcal{B}$	1.229	2.027	0.798	0.034
η	0.0097	0.0278	0.0181	$<$0.0001
γ	1.942	2.808	0.865	$<$0.0001
ψ^2	0.646	0.202	–0.444	$<$0.0001
ζ	3.040	1.316	–1.724	$<$0.0001
$\%manip$	4.05%	2.61%	–1.44%	-
$\overline{manip}$	1.29	1.06	–0.23	-
$\%noise$	54.2%	16.0%	–38.1%	-