

Estimating the marginal cost of manipulation

Jeremy Bertomeu*, Edwige Cheynel†, Edward Li* and Ying Liang*

Preliminary and incomplete.

Abstract

We derive a structural estimator of earnings manipulation based on signaling theory, incorporating cross-sectional properties of earnings and prices. Identification requires partial knowledge of the distribution of true earnings and non-linearities in the relationship between earnings and prices. As an application, we estimate manipulation costs and reporting biases before vs. after the Sarbanes-Oxley Act of 2002. We also characterize implied changes in reporting bias across industries, subsamples of firm size and growth opportunity and over time, and propose an identification strategy that relies on exogenous shocks to misreporting costs. Our empirical model extends to other settings provided reporting benefits are observable to the researcher, such as environments with costly communication.

Keywords: manipulation; discontinuity; earnings; reporting; signalling.

JEL classification: D83; G14; M4.

*J. Bertomeu and E. Li, Associate Professors, and Y. Liang, PhD student, are at Baruch College, City University of New York, One Bernard Baruch Way, New York, NY 10010. J. Bertomeu acknowledges financial support from the PSC-CUNY research fund. Contact Author: jeremy.bertomeu@gmail.com.

†E. Cheynel is an Associate Professor at Columbia Business School, Columbia University, 3022 Broadway, New York, NY 10027.

1. Introduction

How to separate the true earnings process from reported earnings is a fundamental question in accounting research. A vast literature develops a variety of methods to detect and measure earnings manipulation (e.g., Jones 1991; Dechow and Dichev 2002; Burgstahler and Dichev 1997; Degeorge, Patel and Zeckhauser 1999). However, considerable debates exist about the validity of these empirical methods (Dechow, Ge and Schrand 2010). To answer this question, researchers make theoretical restrictions about the true and manipulated earnings processes, which identify one in isolation of the other. However, because these identifying restrictions are rarely stated explicitly, there is continuing disagreement in the literature about the validity of empirical proxies, as well as the existence and magnitude of manipulation.

This paper develops an alternative approach that relies on deriving estimates of a primitive driver of manipulation, the marginal cost of manipulation, from restrictions implied by a simple signalling model on the cross-section of earnings and price. In short, we use a simplified version of the Dye (1988) and Stein (1989) models to recover firm-level estimates of earnings manipulation. Manipulation is identified from economic assumptions about reporting choices - thus, allowing researchers to evaluate, modify or extend the identifying assumptions based on a description of the choice process. The estimation procedure is simple to implement on any statistical package, in closed-form, and does not require researchers to solve for the complete equilibrium of the signalling game. We ask two questions. First, which minimal assumptions about true earnings are required to identify manipulation in the cross-section? Second, how do we recover the expected bias in the cross-section, as well as the bias conditional on a reported earnings? To illustrate a use of this model, we estimate the effects of the Sarbanes-Oxley (SOX) Act of 2002 on the marginal cost of manipulation.

In our model, managers report to increase price and, therefore, manipulation is increasing in the price response. This allows us to map a distribution of true earnings into a distribution of reported earnings by adding the model-predicted manipulation. We invert this mapping to find the distribution of true earnings that best fits the (observed) distribution of reported earnings, solving for the marginal cost of manipulation that fits observed reports. From this cost, we can provide estimates of the amount of reporting bias as well as conditional bias for any reported earnings.

A natural challenge to inferring manipulation is whether data about earnings and price contains enough information to separate true from reported earnings. In particular, like other approaches in the literature, our model requires some identifying a-priori restrictions on true earnings. Our incremental contribution to existing proxies is not in the need for

strong assumptions which as great (if not greater) in our identification strategy, but in that these assumptions are stated on an explicit choice made by the manager. Also, one benefit in our procedure is that it does not imply, by construction that earnings manipulation exist, and we can empirically test whether misreporting costs are large and reporting biases have disappeared in a given data set.

Like other studies, without restrictions on the true earnings process, any reported earnings data might be consistent with a world without manipulation (Dechow, Ge and Schrand 2010). In prior structural models of manipulation, for example, identification is in part obtained from parametric (normal or lognormal) assumptions on economic shocks as well as, in certain cases, the timing of reversals and/or assumptions about earnings and price (Gerakos and Kovrijnykh 2013, Beyer, Guttman and Marinovic 2012, Zakolyukina 2014, Terry 2015, Nikolaev 2014).

In our model, we recover, under certain conditions, the implied earnings management with some knowledge of the shape of true earnings: identification does not require earnings to be necessarily normally (or lognormally) distributed or the association between price and earnings to be linear. Thus, one benefit is that we can let the empirical data inform us about the distribution of reported earnings and relations between price and earnings, plugging it as an input to the estimation procedure. Our procedure may thus guide researchers to input into their model a distribution of true earnings derived from a specified investment model which does not necessarily imply symmetric earnings (Breuer and Windisch 2017, Hemmer and Labro 2015, Hiemann 2016, Dutta and Nezlobin 2016).¹

We develop an application of the theory by estimating manipulation pre-SOX vs. post-SOX: first, under the baseline simplest possible assumption, namely, that true earnings are normally-distributed, and second, under weaker identifying assumptions such as symmetry or a shock to the manipulation cost. Manipulation costs and the average bias decrease post SOX; while this is consistent with insights from other statistical proxies, we derive this finding from a methodology based on choice. Also, our method allows us to quantitatively assess the reduction in reporting biases and if misreporting has disappeared after SOX. In various specifications, the reporting bias appears to have sharply decreased after SOX, from about 5% in ROE surprise pre SOX to below 1% after SOX. In some specifications, we were unable to establish any bias based on observed earnings data after SOX. We also decompose the reduction bias pre vs. post SOX in terms of a change in cost and a

¹We develop this in the econometric model. However, the theory of the distribution of earnings in the presence of investment problems is still in its development stage, i.e., there exists no simple characterization of the set of distributions that would be consistent with such problems, so we do not implement this in our empirical application. Nevertheless, we hope that future research might be able to input a distribution implied by these theories. Note, importantly, that while investment may generate skewness, investment-based theories do not predict that any distribution would be explained by the investment problem.

change in price response to earnings, finding that most of the change is driven by changes in misreporting costs. Estimates by industry, and per subsample of size and growth opportunities (market to book), are consistent with these conclusions. Large firms and firms with high growth opportunities have lower manipulation cost and exhibit higher levels of reporting bias, but also the greater effect from the regulation. We also find that healthcare industry features the largest reduction in reporting biases and find reductions in biases across all industries post SOX.

Literature Review. An extensive foundational literature on earnings manipulation is usually traced to the models of Dye (1988) and Stein (1989), which adapt the Spence education signaling model to a firm biasing a public report to increase its short-term market price. Since these early studies, the theory is key to explaining the possible existence of manipulation in equilibrium, even if investors rationally correct for it. A direct test of signaling is difficult because reported earnings convey both fundamental information and a bias - with a few notable exceptions, we are not aware of many feasible tests of this theory.

Within studies that use formal theory to motivate measures, there are two strands of approaches. The first approach is based on identifying accounting noise (possibly due to manipulation) from assumptions about the time-series of economic earnings, as in Zakolyukina (2014), Nikolaev (2014), Beyer, Guttman and Marinovic (2014) and Terry (2015). This approach is, probably, the most conceptually attractive method to identify manipulation, because it builds on the time-series properties of the accounting process; however, it is also more difficult to implement because it requires an understanding of the time-series properties of economic shocks and incentives.

The second approach is based on identifying manipulation from cross-sectional comparisons between firms, leaving aside any of issue of time-series variation. This approach is, for example, implicit in Burgstahler and Dichev (1997) and presented in the companion model in Degeorge, Patel and Zeckhauser (1999): if the payoff to the manager is a single bonus conditional on attaining a threshold, we should observe manipulation to meet the threshold but (also) downward manipulation far away from the bonus threshold.

Our methodology borrows from both methods. As in the first approach, we draw from observing the relation between earnings and price information about manipulation incentives. Unlike this literature, however, we principally rely on cross-sectional data; this is a caveat but also makes the estimation easy to implement empirically without requiring researchers to formulate and estimate a dynamic choice process.² As in the second ap-

²Having noted this simplification, we will show later on how the framework may be extended to capture manipulation over time, i.e., by estimating the dynamic price response function. A complete analysis of

proach, we draw information from the distribution earnings – specifically, more frequent earnings at points with more price response – to estimate the level of manipulation in the cross-section. On the other hand, we do not require the distribution of earnings to feature discontinuities or to use discontinuity points as proxies for manipulation.

2. The Model

2.1. Overview

We first lay out the four principal assumptions that underlie our estimation, and later on, write these assumptions formally in terms of a formal model.

A1. Managers receive private information about current true earnings, informative about the firm’s economic value, as well as (possibly) their manipulation cost. True earnings are drawn from a continuous distribution.

A2. Managers can issue a manipulated accounting report, for a personal cost increasing in the difference between economic earnings and the accounting reports.

A3. Competitive investors value the firm as a function of reported earnings.

A4. Managers maximize current stock price minus manipulation costs.

Our theoretical framework closely follows the Spence signaling framework and we will later show how this can be extended to noisy signaling (Fischer and Verrecchia 2000; Dye and Sridhar 2004; Frankel and Kartik 2014) without adding much complexity to the estimation method. With only uncertainty about true earnings (our baseline), markets can infer the manipulation from reported earnings (Dye 1988; Stein 1989) while, if there is also uncertainty about the cost of manipulation, differences in reporting strategies among managers with different cost may introduce additional noise in reported earnings so that markets might only recover manipulation in expectation.

2.2. Known manipulation cost

We first introduce the model with a common-knowledge manipulation cost. The firm privately observes a fundamental signal about the value of the firm x (*A1*), hereafter

the dynamic version of our model, however, goes somewhat beyond our current perspective and present some of its own implementation challenges (such as how to estimate policies conditional on past history or which speed of reversals to adopt).

true earnings, drawn from a distribution with full support on \mathbb{R} , p.d.f. $f(\cdot)$ and c.d.f. $F(\cdot)$, and has a personal cost of manipulation (i.e., the time and effort spent in finding obfuscation opportunities, as well as costs related to potential SEC sanction, litigation, or disturbing optimal business operation) defined as a constant cost θ^{-1} (A2) which, for now, we assume is a constant. Note that, for convenience, we use θ^{-1} for the cost so that we may state an infinite cost (perfect enforcement) in terms of a finite parameter $\theta = 0$ (hence, $\theta^{-1} = \infty$) for which we can bootstrap finite standard-errors.

As with many competitive equilibrium models, we assume that the firm's manager reports r , hereafter reported earnings, to maximize its current value as evaluated by investors who only have access to public information. Let $R(x) \in \mathbb{R}$ denote the firm's reporting strategy, and $\gamma : \mathbb{R} \rightarrow \mathbb{R}$ denote a mapping that associates a price $\gamma(r)$ to each report r . For any conjectured increasing reporting strategy $\bar{R}(x)$ and observed report r , investors respond to earnings with

$$\bar{\gamma}(r) = \mathbb{E}(\alpha(\tilde{x}) | \bar{R}(\tilde{x}) = r), \quad (2.1)$$

where $\alpha(\cdot)$ is a continuous function representing the mapping between true earnings and value (A3).

The manager is interested in maximizing market reaction, net of costs (A4). That is, the manager's reporting strategy solves

$$R(x) \in \operatorname{argmax}_r \quad \bar{\gamma}(r) - \frac{1}{\theta} \psi(r - x), \quad (2.2)$$

where $\psi(\cdot)$ is a twice-differentiable convex function with its minimum set at $\psi(0) = \psi'(0) = 0$.³

Definition 2.1 *A fully-separating signaling equilibrium is defined as an increasing reporting function $R(x)$ and market reaction $\gamma(x)$ such that:*

- (i) *investors price the firm according to correct equilibrium conjectures, that is, equation (2.1) is satisfied at $\bar{R}(x) = R(x)$, whenever possible;*
- (ii) *managers optimally select their reporting strategy, that is, equation (2.2) is satisfied for any x at $\bar{\gamma}(x) = \gamma(x)$.*

Note that we focus here on fully-separating equilibria; this model has the single-crossing property of standard Spence signaling which, under traditional refinements (Cho

³We use $1/\theta$ so that arbitrary large costs can be stated as a finite parameter $\theta = 0$. Alternatively, as in Fischer and Verrecchia (2000), the utility function can be rescaled by multiplying by θ so that the manager achieves $\theta\bar{\gamma}(r) - \psi(r - x)$, hence, θ is interpreted as a price benefit.

and Kreps 1987), selects the full-separating equilibrium when it exists. Further, pooling equilibria, if we select them, predict holes in the support of the distribution of reports (Guttman, Kadan and Kandel 2006, 2010).⁴

In an equilibrium, the first-order condition on the manager's problem in (2.2) is

$$\gamma'(R(x)) = \frac{1}{\theta}\psi'(R(x) - x), \quad (2.3)$$

and the second-order condition is satisfied as long as θ is small enough, that is, as long as the solution to (2.3) satisfies

$$\theta \leq \frac{\psi''(R(x) - x)}{\gamma''(R(x))}. \quad (2.4)$$

In general, whether the second-order condition is satisfied depends on the conditions on the primitives of the model (such as distributions and the market pricing of true earnings $\alpha(\cdot)$) or, as we will consider in extensions of this setting, uncertainty about θ . For our current purpose, however, we shall assume that the primitives are such that the second-order condition holds and then let the data inform us about $\gamma(\cdot)$. Of course, we also know (empirically) that the second-order condition holds if there are no holes in the distribution since a violation of (2.4) - if, say, the relationship between price and reports is too steep relative to the cost - would imply that some reports are never observed on the equilibrium path.

Equation (2.3) simply states that the marginal benefit of manipulation $\gamma'(\cdot)$, on the left-hand side, must equal the cost $\frac{1}{\theta}\psi'(R(x) - x)$, on the right-hand side. This expression is intuitive: given any earnings realization x , the equilibrium manipulation $R(x) - x$ will be higher if the market response $\gamma(\cdot)$ is more sensitive to reported earnings. To give an example, if we set the cost of manipulation $\psi(x) = x^2$ as quadratic, then the bias for a given earnings report is

$$R(x) - x = \frac{\theta}{2}\gamma'(R(x)) \quad (2.5)$$

and, therefore, is proportional to the slope of the relation between earnings and price. As is well-known, if $\gamma(\cdot)$ is linear, then the bias is solely a function of $1/\theta$ and is constant in reported earnings (Dye 1988; Stein 1989). Other settings such as Einhorn and Ziv (2012) imply that the equilibrium market response would be convex if the $\alpha(\cdot)$ is linear and true earnings are bounded from below. In the settings of Subramanyam (1996) and Kirschenheiter and Melumad (2002), $\alpha(\cdot)$ is neither convex nor concave as a result of updating the precision of earnings on a fundamental latent process, which would imply

⁴These models can fit a continuous distribution of earnings without hole, but require estimating additional noise in the report, thus adding another source of heterogeneity to our problem. While this could be envisioned in the empirical model (e.g., estimating the location of one or more pooling region as well as the noise process), it goes beyond our current objective.

that $\gamma(\cdot)$ might be neither convex nor concave as well. For this study, because $\gamma(\cdot)$ can be directly estimated from reported earnings and prices, we do not impose theoretical restrictions on $\alpha(\cdot)$.

One problem with (2.3) is that it identifies manipulation up to a proportionality factor θ which we need to estimate. To identify manipulation levels, this is where some a-priori knowledge of the shape of true earnings comes into play. Simply put, if we make a guess about θ , we can reconstruct from (2.3) a predicted distribution of reported earnings from an assumed a-priori shape of true earnings. A good choice of θ should recover a distribution of reported earnings that is as close as possible from the distribution of reported earnings observed empirically.

Formally, let us write the distribution of reported earnings $r = R(x)$, with p.d.f. $g(r)$, in terms of the manipulation cost and $\gamma(\cdot)$. It is well-known that the density of a random variable (r) that is a function of another (x) is

$$g(r) = \frac{1}{|R'(R^{-1}(r))|} f(R^{-1}(r)), \quad (2.6)$$

where, in this type of model, a separating equilibrium implies that the reporting strategy is increasing and differentiable so that we can rewrite $|R'(R^{-1}(r))| = R'(R^{-1}(r))$.

We can further simplify this equation by solving for $R'(\cdot)$ applying the implicit function theorem on equation (2.3),

$$R'(x) = -\frac{\theta^{-1}\psi''(R(x) - x)}{\hat{\gamma}''(R(x)) - c\psi''(R(x) - x)}. \quad (2.7)$$

Substituting in (2.6), the model-predicted p.d.f. of the accounting reports is

$$g(r) = \frac{\theta^{-1}\psi''(r - R^{-1}(r)) - \gamma''(r)}{\theta^{-1}\psi''(r - R^{-1}(r))} f(R^{-1}(r)). \quad (2.8)$$

Equation (2.8) is the (partial) likelihood of an observed report r and points to a natural estimation procedure by maximizing $\sum_{i=1}^n \ln \hat{g}(r_i)$ for any sample $(r_i)_{i=1}^n$ where $\hat{g}(\cdot)$ is obtained by substituting in an estimate $\hat{\gamma}''(\cdot)$ of the theoretical $\gamma(\cdot)$. For example, the researcher can conduct a consistent non-parametric fitting procedure of observed prices on reported earnings, to obtain $\hat{\gamma}$ and numerically differentiate the estimated function twice to recover $\hat{\gamma}'(\cdot)$ and $\hat{\gamma}''(\cdot)$. Once we have constructed an estimator for $\hat{\theta}$, we can recover the implied manipulation for any given period as

$$\hat{b}(r) = r - \hat{R}^{-1}(r) = (\psi')^{-1}(\hat{\theta}\hat{\gamma}'(r)). \quad (2.9)$$

The estimates $\hat{b}(r)$ reflects both the true manipulated amount and (in addition) how

much the market believes has been misstated. Note also that the primitive of the model $\alpha(x)$ can then be estimated from the equilibrium restriction (ii) in definition 2.1 as $\hat{\alpha}(x) = \hat{\gamma}(\hat{R}(x))$, where $\hat{R}(x)$ is the estimated reporting function and is obtained using the estimated price response in (2.3).

2.3. Identification

In appendix A.1., we offer an illustrative example of the estimation with simulated data. In this section, we briefly discuss settings in which disclosure costs may or may not be identified. Formally, let the econometric model $\zeta \in \mathcal{A}$ be denoted as (i) a set of primitives $(\theta, f(\cdot), \alpha(\cdot))$, where $1/\theta$ is the cost, $f(\cdot)$ is the p.d.f. of the true earnings \tilde{x} , $\alpha(\cdot)$ is the pricing function for true earnings, and (ii) a set of observables $(g(\cdot), \gamma(\cdot))$ satisfying the equilibrium conditions (2.6)-(2.8). The model is identified if, given some observables $(g_0(\cdot), \gamma_0(\cdot))$, there exists (at most) a unique econometric model ζ with $g = g_0$ and $\gamma = \gamma_0$. As we will next, identification requires restrictions on the econometric model, often in the form of some partial knowledge of the true earnings distribution.

First, assume that the distribution of true earnings $f(\cdot)$ has an unconditional mean m known to the econometrician and $\gamma(\cdot)$ is increasing. For example, we may have a valuation model that tells us that the mean of true earnings should be equal to a certain multiple of price and its variance connected, from this valuation model, to the variance of cash flows or price. Or, we may assume that, in steady-state, expected earnings are set by the accounting system equal to expected cash flows. Then, the reporting bias is decreasing in the manipulation cost, implying that the difference between the mean of true earnings and reported earnings

$$\Delta(\theta) = \int x f(x) dx - \int r \hat{g}(r) dr \quad (2.10)$$

is increasing in θ . Given the mean report m_r (which can be consistently estimated as the sample mean), there is a unique solution to $\Delta(\theta) = m - m_r$ and the identification of θ follows.

Second, assume that we have no a-priori restriction on true earnings, formally, let the econometric model includes any p.d.f. $f(\cdot)$. For some value θ' that satisfies the second-order condition, we can define

$$f(x; \theta') = g(R(x; \theta')) \frac{\psi''(R(x; \theta') - x)}{\psi''(R(x; \theta') - x) - \theta' \gamma''(R(x; \theta'))}, \quad (2.11)$$

where $R(x; \theta')$ is defined as the solution to $\gamma'(y) = \frac{1}{\theta'} \psi'(y - x)$. Observe next that multiple choices of $(\theta, f(\cdot))$ imply the same observables in the econometric model. We

may set, for example, $\theta' = 0$ and $f(x) = g(x)$, i.e., an infinite manipulation cost and a true distribution that conforms exactly to the observed distribution.⁵ Or, we may set θ' not too large, so that it satisfies the second-order condition, and $f(x) = f(x; \theta')$ from (3.1). There is, however, a possibility to set identify the cost θ for observed distribution of earnings whose support is an interval, using the second-order condition in (2.4). That is, if $g(\cdot)$ has support over a convex set, we know that

$$\theta \leq \frac{\psi''(R(x) - x)}{\gamma''(R(x))} \leq \min_r \frac{\psi''((\psi')^{-1}(\theta\gamma'(r)))}{\gamma''(r)}. \quad (2.12)$$

Third, assume that the observed price function is a polynomial with degree one (linear) or two (quadratic), that is, $\gamma''(\cdot) = v_0$ is constant, implying a p.d.f. of reported earnings given by

$$g(r) = \frac{\psi''(r - R^{-1}(r)) - \theta v_0}{\psi''(r - R^{-1}(r))} f(R^{-1}(r)). \quad (2.13)$$

Assume, further, that (a) $\psi(x) = x^2$ is quadratic, (b) $f(\cdot)$ is unknown up to a location and a scale parameter, that is,

$$f(x) = \frac{1}{\beta} k\left(\frac{x - \alpha}{\beta}\right), \quad (2.14)$$

where $(\alpha, \beta) \in X_\alpha \times X_\beta$ are potentially unknown to the econometrician. This nests families of distributions: for example, in the case of the normal distribution $k(\cdot)$ is the standard-normal, α is the mean and β is the standard error.⁶ The next lemma shows that scale and location cannot be jointly identified in this setting, as noted below.

Lemma 2.1 *The following holds:*

1. *if $\gamma(\cdot)$ is linear and $X_\alpha = \mathbb{R}$, then (θ, α, β) is not identified;*
2. *if $\gamma(x)$ is quadratic with $X_\beta = \mathbb{R}^+$, then only θ/β and α/β are identified.*

Linear pricing functions are widely-used in theoretical studies, as in Dye (1988), Stein (1989), Fischer and Verrecchia (2000) and Dye and Sridhar (2004), but present some identification challenges. To obtain an intuition for this, note that $\gamma''(\cdot) = 0$ implies that

⁵This is a limitation of the approach but it is also true in many commonly-used approaches: discrete choice models for example, are only parametrically identified given a-priori assumptions about the distribution of preference shocks and, even under parametric assumptions, may fail certain primitives such as discount rates (Rust 1987; Magnac and Thesmar 2002).

⁶Many other classes of distributions can be represented with a scale and location parameter, possibly taking the set X_α or X_β as a singleton (for exponential distributions, for example, we would take $X_\alpha = \{1\}$ and represent the parameter of the exponential as different choices of β).

$$g(r) = f(R^{-1}(r)) = f\left(r - \frac{\theta v_1}{2}\right). \quad (2.15)$$

Hence, the predicted distribution of reported earnings only depends on θ via the location parameter α . The intuition here is that misreporting translates the distribution in the same manner as would a location parameter. We may reconcile this identification result with prior literature that identifies manipulation in this setting, by noting that such studies make additional assumptions about the valuation implication of earnings. In Beyer, Guttman and Marinovic (2012), earnings are interpreted as cash flows which, once discounted, have implications about true relationship between price and reported earnings. Any deviation between reported earnings and true earnings helps pin down the manipulation cost even if the price response is linear. This can help sharpen identification, but involves a trade-off using stronger valuation assumptions - our approach, on the other hand, assumes less in terms of how earnings are “converted” into a price. In addition, the empirical relation between price and earnings is not linear, so that the linear case may be more pathological in nature.⁷

The case of quadratic pricing is slightly more subtle. With a quadratic response, misreporting costs shifts the distribution but also transforming it by changing its dispersion. The greater the convexity, the more the distribution of reported earnings becomes dispersed relative to true earnings - the resulting distribution of outcomes remains, however, within the same class of distributions with a different location and scale parameter, so we cannot separate greater misreporting cost from different location and scale in true earnings. Put differently, the model is identified per unit of scale in the original distribution of true earnings.

A different identification procedure is to consider $f(\cdot)$ to be (possibly locally) symmetric, i.e., symmetric around certain known quantiles. This approach assumes less than using classes of distributions that are symmetric (say, the normal distribution).

Lemma 2.2 *Suppose that $\psi(x) = x^2$ and (ii) $f(\cdot)$ is locally symmetric around a real number m . Then, if $\gamma(\cdot)$ is a polynomial with degree two or lower, or symmetric around m , θ is not identified. Otherwise, θ is identified.*

⁷Two additional aspects of Beyer, Guttman and Marinovic (2012) are not (yet) apparent in our baseline model, but can be incorporated without much conceptual difficulty. First, they assume noise in the manager’s objective function which, in a later section, we show can be incorporated into the model with a simple integration of the likelihood function (we actually estimate a random cost model with real earnings data). Second, their model involves serial correlation in true earnings, which can be incorporated by conditioning the estimation on various possible histories. A third aspect of their model does not easily generalize in our model: their maximization problem is on current and future prices while we assume that the manager maximizes current price. Unfortunately, extending the approach to forward-looking problems is non-trivial and no longer in closed-form: absent linearity, would involve computing a misreporting manager’s value function.

One benefit of lemma 2.2 is that it reveals that, given a large enough data set, parametric assumptions such as normality can be weakened, to the extent that some knowledge of symmetry around a certain point is sufficient. Further, symmetry need only be local and need not hold over the entire distribution if, say, extreme events may be tied to extreme risks or, on the other hand, capture a liquidation option.

2.4. Random manipulation cost

Estimation of the baseline model can be extended to a random θ that follows a distribution $h(\cdot)$, where this distribution may be a function of parameters to be estimated. Unfortunately, we are not able to provide sharp identification results in this case.

Let $R(x, \theta) \in \mathbb{R}$ denote the firm's reporting strategy, and $\gamma : \mathbb{R} \rightarrow \mathbb{R}$ denote a mapping that associates a price $\gamma(r)$ to each report r . For any r and conjectured reporting strategy $\bar{R}(x, \theta)$, the price is assumed to satisfy, whenever possible,

$$\bar{\gamma}(r) = \mathbb{E}(\alpha(\tilde{x}) | \bar{R}(\tilde{x}, \theta) = r). \quad (2.16)$$

For any (x, θ) , the manager chooses a report to maximize a utility function

$$R(x, \theta) \in \operatorname{argmax}_r \quad \bar{\gamma}(r) - \frac{\psi(r - x)}{\theta}. \quad (2.17)$$

As before, an equilibrium is such that $\gamma(r) \equiv \bar{\gamma}(r) = \mathbb{E}(\alpha(\tilde{x}) | R(\tilde{x}, \theta) = r)$. Differentiating (2.17) and evaluating at $r = R(x, \theta)$, the first-order optimality condition implies

$$\gamma'(R(x, \theta)) - \frac{\psi'(R(x, \theta) - x)}{\theta} = 0. \quad (2.18)$$

We restrict the analysis to equilibria in which (2.18) has a unique solution $R(x, \theta)$ for any given (x, θ) and this solution satisfies $\partial R / \partial x > 0$ and $\partial R / \partial \theta > 0$.⁸

Data about prices and earnings identifies the pricing function $\gamma(\cdot)$ and the distribution of reported earnings $R(x, \theta)$, whose p.d.f. is denoted $g(\cdot)$, and follows immediately by integration over γ :

$$g(r) = \int \frac{\psi''(r - R^{-1}(r, \theta)) - \theta \hat{\gamma}''(r)}{\psi''(r - R^{-1}(r, \theta))} f(R^{-1}(r, \theta)) h(\theta) d\theta. \quad (2.19)$$

⁸Since a corner solution $r \in \{-\infty, \infty\}$ is suboptimal, this assumption also implies that the first-order condition is also sufficient and a second-order condition $\gamma''(R(x, \theta)) - \frac{1}{\theta} \psi''(R(x, \theta) - x) \leq 0$ must hold as well.

3. Estimation

3.1. Sample Selection

To estimate the structural model, we use the price data from CRSP, accounting data from Compustat and forecast data from IBES (see Table 1). Our sample starts with 249,293 firm-years from the intersection of Compustat Annual and CRSP for the period 1976-2014. To calculate annual earnings scaled by market value of equity, we require the following data to be non-missing: income before extraordinary items (item *IB*) and beginning-of-the-year market value of equity (shares outstanding (*CSHO*) * stock price (*PRCC_F*)). This eliminates 70,147 observations. Consistent with prior research, we drop 24,333 observations from financial (SIC codes 4400-5000) and regulated (SIC codes 6000-6500) industries. This is done primarily to be comparable to samples in other studies although, to be fair, our theory does not provide justification for this sample selection.

Because we need to calculate market reactions around earnings announcement dates, we further eliminate 30,947 observations that have missing earnings announcement dates (*RDQ*) in any of the four quarters. Finally, we remove 57,275 observations that have missing data from IBES to calculate quarterly earnings surprise. Our final sample comprises 65,023 observations. Similar to Gilliam, Heflin and Paterson (2015), we also split our sample period into a pre-SOX period (1976-2001) and a post-SOX period (2003-2014). We trim the top and bottom 1% of the data in the actual estimation.

Table 1: Sample Selection

Panel A: Initial Sample Selection	
Firm-years at the intersection of Compustat and CRSP, 1976-2014	249,293
Less: Missing data to calculate annual earnings scaled by market value of equity	(70,147)
Firms in financial and regulated industries	(24,333)
Missing Earnings announcement dates (RDQ)	(30,947)
Missing IBES data to calculate earnings surprise	(57,275)
Initial sample	65,023
Panel B: Pre-SOX Sample Selection	
Firm-years at the intersection of Compustat and CRSP, 1976-2001	32,215
Less: Trimmed top and bottom 1%	(644)
Pre-SOX Sample	31,571
Panel C: Post-SOX Sample Selection	
Initial Sample, 2003-2014	30,111
Less: Trimmed top and bottom 1%	(407)
Post-SOX Sample	29,704

To compare our sample to Gilliam, Hefflin and Paterson (2015), we calculate *ROE*, the net income scaled by beginning-of-the-year market value of equity. There is no perfect guidance for the appropriate scaling and it is worth noting that (like any study using this form of scaling) this scaling choice supposes that properties of accounting return to an investor buying equity is similar across time periods and firms.

Table 2: Descriptive Statistics

variable	N	Mean (%)	sd	Pctl 1 (%)	Pctl 5 (%)	Lower Quartile (%)	Median (%)	Upper Quartile (%)	Pctl 95 (%)	Pctl 99 (%)
ROE	65023	0.004	0.152	-0.799	-0.285	-0.010	0.044	0.074	0.142	0.263
ROA	65023	0.010	0.184	-0.866	-0.368	-0.010	0.048	0.097	0.206	0.356
ESurp_4q	65023	-0.007	0.046	-0.286	-0.072	-0.007	0.001	0.006	0.033	0.108
ESurp_ann	64924	-0.010	0.058	-0.440	-0.059	-0.003	0.000	0.002	0.018	0.092
R_year	65023	0.140	0.621	-0.861	-0.658	-0.239	0.055	0.367	1.255	2.969
CAR_3day_4q	65023	0.005	0.177	-0.456	-0.284	-0.094	-0.001	0.095	0.313	0.581
CAR_3day_ann	65023	0.002	0.083	-0.249	-0.136	-0.039	0.000	0.042	0.141	0.268
CAR_sa_3d_4q	65023	0.006	0.176	-0.454	-0.281	-0.092	0.000	0.096	0.313	0.578
MVE	65023	2838.9	8184.2	16.5	38.6	147.6	466.7	1633.3	12890.8	60464.3
BM	65008	0.520	0.400	-0.17	0.078	0.255	0.429	0.679	1.284	2.236

ROE is the net income (item NI) scaled by beginning-of-the-year market value of equity (shares outstanding (CSHO) * stock price (PRCC.F)). ROA is the net income (item NI) scaled by beginning-of-the-year total assets (AT). $ESurp_{4q}$ is the sum of four quarterly earnings surprises in the same calendar year. $ESurp_{ann}$ is the annual earnings surprise. R_{year} is the buy-and-hold return in the twelve month window ended three months after the fiscal year end. CAR_{3day_4q} is the difference between the buy-and-hold return of a firm and that of CRSP value-weighted market portfolio over the four 3-day windows centered on quarterly earnings announcement dates. $CAR_{3day_{ann}}$ is the difference between the buy-and-hold return of a firm and that of CRSP value-weighted market portfolio over the three-day window from Day -1 to Day +1 of earnings announcement date. $CAR_{sa_3d_4q}$ is the difference between the sum of four quarterly buy-and-hold returns in the window from Day -1 to Day +1 of earnings announcement date and that of the corresponding size-decile portfolio (NYSE/AMEX/NASDAQ).

Our main empirical construct for earnings in the structural model is $ESurp_{4q}$, which is the sum of earnings surprises over all four quarters. For each quarter, the earnings surprise is the difference between IBES actual earnings and the last analyst consensus (mean) from the IBES summary file, deflated by the stock price two trading days before earnings announcements.

As for the empirical construct for the price response, we use the sum of 3-day excess returns around all four quarterly earnings announcements ($CAR_{3day_{4q}}$) to measure the information conveyed by reported earnings over the year. Excess return is calculated as the difference between the buy-and-hold return of a firm and that of CRSP value-weighted market portfolio over the four 3-day windows centered on quarterly earnings announcement dates (i.e., 12 trading days in total). Note that we use a short-term market response to earnings announcements to get as close as we can to a causal effect of earnings surprises on returns.⁹

We develop first estimates of the model assuming that the sample is homogenous on scaled variables. This is not meant for realism but, in the same way that it is cumbersome to conduct separate regressions for each possible portfolio of characteristics in reduced-form analyses, it offers a simple summary analysis of the data. We shall later on provide more detailed estimates where observable heterogeneity is controlled by estimating by subsamples of industry, size and growth. Other possibilities, which we did not implement, would involve controlling for observed heterogeneity by making the cost and/or the market response a function of various observable characteristics.

We plot in Figure 1 the market response, i.e., the estimated $\hat{\gamma}(\cdot)$, to earnings surprises in the full sample, pre-SOX and post-SOX. Throughout, we fit γ using a higher-order polynomials, specifically, we regress price on earnings on a polynomial function of earnings.¹⁰

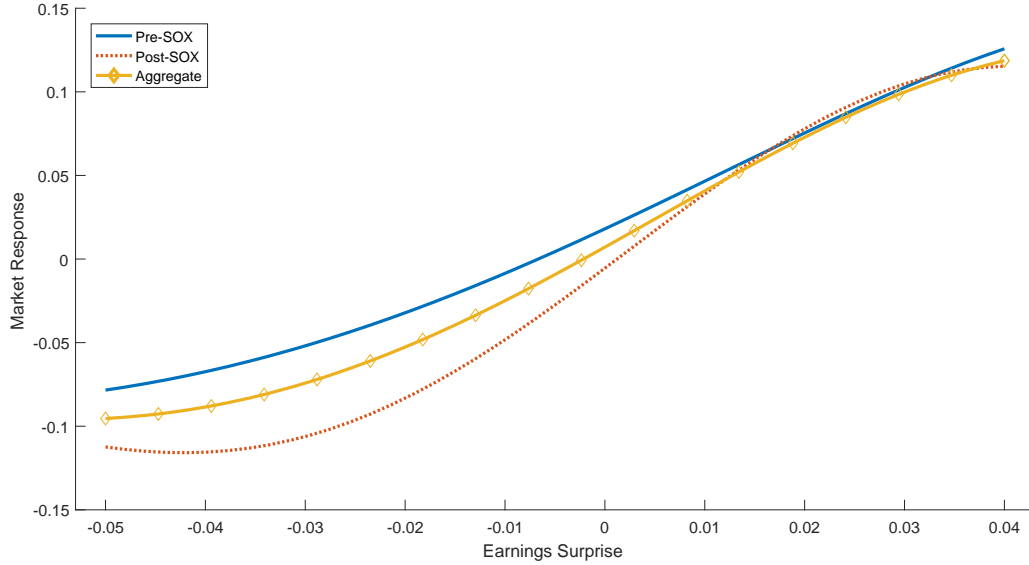
The plot of the market responses is steeply sloped for small absolute earnings surprises and approximately flat for large negative earning surprises for both periods. In particular, the market response to earnings surprises pre SOX is stronger than post SOX for a small interval around the zero $ESurp_{4q}$ threshold, providing higher incentives for firms to manage earnings pre SOX. The post SOX reaction is slightly stronger than pre SOX for

⁹An alternative approach would be a value-relevance design that measures a long term (contemporaneous) association between annual earnings and return. However, this method may overestimate the price response to earnings if earnings respond to market news and it does not allow us to use surprises.

¹⁰We also used a Nadarya-Watson kernel regression. However, the kernel method presents some non-trivial challenges. First, with data sets of up to 100k observations and the need to estimate first and second derivatives, it is very slow; with polynomials, by contrast, we use the analytical derivatives. Second, there is no clear guidance for the optimal bandwidth to estimate first and second derivative, and prior research suggests that differentiating the estimated $\hat{\gamma}(\cdot)$ can be very noisy (Newell and Einbeck 2007). Our more applied choice toward a simple polynomial is primarily driven from the fact that is naturally extends a degree one regression as used in the literature.

large positive $ESurp_{4q}$, which could encourage more upward manipulation.

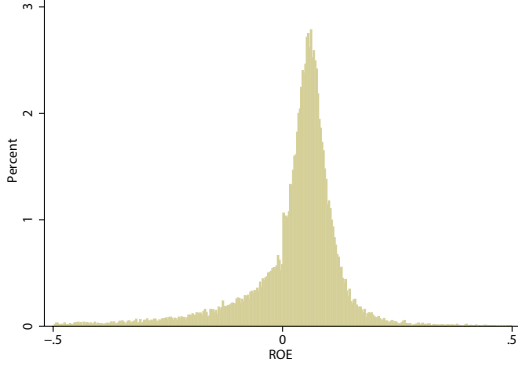
Figure 1: Market Response Function



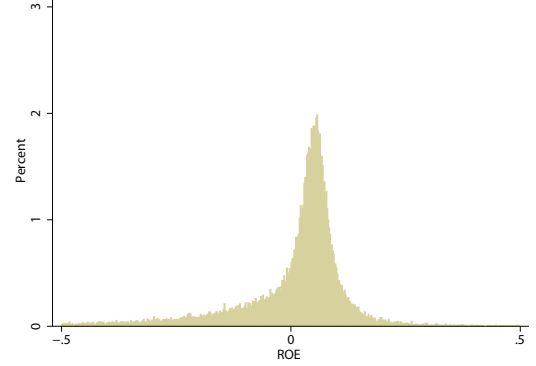
We also plot the distribution of earnings surprises ($ESurp_{4q}$) and reported earnings (ROE) in Figure 2, which illustrates some of the debates about how to interpret the earnings distribution without theory. There is a discontinuity both pre and post SOX in surprise space, but the discontinuity no longer exists in ROE in the post period (Gilliam, Heflin and Paterson 2015). There is also debate in the literature about whether the discontinuity is sufficient to conclude that manipulation exist and/or if it is economically significant. An unexpected output of our analysis is to offer a contribution to this debate.

Figure 2: Cross-Sectional Earnings Distributions for Pre-SOX and Post-SOX Eras

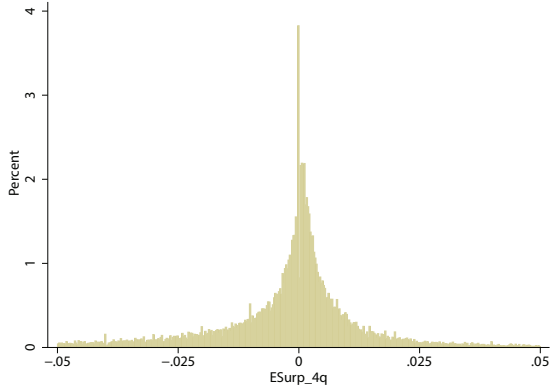
Panel A: Pre-SOX ROE



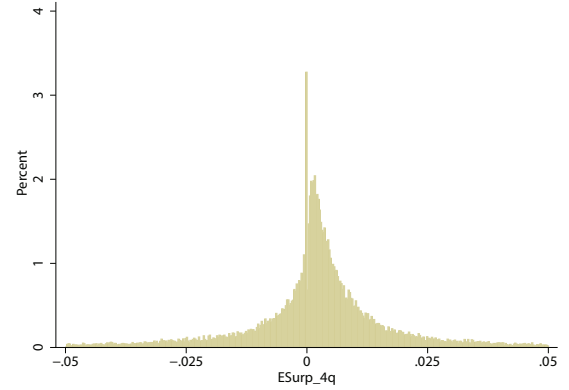
Panel B: Post-SOX ROE



Panel C: Pre-SOX Earnings Surprise



Panel D: Post-SOX Earnings Surprise



These figures show the cross-sectional distributions of earnings. Panel A and Panel B are the distributions ROE, which are calculated using annual earnings scaled by beginning-of-the-year market value of equity. Panel C and Panel D are distribution of earnings surprises, which are the sum of the four quarterly earnings surprise in the same calendar year. They are based on 31,609 and 29,451 firm-years in the pre-SOX era and the post-SOX era, respectively. Interval widths are 0.003 with 400 bins from -0.5 to 0.5.

3.2. Estimation with known manipulation cost

We first focus on a baseline model where cost of manipulation is constant across firms. Assume the cost function $\psi(\cdot)$ takes a quadratic form: $\psi(r - x) = (r - x)^2$. Thus, the density function of the accounting reports is

$$\hat{g}(r) = (1 - \gamma''(r)\frac{\theta}{2})f(r - \frac{\theta}{2}\gamma'(r)). \quad (3.1)$$

The likelihood of the sample is

$$\mathcal{L}(\theta, \mu_x, \sigma_x) = \sum_{i=1}^N \log((1 - \gamma''(r_i)\frac{\theta}{2})f(r_i - \frac{\theta}{2}\gamma'(r_i))). \quad (3.2)$$

The second-order condition of the manager's problem requires a lower bound of the cost of manipulation:

$$\theta \leq \min \frac{2}{\gamma''(r_i)}, \quad (3.3)$$

which we impose as a bound on the estimation.

For our baseline model, we select $f(\cdot)$ to be normally distributed with mean μ and σ^2 , which are two parameters to estimate. Of course, it would be possible to use a different class of distributions as input which is not necessarily symmetric. We have not done so at this stage because we do not yet have theoretical guidance about classes of distributions generated by a decision problem. To be more precise, the normal or lognormal distribution, which is implied in various other studies (e.g., Beyer, Guttman and Marinovic 2012 and Zakolyukina 2014), presumes that the setting is close enough to one of pure-exchange where exogenous shocks to an information environment get released to the market and the main focus is about the communication of that information.¹¹ We take it primarily as a first-step on this problem and the simplest possible illustration of the structural model without adding more complexity to the model.

Table 3 reports the results of maximum likelihood estimation using the baseline model, as well as the standard errors. We estimate the model using pre-SOX sample, post-SOX sample as well as the entire sample. The mean and standard deviation of the true earnings are stable across samples. The means of the true earnings signal are all slightly negative: -0.04 for pre-SOX period (1976 - 2001) and -0.01 for post-SOX period (2003 - 2014).¹² The standard deviations of managers' private information are 0.005 for the pre-SOX period and 0.012 for the post-SOX period and the whole sample, showing a decrease in the information environment after the passage of SOX. Note that similarly low accounting returns are also found under different approaches for estimating the level of manipulation (Beyer, Guttman and Marinovic 2012, Zakolyukina 2014,) and, while not found under purely regression-based (abnormal-accrual) approaches, this occurs only because these approaches do not capture manipulation levels. Perhaps, one rationalization for low accounting returns is that conservatism in what accounting numbers mean would imply a negative or zero accounting returns for firms that have non-zero growth (Rajan, Reichelstein and Soliman 2007), as is the case for most of the firms in our sample.

The cost of manipulation is significantly larger by a sixfold increase in the difficulty to manipulate earnings after the passage of SOX. Comparing the mean of true earnings (as

¹¹In studies that fall outside the pure manipulation literature, Bertomeu, Ma and Marinovic (2015) and Zhou (2016) also assume the normal distribution in their structural estimation.

¹²One potential explanation for negative true earnings is that these are not value-weighted, so that firms with sequences of low earnings will have lower value of equity and, hence, a return on asset that is extremely negative.

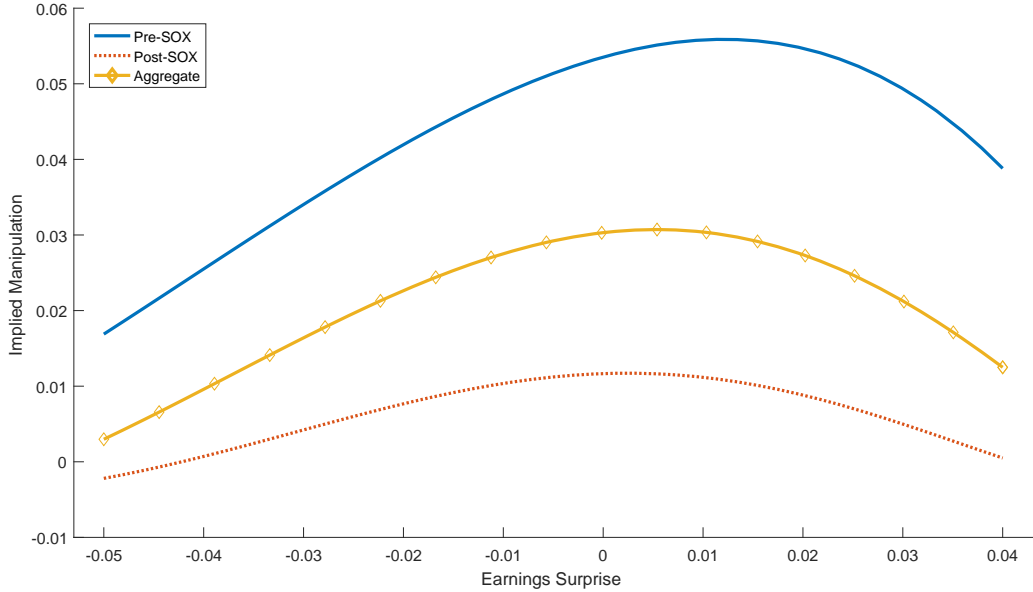
Table 3: Estimation Results

Maximum likelihood estimation				
	1/cost	Mean _x	Standard deviation _x	Likelihood
	θ	μ_x	σ_x	
Pre-SOX Period (1976-2001)	0.039 (0.012)	-0.054 (0.015)	0.007 (0.003)	2.985
Post-SOX Period (2003-2014)	0.005 (0.001)	-0.010 (0.002)	0.012 (0.000)	2.978
Aggregate Period (1976-2014)	0.018 (0.007)	-0.030 (0.008)	0.010 (0.001)	2.976

Standard errors are in parentheses, which is calculated by using firm block bootstrap.

estimated) with the mean of reported earnings (as obtained from the data), the average bias decreases by three times, from 0.055 pre-SOX to 0.007 post-SOX. This significant decrease in average bias, together with the increase in manipulation cost and a reduction in the steepness of the market response for earnings surprises around zero.

Note that this is an average over the entire sample because the bias is different for each reported earnings. We examine the manipulation at any level of reported earnings in Figure 3. SOX seems to have changed managers' strategic behavior when it comes to manipulating earnings. Post-SOX manipulation is consistently lower than pre-SOX manipulation. In particular, at zero earnings surprise, the post-SOX manipulation of 0.012 has decreased by 0.042 relatively to the pre-SOX manipulation 0.054. The highest estimated amount of manipulation for post-SOX is when earnings surprise is 0.002, almost coinciding with the zero earnings threshold, pre-SOX the highest manipulation is slightly above the zero earnings threshold, at 0.012, with an expected bias of 0.056.

Figure 3: Implied Manipulation using MLE

Implied Manipulation is obtained by $\hat{R}(x) - x = \frac{1}{2c}\hat{\gamma}'(\hat{R}(x))$.

3.3. Counterfactuals

Our structural approach allows us to measure some counterfactuals. Namely, the reporting bias might have been reduced due to a reduction in price response and/or an increase in cost. Only the latter may be directly traced to improvements in internal controls in SOX, while the former may reflect changes in the price sensitivity to the underlying information being reported.

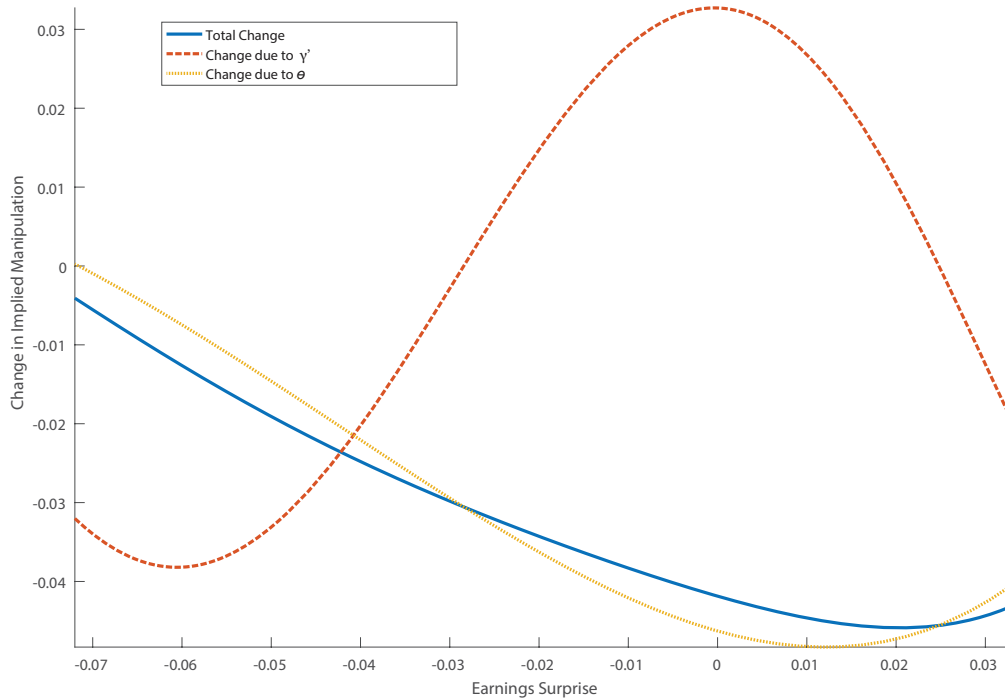
To quantify the change in manipulation caused by the change in cost and $\gamma'(\cdot)$ respectively, we examine the change in manipulation using the pre-SOX $\gamma'(\cdot)$ and keeping the estimated post-SOX manipulation cost, and then, the change in manipulation using the estimated pre-SOX manipulation cost and keeping the post-SOX $\gamma'(\cdot)$. The post-SOX decrease in manipulation of 0.048 in the entire sample can be explained by the increase in cost for 0.042 and an overall decrease in $\gamma'(\cdot)$ for 0.013.¹³ So both forces appear to have been about as important in explaining the reduction in the bias. Naturally, the interpretation of this decomposition comes with a caveat, to the extent that the change $\gamma'(\cdot)$ is at least partly driven by a change in the cost.

We plot next in Figure 4 the post-SOX change in manipulation for every level of

¹³Note that the sum of the two reductions due to the changes in manipulation cost and γ' do not add up to the total change in manipulation because the change due to the interaction between the manipulation cost and γ' is not taken into account.

reported earnings. For the largest earnings surprise, both the reduction in $\gamma'(\cdot)$ and the increase in the manipulation cost contribute to the decrease in manipulation. For earnings surprises close to zero, the increase in manipulation cost gets larger but the change in $\gamma'(\cdot)$ is positive and becomes larger as we approach the zero earnings surprise. The increase in cost consistently reduces the manipulation and outweighs the increase in $\gamma'(\cdot)$. As earnings surprises become more positive, the positive change in $\gamma'(\cdot)$ starts to steeply decrease thereby lowering manipulation that attains the lowest reduction at 0.025 of earnings surprise. Moreover when earnings surprises are slightly positive, at about 0.012, the cost increase achieves its highest level and contributes to the decrease in manipulation by -0.048 . For larger earnings surprises above 0.028, the change in cost is less negative but the larger negative change in $\gamma'(\cdot)$ helps maintain a large reduction in manipulation. To summarize, changes in cost tend to matter more for intermediate levels of earnings.

Figure 4: Breakdown Post-SOX Change in Manipulation



The solid curve is the difference of implied manipulation between the post-SOX and pre-SOX period. The dashed curve is the change in manipulation caused by the change of γ' after SOX, i.e., the manipulation estimated by keeping the estimated pre-SOX manipulation cost and using post-SOX γ' , and subtract the pre-SOX manipulation. The dotted curve is the change in manipulation caused by the change in cost after SOX, which is obtain by taking the difference between the manipulation calculated by using the pre-SOX γ' and use the estimated post SOX cost and subtract the pre-SOX manipulation.

4. Extensions

4.1. Relaxing distributional assumptions

Symmetric distributions

Consistent with our identification results, we weaken next the assumption of normal distributions, assuming that $f(\cdot; \beta)$ is symmetric, where β is a J -dimensional set of parameters such that each β implies a different set of moments $(m_1(\beta), \dots, m_J(\beta))$. Note that m_i may refer to any moment, not necessarily the i^{th} moment.

For any parameter θ , let the inferred true earnings be denoted

$$\hat{x}_i(\theta) = r_i - \frac{\theta}{2} \hat{\gamma}'(r_i), \quad (4.1)$$

and denote $\mu_i(\theta)$ as the same moment as m_i calculated with $\hat{x}_i(\theta)$. We can then minimize the following objective function to recover θ :

$$Obj = \sum_{j=1}^J (m_j(\beta) - \hat{\mu}_j(\theta))^2 + (\tau_0(\theta) - \tau_1(\theta))^2, \quad (4.2)$$

where $\tau_0(\theta)$ (resp., $\tau_1(\theta)$) is the mean (resp., median) of the $(\hat{x}_i(\theta))$. Note that this estimation procedure can be applied more generally on a distribution of x that is only symmetric over a range, by suitably trimming the data.

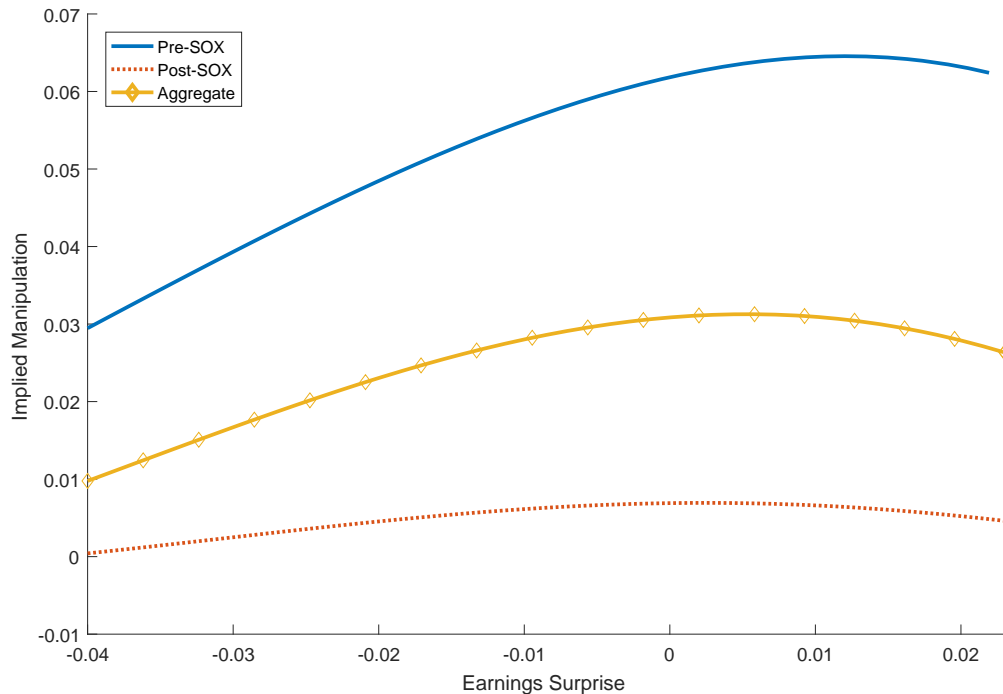
Table 4 reports the results using the method of moments. In the pre-SOX sample, θ is 0.045, but not statistically significant. In the post-SOX sample, the manipulation cost increases significantly and is about 15 times more than the cost in the pre-SOX period. Hence the average manipulation goes down post SOX, roughly divided by 10. The estimated mean of the underlying earnings m_1 is higher post SOX than pre SOX. Note that, as one would expect, using non-parametric identification puts a greater burden on data and increases uncertainty in the estimation. We may no longer conclude with high certainty, in this design, that reporting bias was non-zero even in the pre-SOX period.

Table 4: Method of Moments Estimation

	$1/\text{cost}$ θ	$Mean_x$ m_1	$Avg\ Bias$ $Mean(r) - m_1$	$Std\ Dev_x$ $\sqrt{m_2}$	Obj.
Pre-SOX Period (1976-2001)	0.045 (0.035)	-0.062 (0.028)	0.057	0.005 (0.002)	1.09E-16
Post-SOX Period (2003-2014)	0.003 (0.001)	-0.005 (0.002)	0.006	0.012 (0.000)	2.24E-20
Aggregate Period (1976-2014)	0.019 (0.005)	-0.03 (0.006)	0.028	0.01 (0.001)	5.08E-17

Standard errors are in parentheses, which is calculated by using firm block bootstrap.

Figure 5 plots the implied point estimates manipulation using method of moments. The highest level of manipulation pre-SOX of 0.065 is attained at an earnings surprise 0.012, whereas the highest level of manipulation post-SOX of 0.007 is much lower and attained at a lower earnings surprise of 0.0024.

Figure 5: Implied Manipulation Using the Method of Moments

Exogenous shock to manipulation costs

We relax next the requirement of symmetric distributions. Instead, we assume that a shock may have changed both θ and a subset of parameters in β . However, the identification restriction that we impose is that at least two moments of the distribution of true earnings remained unchanged. Note that the symmetry assumption made earlier implies that this is the case but, vice-versa, non-symmetric distributions may be identified as well.

Nevertheless, the estimation procedure borrows similar steps to the procedure with the symmetric distribution x : let the variables pre and post shocks be indexed by $k = 1, 2$, implying that inferred true earnings in each sample are

$$\hat{x}_i^k(\theta) = r_i^k - \frac{\theta^k}{2} \hat{\gamma}'(r_i^k). \quad (4.3)$$

Assuming that two moments $j = 1, 2$ are unchanged, we can minimize an objective function given by

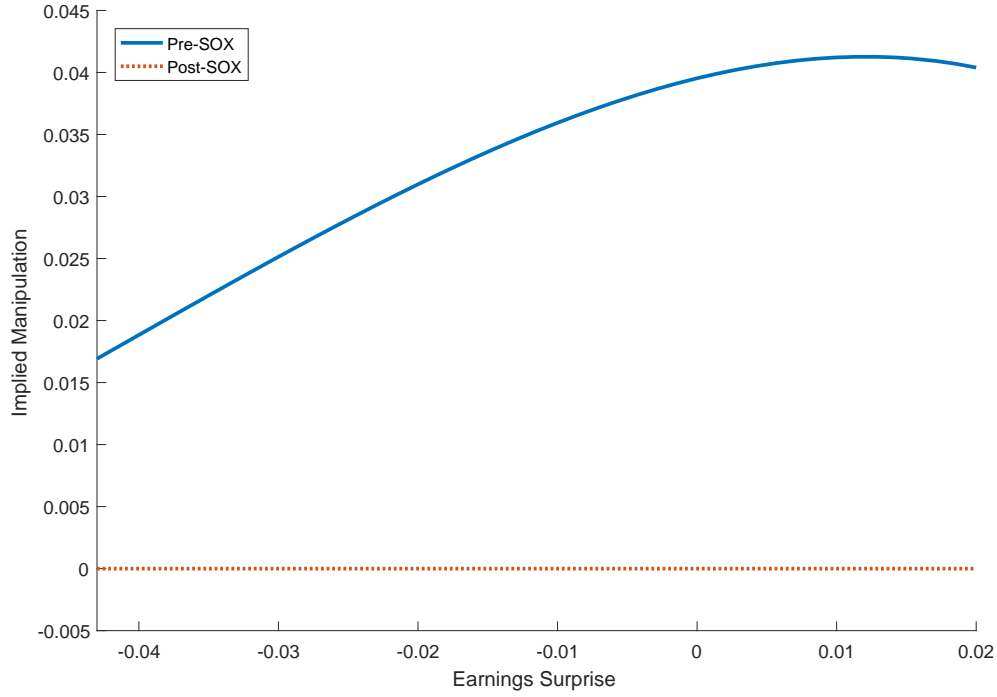
$$Obj = \sum_{j=1}^2 (\hat{\mu}_j^1(\theta^1) - \hat{\mu}_j^2(\theta)(\theta^2))^2. \quad (4.4)$$

For the estimation, we use the pre and post SOX samples, viewing SOX as a shock, and use as an identification restriction an assumption that the variance and skewness of earnings did not change (but mean earnings might have changed).

Table 5: Estimation Using Non-parametric Estimation

	Pre-SOX	Post-SOX	Obj
θ	0.028	0	1.48E-05
Std.Err	(0.010)	(0.000)	

In Table 5, we find that under this restriction, manipulation was non-zero in the pre-SOX period but then was zero in the post-SOX period. Jointly with the prior symmetric results, it follows that the non-parametric tend to feature lower manipulation in general (which we would expect from more generality in fitting the earnings distribution) and an effect of SOX that is more consistent on eliminating all manipulation as illustrated in Figure 6.

Figure 6: Implied Manipulation Using Non-parametric Estimation

4.2. Subsamples

So far, we have estimated the manipulation cost without plausibly allowing for heterogeneity across firms and/or industries, which has the benefit of offering one summary set of estimates. However, the estimation may be made to capture the heterogeneity more precisely by dividing into subsamples, which is admittedly the simplest manner to incorporate heterogeneity into a non linear model (see, e.g., Gayle, Li and Miller 2015; Li 2016). We chose industry (Fama French 12), size and book-to-market since these are characteristics that are likely to affect cost and incentives and have been documented to be associated to existing proxies (Dechow, Ge and Schrand 2010).

Industries

We divide our sample using the Fama-French 12 industry classification and since finance and regulation industries are excluded from our main sample, we count nine industries in our sample. The estimated θ and average bias per industries are reported in Table 6.

Table 6: Industry Estimation Results

Fama-French 12 Industries	θ			Average bias		
	Pre-SOX	Post-SOX	Aggregate	Pre-SOX	Post-SOX	Aggregate
Consumer NonDurables	0.029 (0.012)	0.003 (0.008)	0.015 (0.008)	0.033	0.029	0.031
Consumer Durables	0.035 (0.011)	0.01 (0.010)	0.028 (0.012)	0.040	0.035	0.038
Manufacturing	0.002 (0.022)	0.004 (0.002)	0.023 (0.007)	0.006	-0.000	0.003
Oil, Gas, and Coal	0.002 (0.153)	0.089 (0.099)	0.033 (0.192)	0.006	0.003	0.004
Chemicals and Allied Products	0.037 (0.045)	0.011 (0.012)	0.026 (0.011)	0.039	0.035	0.037
Business Equipment	0.023 (0.013)	0.004 (0.003)	0.032 (0.011)	0.030	0.019	0.024
Wholesale, Retail, and Some Services	0.004 (0.008)	0.003 (0.002)	0.008 (0.002)	0.009	0.003	0.006
Healthcare, Medical Equipment, and Drug	0.140 (0.056)	0.005 (0.018)	0.007 (0.041)	0.143	0.138	0.140
Other	0.036 (0.019)	0.011 (0.020)	0.035 (0.016)	0.042	0.038	0.039

Figure 7 plots the industry θ and the 90% confidence intervals in the pre- and post-SOX period. Most of the industries experience a higher manipulation post SOX except for the *Manufacturing* and the *oil, gas and coal industries*, which returns anomalous high standard errors. *Healthcare, Medical Equipment and Drug* industry experiences the most significant increase in manipulation.

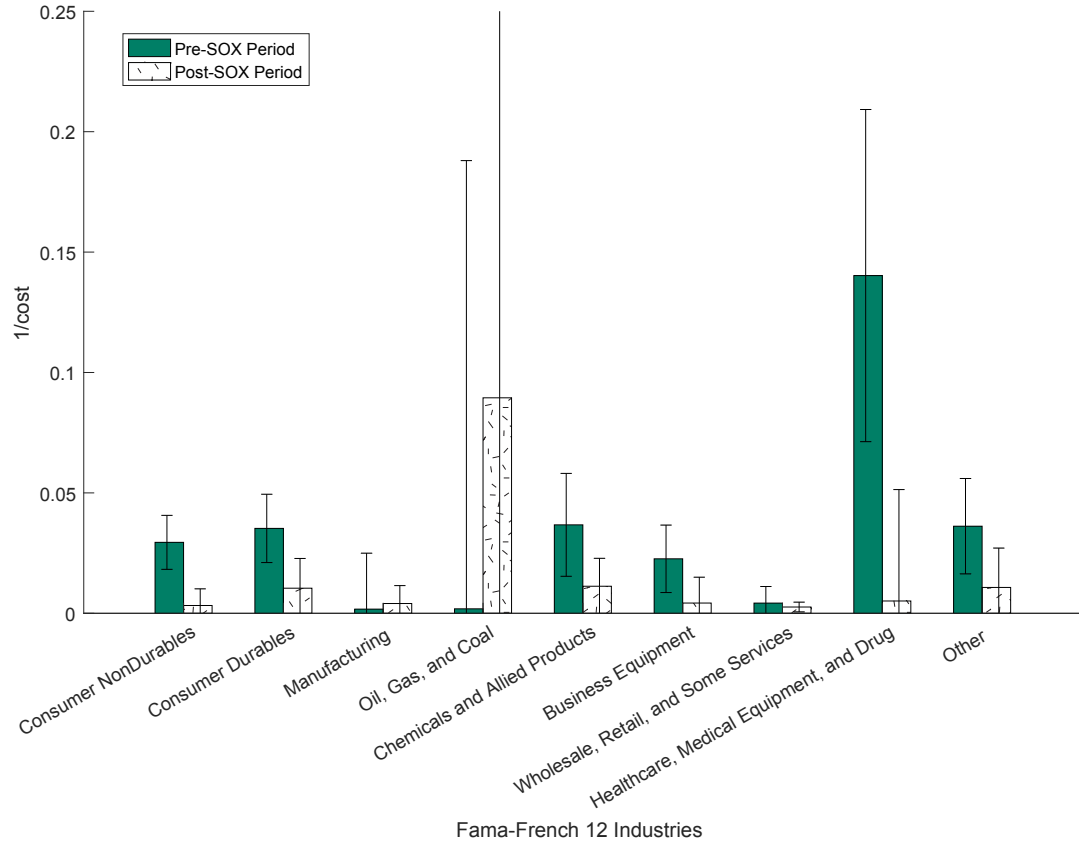
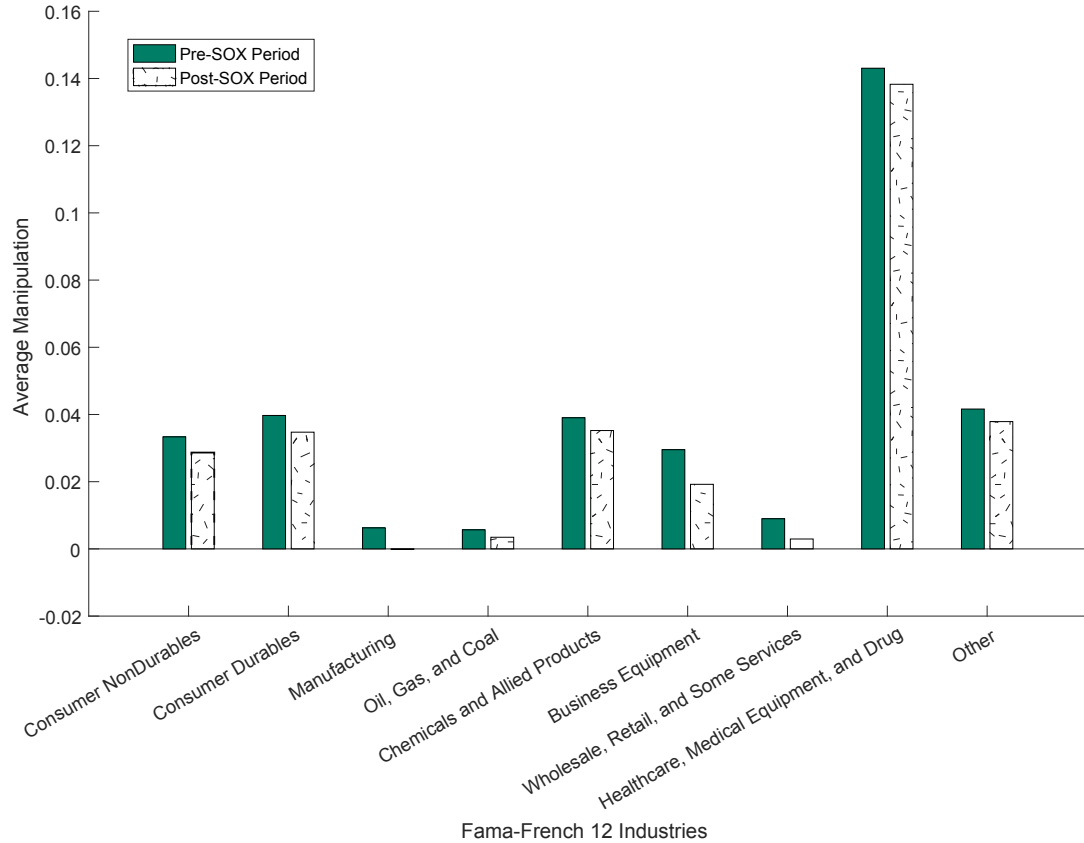
Figure 7: Estimated θ by Industry

Figure 8 plots the average manipulation pre and post SOX periods by industry, which is the difference between the mean of $ESurp_{4q}$ and the estimated mean of underlying earnings. All industries exhibit a lower manipulation post SOX, though small in magnitude. Except for the *Manufacturing* and the *oil, gas and coal industries*, this is mainly driven by the increase in manipulation cost. *Healthcare, Medical Equipment, and Drug* has the highest average manipulation in both periods, although the post SOX manipulation cost is large, the price response increases to outweigh the cost. *Wholesale, Retail and Some Service* has the lowest average bias among all industries. The average bias are relatively similar among *Consumer NonDurables*, *Consumer Durables*, *Chemicals and Allied Products* and *Business Equipment* industries.

Figure 8: Average Manipulation by Industry

Size and growth

We pursue our analysis by further estimating the manipulation when we take into account two major sources of heterogeneity across firms: size and growth opportunities. We use firms' market value equity (mve) in the previous year as the measure of size, and growth opportunity is proxied by book-to-market ratio (btm) in the previous year. We sort the sample by sizes and growth opportunities and generate nine portfolios which are re-balanced every year.

Table 7 summarizes the estimated θ for the nine portfolios. Among the three samples, the manipulation cost is larger for firms with lower growth opportunities and smaller for larger firms. Firms with low growth opportunities (large btm) experience higher cost when they are larger. In particular, the difference between large and small sized portfolios (LmS) is the lowest for firms with low growth opportunities, thereby implying the highest cost. The cost of manipulation increases post SOX relatively to pre SOX for all portfolios.

Table 7: Estimated θ by Size and Growth Opportunity

Panel A. Pre-SOX Period					
		Book-to-Market			
		Low	Medium	High	<i>HmL</i>
Size	Small	0.021	0.005	0.001	<i>-0.020</i>
	Medium	0.019	0.006	0.004	<i>-0.015</i>
	Large	0.129	0.029	0.029	<i>-0.100</i>
	<i>LmS</i>	<i>0.107</i>	<i>0.024</i>	<i>0.027</i>	
Panel B. Post-SOX Period					
		Book-to-Market			
		Low	Medium	High	<i>HmL</i>
Size	Small	0.002	0	0	<i>-0.002</i>
	Medium	0.006	0	0	<i>-0.006</i>
	Large	0.027	0.005	0.001	<i>-0.026</i>
	<i>LmS</i>	<i>0.025</i>	<i>0.005</i>	<i>0.001</i>	
Panel C. Aggregate Period					
		Book-to-Market			
		Low	Medium	High	<i>HmL</i>
Size	Small	0.016	0.002	0	<i>-0.016</i>
	Medium	0.018	0.004	0.001	<i>-0.017</i>
	Large	0.062	0.014	0.004	<i>-0.058</i>
	<i>LmS</i>	<i>0.046</i>	<i>0.012</i>	<i>0.004</i>	

Table 8 reports the implied manipulation of each portfolio when the underlying earnings are 0. Consistent with the evidence on the cost of manipulation, firms with larger sizes and higher growth opportunities manipulate more. Post SOX, the implied manipulation has significantly decreased mainly due to the increase in the cost of manipulation. These results are consistent with the literature: larger firms might be exposed to more capital market pressure and are more likely to manage their earnings than smaller firms (e.g., Myers et al. (2007), Nelson et al. (2002)).

Table 8: Implied Manipulation at 0 by Size and Growth Opportunity

Panel A. Pre-SOX Period					
Book-to-Market					
		Low	Medium	High	<i>HmL</i>
Size	Small	0.032	0.008	0.002	-0.030
	Medium	0.029	0.01	0.004	-0.024
	Large	0.244	0.034	0.029	-0.216
	<i>LmS</i>	0.212	0.026	0.027	
Panel B. Post-SOX Period					
Book-to-Market					
		Low	Medium	High	<i>HmL</i>
Size	Small	0.003	0	0	-0.003
	Medium	0.016	0	0	-0.016
	Large	0.058	0.01	0.002	-0.056
	<i>LmS</i>	0.055	0.01	0.002	
Panel C. Aggregate Period					
Book-to-Market					
		Low	Medium	High	<i>HmL</i>
Size	Small	0.026	0.003	0	-0.026
	Medium	0.03	0.006	0.002	-0.028
	Large	0.103	0.023	0.006	-0.097
	<i>LmS</i>	.077	0.021	0.006	

4.3. Unobserved cost heterogeneity

We conduct the estimation by considering the alternative assumption that $1/\theta$ is a random variable that is observed only by the firm, using the integral of the likelihood developed earlier. Specifically, we assume that $1/\theta$ is exponential with truncation \underline{c} and mean c_m . We estimate the truncation but requiring that the second-order condition be satisfied, that is, $\underline{c} \geq \max \frac{r_i}{2\hat{\gamma}''(r_i)}$. The log likelihood of the total sample can be written as:

$$\mathcal{L}(\underline{c}, c_m, \mu_x, \sigma_x) = \sum_{i=1}^N \log \left(\int (1 - \gamma''(r_i) \frac{\theta}{2}) f(R^{-1}(r, \theta)) h(\theta) d\theta \right). \quad (4.5)$$

The estimation on the distribution of private earnings signals provides some interesting findings about the distribution of cost. In the post-SOX period, the lower bound of manipulation cost \underline{c} increases while the distribution is more concentrated as the mean is

lower. The average bias reduces significantly, from 0.055 before SOX to 0.007 after SOX.

Table 9: Exponential distribution with unknown boundary

	Mean c_m	Lower bound \underline{c}	Mean $_x$ μ_x	Average Bias $\mathbb{E}(r) - \mu_x$	Standard deviation σ_x	Likelihood
Pre-SOX (1976-2001)	1795.20 (632.41)	25.96 (9.58E+04)	-0.054 (0.015)	0.055	0.007 (0.003)	2.89
Post-SOX (2003-2014)	388.89 (91.13)	191.89 (32.59)	-0.009 (0.002)	0.007	0.012 (0.000)	2.98
Aggregate (1976-2014)	713.94 (232.25)	54.99 (22.49)	-0.030 (0.006)	0.027	0.010 (0.001)	2.98

Another important fact is that the random cost model, when used in conjunction with an exponential conjecture about the distribution, does not fit the observed distribution of earnings better than the fixed cost model. The likelihood is very similar and actually slightly lower in this case. In particular, this counter-intuitive implication suggests that more work might be needed to narrow down a suitable parametric specification for the random model.

5. Concluding remarks

An attempt to reconcile with (a few) existing approaches to identifying manipulation

In this study, we provide the simplest approach to identifying misreporting cost in a signalling model, but an approach that nevertheless brings into the estimation flexible information contained in reported earnings and prices. Our main objective here is not as much to provide definite estimate of manipulation but to open the door to the formalization of richer economic assumptions on a choice process, as a basis for estimation. In doing so, we demonstrate how this approach can offer predictions which earnings reports are manipulated with quantifiable magnitudes, and offer researchers with a tool for conducting simple counterfactuals. Furthermore, it provides an empirical path to rejecting the existence of manipulation if it is indeed true that manipulation has become small - noting that our current results are far too limited to reach such a strong conclusion.

?

As such our analysis contains many caveats that we hope may offer paths for continuing work in this area. First, by using earnings announcement surprises, we do not use any information that comes from predictable reversals, i.e., dynamics predictions implied by the manipulation over time, or information from restatements. Second, a more complete of other choices, such as investment, may help researchers separate accrual from real manipulation, also in terms of choice process. Third, the setting offers many possible paths to parametrize the model, such as misreporting costs that depend on true earnings or other observable characteristics.

In the second part of this conclusion, we attempt to reconcile our approach with prior approaches, and offer some thoughts as to future research may come closer to merging insights from the various strands of research on manipulation. Note that our objective is not to offer a comprehensive survey of all studies on manipulation, but to obtain some guidance as to how to build models that integrate all prior results which, admittedly, take the current approach as the starting block.

Abnormal accruals. A vast literature relies on abnormal accruals, broadly defined as the residual from earnings minus cash flows regressed on a set of explanatory variables, as a proxy for manipulation (see Dechow, Ge and Schrand 2010 for various examples). The assumptions that underlies this approach is that (i) cash is hard to manipulate, (ii) earnings should be “cleaned” of economic determinants of earnings, so that what remains may contain a greater portion of manipulation. This approach does not offer measures of actual manipulation but, nevertheless, its insights are relevant to implementing our estimation. Specifically, under the same assumptions, one may prefer to use abnormal accruals as the input variable for earnings and price responses which would serve as a refinement of our approach and, incrementally, help researchers derive the implied effective manipulation implied by these models.

Restatement-based. Another set of approaches involves predictive models of restatements or frauds Dechow, Ge and Schrand (2010), which allows researchers to capture a parsimonious set of variables that may indicate accounting risks. Within a structural model, Zakolyukina (2014) uses data from actual restatements to capture the probability that a fraud is eventually observed (and when). The detection probability is, implicitly, contained in reduced-form in terms of manipulation cost; nevertheless, it may be useful to test whether the predicted manipulation in our model appears to correlate with the restatements and frauds.

Time-series. Various other approaches rely on time-series properties of earnings and/or cash flows to identify manipulation. We discuss, below, a few notable examples. In Dechow and Dichev (2002), next-period cash flows used in a regression analysis can help identify current earnings that are not converted into earnings. This approach is further refined by Bloomfield, Gerakos and Kovrijnykh (2015) who measure the speed at which earnings are converted into cash payments. These approaches can allow for an ex-post match of manipulation predicted ex-ante to whether, later on, we observe lower cash flows. In Gerakos and Kovrijnykh (2013), reversal of the current misreported earnings introduces a second-order auto-correlation in earnings. While we do not specifically model persistence in earnings or reversals, their model suggests that we should observe the highest second-order auto-correlation for earnings levels with the greatest predicted manipulation. Within a different model, Nikolaev (2016) models the persistence of earnings and cash flows to identify the noise in accounting numbers. Lastly, Beyer, Guttman and Marinovic (2012) estimate a generalization of Dye and Sridhar (2004) with persistence in earnings and forward-looking managers, and demonstrate that manipulation can be estimated from a regression of prices on past prices and past earnings. Within their dynamic setting, we should conduct the estimation by histories of past prices and earnings. To summarize, these approaches suggest different manners we should group observations together in terms of different time-series history.

Appendix

A.1. Simulation

Below, we provide a simple example that shows how the model extracts information from the cross-section of earnings to recover information about manipulation costs. For expositional reasons, this is most apparent using a simple moment restriction implied by (2.8) and making the (infeasible) assumption that we know the distribution of true earnings, say, $x \sim N(0, 1)$ but do not know θ .¹⁴

Define m as the mean of the accounting reports, which must satisfy:

$$0 = m_1 - \int g(r)rdr,$$

so that we can select estimates for θ that minimize $(\hat{m}_1 - \int \hat{g}(r)rdr)^2$ where \hat{m}_1 is the sample mean of reported earnings and $\hat{g}(r)$ is obtained from (2.8).

¹⁴We use this assumption as an expositional tool, similar to Arcidiacono and Miller (2011), to take aside the estimation of this distribution. For the baseline model with more than a single unknown parameter, it is more practical to use the entire density.

To validate and illustrate this approach, as in Gow, Larcker and Reiss (2015), we simulate a sample of earnings and market responses from a researcher-known manipulation cost, and then, verify whether our GMM estimation procedure can precisely re-estimate the cost using only the simulated sample. This laboratory setting has two objectives: (i) to (numerically) verify identification in our model, and (ii) to illustrate conceptually how the estimator recovers the cost. We generate a sample with the following additional assumptions:

1. The parameter of the manager's engaging management is $\theta = 0.5$ and the function $\psi(\cdot)$ is a quadratic function, i.e., $\psi(r - x) = (r - x)^2$.
2. The market response is non-linear due to learning about the mean and variance of economic value. Specifically, Subramanyam (1996) and Kirschenheiter and Melumad (2002) show that if earnings are a signal about true value, about which the prior is normally distributed with unknown mean and variance, then the price response is S-shaped with a form

$$\gamma(r) = \frac{\Gamma(r + 1.5, 1)}{\Gamma(r + 0.5, 1)}, \quad (5.1)$$

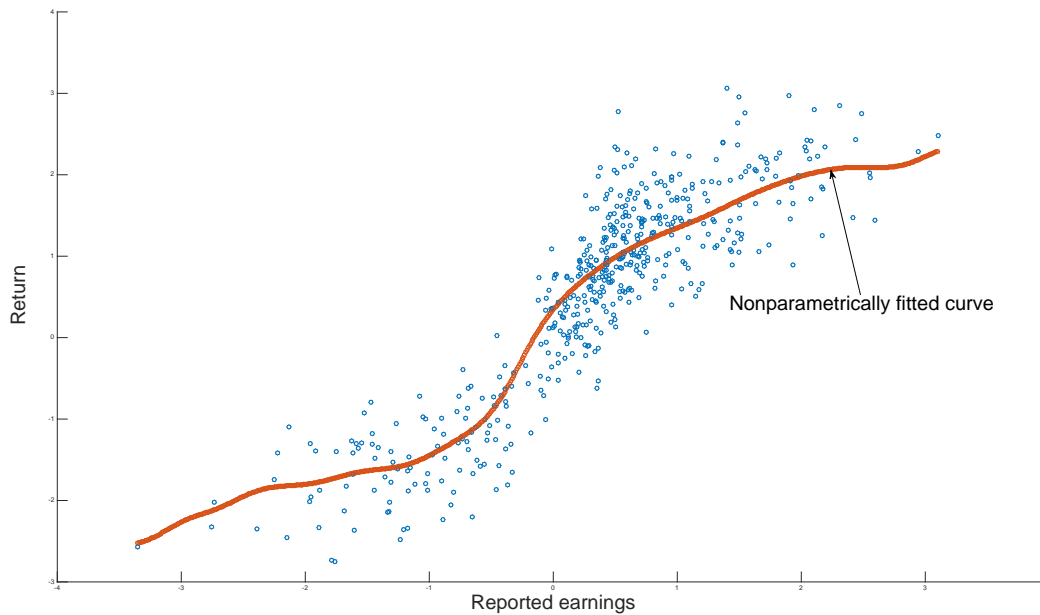
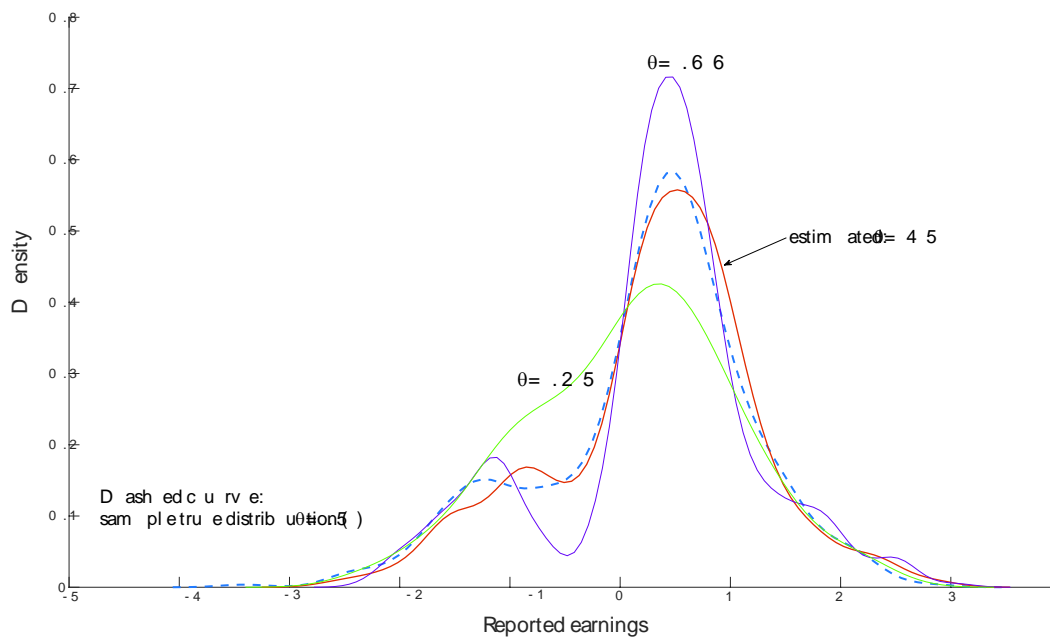
where e is the reported earnings and $\Gamma(\cdot, 1)$ represents the incomplete gamma integral with support $(0, 1)$. We have also tried - and the results are very similar - using the ad-hoc arctan S-shape in Freeman and Tse (1992).

We create a sample with $\theta = 0.5$ and 2,500 observations for the market response and earnings. We simulate the reported earnings by solving the implicit equation (2.5), and, lastly, the simulated market response is the market response defined in equation (5.1) plus some noise to make the estimation less than ideal.

The simulated market response or return and reported earnings are given in a scatter plot in Figure 1. The dots represent (reported earnings, market response) pairs and the red line is a non-parametric fit of this curve using the Nadarya-Watson estimator with optimal bandwidth. Note that this curve is estimated non-parametrically because we do not assume that the true nature of the price function in (5.1) is known to the estimation.

We compute the model-implied $\hat{g}(\cdot)$ from equation (2.8), which depends on the estimated price response $\hat{\gamma}$ and its first and second derivatives. Finally, the moment is matched to recover an estimated $\hat{\theta}$.

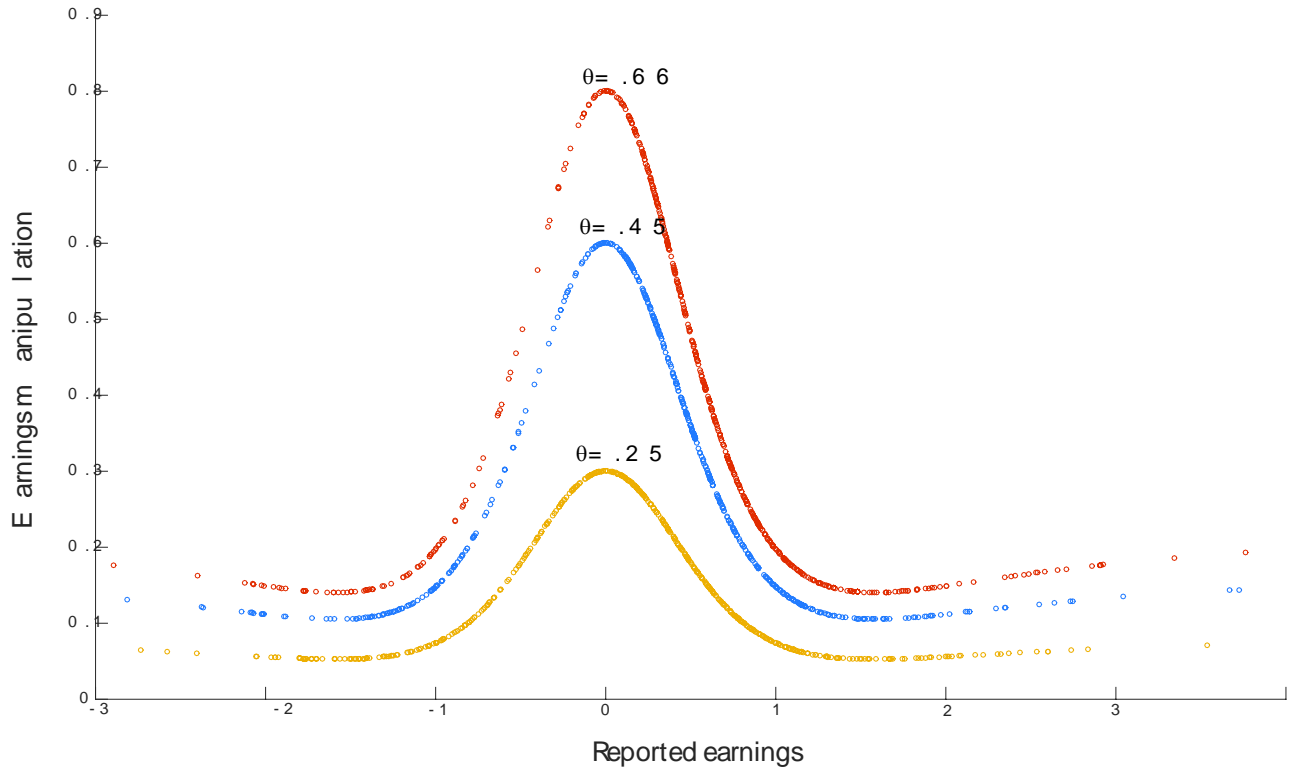
In Figure 2, we illustrate how the estimation procedure seeks θ to match the observed distribution of earnings and the estimated model-implied distribution. We plot the sample density of earnings in dots which, by construction, features a hump-shape form. The price

Figure 9: Market Response to Earnings**Figure 10: Finding the true θ** 

benefit $\hat{\theta} = 0.450$, in bold, yields a predicted density that is close to its true value set in the simulation. Consider other choices of $\hat{\theta}$, that do not fare as well. For high $\hat{\theta} = 0.25$, the distribution is much smoother and looks more like a hill. For high $\hat{\theta} = 0.666$, on the other hand, we observe an extreme skewness to the right of the distribution. Neither of these choices match the dotted line very well.

In Figure 3, we illustrate the model-implied manipulation $r - x$ for any observed report r . The bold curve illustrates the manipulation for our baseline simulation with $\theta = 0.450$; this is compared to the actual manipulation if one knew the true $\theta = 0.5$. To get some further idea about the sensitivity of the measure to other values of θ , we simulate and estimate two samples, one with $\theta = 0.666$ and one with $\theta = 0.25$. Perhaps the most interesting aspect of these plots is that the earnings manipulation is indeed strongest around the threshold, but it is relatively spread out both before, and after the threshold.

Figure 11: Magnitude of earning manipulation



A.2. Omitted proofs

Proof Lemma 2.1:

We show that (θ, α, β) is not identified; when $\psi(x) = x^2$ and (i) the true distribution of earnings is of the form $f(x) = \frac{1}{\beta}k(\frac{x-\alpha}{\beta})$ and (ii) $\gamma(x) = v_0 + v_1x$:
The distribution of reported earnings $g(r)$ is equal to:

$$g(r) = f(r - \frac{\theta}{2}v_1) \quad (5.2)$$

$$g(r) = \frac{1}{\beta}k(\frac{r - \frac{\theta}{2}v_1 - \alpha}{\beta}). \quad (5.3)$$

Let us define two vectors solution to equation (5.3) such that $(\alpha_1, \beta_1, \theta_1) \neq (\alpha_2, \beta_2, \theta_2)$. We define a report r_0 such that $\frac{r - \frac{\theta_1}{2}v_1 - \alpha_1}{\beta_1} = \frac{r_0 - \frac{\theta_2}{2}v_1 - \alpha_2}{\beta_2}$. We obtain:

$$g(r) = \frac{\beta_2}{\beta_1 z_2} g(r_0) \quad (5.4)$$

$$g(r) = \frac{\beta_2}{\beta_1} g(\frac{\beta_2}{\beta_1}(r - \alpha_1 - \frac{\theta_1}{2}v_1) + \frac{\theta_2}{2}v_1 + \alpha_2) \quad (5.5)$$

We define $d_1 = \frac{\beta_2}{\beta_1}$ and $d_2 = -\frac{\beta_2}{\beta_1}(\frac{\theta_1}{2}v_1 + \alpha_1) + \frac{\theta_2}{2}v_1 + \alpha_2$, and rewrite $g(r) = d_1 g(d_1 r + d_2)$. We know that $d_1 > 0$ because $g(r) > 0$.

Let $r_{max} \in \text{argmax } g(r)$ and $g(r)$ be single peaked.

If $d_1 < 1$, $g(d_1 r_{max} + d_2) = g(r_{max})/d_1 > g(r_{max})$, which is a contradiction.

If $d_1 > 1$, $r_{max} = d_1 r^* + d_2$ and $d_1 g(d_1 r^* + d_2) = g(r^*) > g(r_{max})$, which is a contradiction.

Thus $d_1 = 1$ and $g(r_{max}) = g(r_{max} + d_2)$, which can only be true if $d_2 = 0$.

If $d_1 = 1$, $\beta_1 = \beta_2$. If $d_2 = 0$, $\frac{\theta_2}{2}v_1 + \alpha_2 = \frac{\theta_1}{2}v_1 + \alpha_1$. Hence we cannot identify (β, α, θ) .

We show that we can only identify only θ/β and α/β when $\psi(x) = x^2$ and (i) the true distribution of earnings is of the form $f(x) = \frac{1}{\beta}k(\frac{x-\alpha}{\beta})$ and (ii) $\gamma(x) = v_0 + v_1x + v_2x^2$.
The distribution of reported earnings $g(r)$ is equal to:

$$g(r) = (1 - v_2\theta)f(r(1 - v_2\theta) - \frac{\theta}{2}v_1) \quad (5.6)$$

$$g(r) = (1 - v_2\theta)\frac{1}{\beta}k(\frac{r(1 - v_2\theta) - \frac{\theta}{2}v_1 - \alpha}{\beta}). \quad (5.7)$$

We recast the same reasoning as before. Let us define $z = (1 - v_2\theta)$, $g(r) = \frac{z}{\beta}k(\frac{rz}{\beta} - \frac{1}{\beta}(\frac{\theta}{2}v_1 + \alpha))$. Now let us define two vectors solution to equation (5.7) such that $(\alpha_1, \beta_1, z_1, \theta_1) \neq (\alpha_2, \beta_2, z_2, \theta_2)$. We define a report r_0 such that $\frac{rz_1}{\beta_1} - \frac{1}{\beta_1}(\frac{\theta_1}{2}v_1 + \alpha_1) = \frac{r_0 z_2}{\beta_2} - \frac{1}{\beta_2}(\frac{\theta_2}{2}v_1 + \alpha_2)$. We obtain:

$$g(r) = \frac{\beta_2 z_1}{\beta_1 z_2} g(r_0) \quad (5.8)$$

$$g(r) = \frac{\beta_2 z_1}{\beta_1 z_2} g\left(\frac{\beta_2}{z_2} \left(\frac{r z_1}{\beta_1} - \frac{1}{\beta_1} \left(\frac{\theta_1}{2} v_1 + \alpha_1\right) + \frac{1}{\beta_2} \left(\frac{\theta_2}{2} v_1 + \alpha_2\right)\right)\right) \quad (5.9)$$

We define $d_1 = \frac{\beta_2 z_1}{\beta_1 z_2}$ and $d_2 = -\frac{\beta_2}{\beta_1 z_2} \left(\frac{\theta_1}{2} v_1 + \alpha_1\right) + \frac{1}{z_2} \left(\frac{\theta_2}{2} v_1 + \alpha_2\right)$, and rewrite $g(r) = d_1 g(d_1 r + d_2)$. We know that $d_1 > 0$ because $g(r) > 0$.

Let $r_{max} \in \operatorname{argmax} g(r)$ and $g(r)$ be single peaked.

If $d_1 < 1$, $g(d_1 r_{max} + d_2) = g(r_{max})/d_1 > g(r_{max})$, which is a contradiction.

If $d_1 > 1$, $r_{max} = d_1 r^* + d_2$ and $d_1 g(d_1 r^* + d_2) = g(r^*) > g(r_{max})$, which is a contradiction.

Thus $d_1 = 1$ and $g(r_{max}) = g(r_{max} + d_2)$, which can only be true if $d_2 = 0$.

If $d_1 = 1$, $\frac{\theta_1}{\beta_1} = \frac{\theta_2}{\beta_2}$. If $d_2 = 0$, $-\frac{1}{z_1} \left(\frac{\theta_1}{2} v_1 + \alpha_1\right) + \frac{\theta_1 v_1}{2 z_1} + \frac{\beta_1 \alpha_2}{\beta_2 z_1} = 0$. Replacing $z_1 = (1 - v_2 \theta_1)$, $\frac{\alpha_1}{\beta_1} = \frac{\alpha_2}{\beta_2}$.

Proof Lemma 2.2: The reported distribution equals:

$$g(r) = (1 - \gamma''(r)\theta) f\left(r - \frac{\theta}{2} \gamma'(r)\right) \quad (5.10)$$

$$(5.11)$$

Because it is symmetric around m : $f\left(m + r - \frac{\theta}{2} \gamma'(r + m)\right) = f\left(m - r - \frac{\theta}{2} \gamma'(r - m)\right)$.

$$g(m + r) = (1 - \gamma''(r + m)\theta) f\left(m + r - \frac{\theta}{2} \gamma'(r + m)\right)$$

$$g(m - r) = (1 - \gamma''(r - m)\theta) f\left(m - r - \frac{\theta}{2} \gamma'(r - m)\right)$$

If $(1 - \gamma''(r + m)\theta) = (1 - \gamma''(r - m)\theta)$, $g(m + r) = g(m - r)$, and θ is not identified.

Thus if $\gamma''(\cdot)$ is a constant or $\gamma(\cdot)$ is symmetric around m , θ is not identified. Otherwise,

$g(m + r) = \frac{1 - \gamma''(r + m)\theta}{1 - \gamma''(r - m)\theta} g(m - r)$ and θ is identified.

References

Arcidiacono, Peter and Robert A Miller, “Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity,” *Econometrica*, 2011, 79 (6), 1823–1867.

Bertomeu, Jeremy, Paul Ma, and Iv'an Marinovic, “How often do managers withhold information?,” *Stanford GSB Working Paper*, 2015.

- Beyer, Anne, Ilan Guttman, and Ivan Marinovic**, “Optimal Contracts with Performance Manipulation,” 2012.
- , – , and **Iván Marinovic**, “Earnings management and earnings quality: Theory and evidence,” *Available at SSRN 2516538*, 2014.
- Bloomfield, Matthew J, Joseph J Gerakos, and Andrei Kovrijnykh**, “Accrual reversals and cash conversion,” 2015.
- Breuer, Matthias and David Windisch**, “Investment Dynamics and Earnings-Return Properties: A Structural Approach,” 2017.
- Burgstahler, David and Ilia Dichev**, “Earnings management to avoid earnings decreases and losses,” *Journal of accounting and economics*, 1997, *24* (1), 99–126.
- Cho, In-Koo and David M Kreps**, “Signaling games and stable equilibria,” *The Quarterly Journal of Economics*, 1987, pp. 179–221.
- Dechow, Patricia M and Ilia D Dichev**, “The quality of accruals and earnings: The role of accrual estimation errors,” *The accounting review*, 2002, *77* (s-1), 35–59.
- Dechow, Patricia, Weili Ge, and Catherine Schrand**, “Understanding earnings quality: A review of the proxies, their determinants and their consequences,” *Journal of accounting and economics*, 2010, *50* (2), 344–401.
- Degeorge, Francois, Jayendu Patel, and Richard Zeckhauser**, “Earnings Management to Exceed Thresholds,” *Journal of Business*, 1999, *72* (1), 1–33.
- Dutta, Sunil and Alexander Nezlobin**, “Dynamic Effects of Information Disclosure on Investment Efficiency,” *Journal of Accounting Research*, 2016.
- Dye, Ronald A.**, “Earnings Management in an Overlapping Generations Model,” *Journal of Accounting Research*, 1988, *26* (2), 195–235.
- and **Sri Sridhar**, “Reliability-Relevance Trade-Offs and the Efficiency of Aggregation,” *Journal of Accounting Research*, 2004, *42* (1), 51–88.
- Einhorn, Eti and Amir Ziv**, “Biased voluntary disclosure,” *Review of Accounting Studies*, 2012, *17* (2), 420–442.
- Fischer, Paul E. and Robert E. Verrecchia**, “Reporting Bias,” *The Accounting Review*, 2000, *75* (2), 229–245.

- Frankel, Alex and Navin Kartik**, “Muddled Information,” *Columbia University*, 2014.
- Freeman, Robert N. and Senyo Y. Tse**, “A Nonlinear Model of Security Price Responses to Unexpected Earnings,” *Journal of Accounting Research*, 1992, *30* (2), 185–209.
- Gayle, George-Levi, Chen Li, and Robert A Miller**, “Was Sarbanes-Oxley Costly? Evidence from Optimal Contracting on CEO Compensation,” 2015.
- Gerakos, Joseph and Andrei Kovrijnykh**, “Performance shocks and misreporting,” *Journal of Accounting and Economics*, 2013, *56* (1), 57–72.
- Gilliam, Thomas A, Frank Heflin, and Jeffrey S Paterson**, “Evidence that the zero-earnings discontinuity has disappeared,” *Journal of Accounting and Economics*, 2015, *60* (1), 117–132.
- Gow, Ian D, David Larcker, and Peter Reiss**, “Causal Inference in Accounting Research,” Technical Report, 2015 Conference of the Journal of Accounting Research 2015.
- Guttman, Ilan, Ohad Kadan, and Eugene Kandel**, “A Rational Expectations Theory of Kinks in Financial Reporting,” *The Accounting Review*, 2006, *81* (4), 811–848.
- , – , and – , “Dividend stickiness and strategic pooling,” *Review of Financial Studies*, 2010, *23* (12), 4455–4495.
- Hemmer, Thomas and Eva Labro**, “Management by the Numbers: A Formal Approach to Deriving Informational and Distributional Properties of Un-Managed Earnings,” 2015.
- Hiemann, Moritz**, “Earnings and Firm Value in the Presence of Real Options,” 2016.
- Jones, Jennifer J**, “Earnings management during import relief investigations,” *Journal of accounting research*, 1991, pp. 193–228.
- Kirschenheiter, Michael T. and Nahum D. Melumad**, “Can Big Bath and Earnings Smoothing Co-Exist as Equilibrium Financial Reporting Strategies?,” *Journal of Accounting Research*, 2002, *40* (2), 761–796.
- Li, Chen**, “Are top management teams compensated as teams? A Structural Modeling Approach,” in “in” 2016.

- Magnac, Thierry and David Thesmar**, “Identifying dynamic discrete decision processes,” *Econometrica*, 2002, *70* (2), 801–816.
- Myers, James N, Linda A Myers, and Douglas J Skinner**, “Earnings momentum and earnings management,” *Journal of Accounting, Auditing & Finance*, 2007, *22* (2), 249–284.
- Nelson, Mark W, John A Elliott, and Robin L Tarpley**, “Evidence from auditors about managers’ and auditors’ earnings management decisions,” *The accounting review*, 2002, *77* (s-1), 175–202.
- Newell, John and Jochen Einbeck**, “A comparative study of nonparametric derivative estimators,” in “Proc. of the 22nd International Workshop on Statistical Modelling” 2007.
- Nikolaev, Valeri V**, “Identifying Accounting Quality,” *Chicago Booth Research Paper*, 2014, (14-28).
- , “Identifying accounting quality,” 2016.
- Rajan, Madhav V, Stefan Reichelstein, and Mark T Soliman**, “Conservatism, growth, and return on investment,” *Review of Accounting Studies*, 2007, *12* (2-3), 325–370.
- Rust, John**, “Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher,” *Econometrica: Journal of the Econometric Society*, 1987, pp. 999–1033.
- Stein, Jeremy C.**, “Efficient Capital Markets, Inefficient Firms: A Model of Myopic Corporate Behavior,” *Quarterly Journal of Economics*, 1989, *104* (4), 655–669.
- Subramanyam, KR**, “Uncertain precision and price reactions to information,” *Accounting Review*, 1996, pp. 207–219.
- Terry, Stephen**, “The macro impact of short-termism,” *Work. Pap.*, 2015.
- Zakolyukina, Anastasia A**, “Measuring intentional gaap violations: A structural approach,” *Chicago Booth Research Paper*, 2014, (13-45).
- Zhou, Frank**, “Disclosure Dynamics and Investor Learning,” 2016.