Example 1 *Earnings management: reserve accounting revisited*

Revisit the reserve accounting-based earnings management example based on Demski [2004] initially discussed in chapter 3. The objective is again to track the relation between a firm's value P_t and its accruals z_t. Everything remains as in that setting except the entrepreneur or manager with an equity stake endogenously manipulates reserves subject to auditor discipline but the auditor can never fully eliminate mis-reporting (described below). As this is a multiperiod setting, the entrepreneur or manager continues to retain equity ownership with intermittent and stochastic forced or liquidity sales of a fraction of his holdings.

The entrepreneur reveals his private information y_t^p through income I_t (cash flows plus change in accruals) where fair value accruals

$$z_t = E\left[\widetilde{d}_{t+1} \mid \widetilde{y}_t^p = y_t^p\right] = \frac{1}{2}y_t^p$$

are reported, income is

$$\begin{aligned} I_t &= d_t + (z_t - z_{t-1}) \\ &= d_t + \frac{1}{2}\left(y_t^p - y_{t-1}^p\right) \end{aligned}$$

Since $\widetilde{y}_t^p = \widetilde{d}_{t+1} + \widetilde{\varepsilon}_t$ where $\widetilde{d}_{t+1}$ and $\widetilde{\varepsilon}$ are Normal *iid* with mean zero and variance σ^2, the *DGP* for dividends given accruals is

$$\begin{aligned} \widetilde{d}_{t+1} &= \widetilde{z}_t + \nu_t \\ &= \frac{1}{2}y_t^p + \nu_t \end{aligned}$$

where ν_t are Normal *iid* with mean zero and variance $\frac{1}{2}\sigma^2$. With no mis-reporting, there is a linear relation between price and fair value accruals

$$\begin{aligned} P_t &\equiv E\left[\widetilde{d}_{t+1} \mid \widetilde{y}_t^p = y_t^p\right] \\ &= \frac{1}{2}y_t^p \end{aligned}$$

Now, suppose the entrepreneur mis-reports via θ. The auditor naturally focuses on upward mis-statement and θ is limited to $\frac{1}{2}\Delta$ as anything smaller is typically undetectable (except as noted below). However, the entrepreneur may choose to increase reserves for future periods when he is unable to upward mis-report (think "big bath").[1] As investors naturally price protect against upward mis-reporting, occasional downward mis-reporting builds reserves and consequently increases the opportunity for the entrepreneur to follow through as anticipated in future periods. The idea is that if current results are to be

[1] Financial reports routinely attempt to differentiate downward accrual adjustments such as "unusual" charges or write downs from results of "normal" or continuing operations.

understated in favor of reserves for the future, then the entrepreneur wants to increase the likelihood that investors recognize the understatement and avoid any stock price penalty. The auditor restricts downward mis-reports to -5Δ. Finally, there is an interior region in which the auditor limits upward mis-reporting to a boundary on reserves $\underline{r}$; hence accrual mis-reporting is less than $\frac{1}{2}\Delta$ in this partial upward reporting region. Reported accruals are

$$z_t = \frac{1}{2}y_t^p + \theta_t$$

Accrual manipulation is limited by auditor-disciplined reserves, where favorable (unfavorable) results increase (decrease) reserves and upward (downward) mis-reports decrease (increase) reserves

$$R_t = R_{t-1} + \frac{1}{2}y_t^p - \theta_t$$

Further, accruals may not be upwardly manipulated so that reserves fall below some auditor detectable limit $R \geq \underline{r}$. Putting this together the resultant mis-reporting (and shorthand notation) is

$$\theta_t \leq \frac{1}{2}\Delta \text{ if } \underline{r} \leq R_{t-1} + \frac{1}{2}y_t^p - \frac{1}{2}\Delta \qquad (\theta_t^u)$$

$$\theta_t \geq -5\Delta \text{ if } \underline{r} \geq R_{t-1} + \frac{1}{2}y_t^p \qquad (\theta_t^d)$$

$$\theta_t = \frac{\gamma_t}{2}\Delta \text{ if } \underline{r} - R_{t-1} < \frac{1}{2}y_t^p < \underline{r} - R_{t-1} + \frac{1}{2}\Delta \qquad (\theta_t^0)$$

where $\gamma_t = \alpha_t + (1-\alpha_t)\frac{R_{t-1} - \underline{R}_{t-1}^0}{\overline{R}_{t-1}^0 - \underline{R}_{t-1}^0}$ and $\alpha_t = \max\left\{\underline{r}, \frac{\underline{R}_{t-1}^0 - \underline{r} + z_t}{\Delta}\right\}$.

Investors process the entrepreneur's report with mis-reporting in mind. The likelihood of mis-reporting depends on the distribution of y_t^p, the distribution of reserves R_{t-1}, and Δ. Assume the distribution of reserves is uniformly distributed over the interval $\left[\underline{R}_{t-1}, \overline{R}_{t-1}\right]$, the probability of upward mis-reporting given accrual report z_t is

$$\begin{aligned} p_t^u(z_t) &\equiv \Pr\left(\theta_t = \frac{1}{2}\Delta \mid \widetilde{z}_t = z_t, \underline{R}_{t-1} \leq R_{t-1} \leq \overline{R}_{t-1}\right) \\ &= \Pr\left(\underline{r} \leq R_{t-1} + z_t - \Delta \mid \widetilde{z}_t = z_t, \underline{R}_{t-1} \leq R_{t-1} \leq \overline{R}_{t-1}\right) \\ &= \frac{\int_{\underline{R}_{t-1}}^{\overline{R}_{t-1}} \Pr\left(\underline{r} \leq R_{t-1} + z_t - \Delta \mid \widetilde{z}_t = z_t, R_{t-1}\right) \frac{f(z_t \mid R_{t-1})}{\overline{R}_{t-1} - \underline{R}_{t-1}} dR_{t-1}}{f\left(z_t, \underline{R}_{t-1} \leq R_{t-1} \leq \overline{R}_{t-1}\right)} \end{aligned}$$

where

$$f\left(z_t, \underline{R}_{t-1} \leq R_{t-1} \leq \overline{R}_{t-1}\right) = \int_{\underline{R}_{t-1}}^{\overline{R}_{t-1}} \frac{f(z_t \mid R_{t-1})}{\overline{R}_{t-1} - \underline{R}_{t-1}} dR_{t-1}$$

a density function. The lead term $\Pr\left(\underline{r} \leq R_{t-1} + z_t - \Delta \mid \widetilde{z}_t = z_t, R_{t-1}\right)$ is an indicator function $I\left(\theta_t^u\right) \equiv I\left(\theta_t^u \mid z_t, R_{t-1}\right)$ and

$$\begin{aligned} f\left(z_t \mid R_{t-1}\right) &= \sum_{k \in \{u,0,d\}} f\left(z_t \mid \theta_t^k, R_{t-1}\right) \Pr\left(\theta_t^k \mid R_{t-1}\right) \\ &= \sum_{k \in \{u,0,d\}} \frac{\phi\left(\frac{z_t - \theta_t^k}{\sqrt{0.5}\sigma} \mid \theta_t^k, R_{t-1}\right) I\left(\theta_t^k\right)}{\Pr\left(\theta_t^k \mid R_{t-1}\right)} \Pr\left(\theta_t^k \mid R_{t-1}\right) \\ &= \sum_{k \in \{u,0,d\}} \phi\left(\frac{z_t - \theta_t^k}{\sqrt{0.5}\sigma} \mid \theta_t^k, R_{t-1}\right) I\left(\theta_t^k\right) \end{aligned}$$

a truncated standard normal density over the support for θ_t^k that is,

$$\begin{aligned} I\left(\theta_t^u\right) &= 1 \text{ if } \underline{R}_{t-1}^u \leq R_{t-1} \leq \overline{R}_{t-1}^u \\ &= 0 \quad \text{otherwise} \\ I\left(\theta_t^0\right) &= 1 \text{ if } \underline{R}_{t-1}^0 \leq R_{t-1} \leq \overline{R}_{t-1}^0 \\ &= 0 \quad \text{otherwise} \\ I\left(\theta_t^d\right) &= 1 \text{ if } \underline{R}_{t-1}^d \leq R_{t-1} \leq \overline{R}_{t-1}^d \\ &= 0 \quad \text{otherwise} \end{aligned}$$

where $\overline{R}_{t-1}^u = \overline{R}_{t-1}$, $\underline{R}_{t-1}^d = \underline{R}_{t-1}$, $\underline{R}_{t-1}^u = \overline{R}_{t-1}^0 = \max\left\{\underline{r} - z_t + \Delta, \underline{R}_{t-1}\right\}$, $\underline{R}_{t-1}^0 = \overline{R}_{t-1}^d = \max\left\{\underline{r} - z_t, \underline{R}_{t-1}\right\}$, and $\overline{R}_{t-1}^d - \underline{R}_{t-1}^d + \overline{R}_{t-1}^0 - \underline{R}_{t-1}^0 + \overline{R}_{t-1}^u - \underline{R}_{t-1}^u = \overline{R}_{t-1} - \underline{R}_{t-1}$. Integration and simplification yield

$$p_t^u\left(z_t\right) = \frac{\frac{\overline{R}_{t-1}^u - \underline{R}_{t-1}^u}{\overline{R}_{t-1} - \underline{R}_{t-1}} \phi\left(\frac{z_t - 0.5\Delta}{\sqrt{0.5}\sigma}\right)}{\Psi_t}$$

where $\phi(\cdot)$ is the standard normal density function, $\Phi(\cdot)$ is the cumulative standard normal distribution function,

$$\begin{aligned} \Psi_t &= \frac{\overline{R}_{t-1}^u - \underline{R}_{t-1}^u}{\overline{R}_{t-1} - \underline{R}_{t-1}} \phi\left(\frac{z_t - .5\Delta}{\sqrt{.5}\sigma}\right) + \frac{\overline{R}_{t-1}^d - \underline{R}_{t-1}^d}{\overline{R}_{t-1} - \underline{R}_{t-1}} \phi\left(\frac{z_t + 5\Delta}{\sqrt{.5}\sigma}\right) \\ &\quad + \sigma_t^0 \frac{\Phi\left(\frac{\overline{R}_{t-1}^0 - \mu_t^0}{\sigma_t^0}\right) - \Phi\left(\frac{\underline{R}_{t-1}^0 - \mu_t^0}{\sigma_t^0}\right)}{\overline{R}_{t-1} - \underline{R}_{t-1}} \end{aligned}$$

$\sigma_t^0 = \left|\frac{\sqrt{0.5}\sigma}{\frac{1-\alpha_t}{\overline{R}_{t-1}^0 - \underline{R}_{t-1}^0} \frac{1}{2}\Delta}\right|$, and $\mu_t^0 = \frac{z_t - \left[\alpha_t - (1-\alpha_t) \frac{\underline{R}_{t-1}^0}{\overline{R}_{t-1}^0 - \underline{R}_{t-1}^0}\right] \frac{1}{2}\Delta}{\sqrt{0.5}\sigma} \sigma_t^0$.

The probability of downward mis-reporting given accrual report z_t is

$$\begin{aligned} p_t^d(z_t) &\equiv \Pr\left(\theta_t = -5\Delta \mid \widetilde{z}_t = z_t, \underline{R}_{t-1} \leq R_{t-1} \leq \overline{R}_{t-1}\right) \\ &= \Pr\left(\underline{r} \geq R_{t-1} + z_t \mid \widetilde{z}_t = z_t, \underline{R}_{t-1} \leq R_{t-1} \leq \overline{R}_{t-1}\right) \\ &= \frac{\frac{\overline{R}_{t-1}^d - \underline{R}_{t-1}^d}{\overline{R}_{t-1} - \underline{R}_{t-1}} \phi\left(\frac{z_t + 5\Delta}{\sqrt{.5}\sigma}\right)}{\Psi_t} \end{aligned}$$

Clearly, the probability of partial upward mis-reporting over the interior region given an accruals report of z_t is $p_t^0(z_t) = 1 - p_t^u(z_t) - p_t^d(z_t)$.

For example, $p_t^u(z_t) = 1$ ($\overline{R}_{t-1}^d = \underline{R}_{t-1}^d = \overline{R}_{t-1}^0 = \underline{R}_{t-1}^0 = \underline{R}_{t-1}^u = \underline{R}_{t-1}$ and $\overline{R}_{t-1}^u = \overline{R}_{t-1}$) if the entrepreneur is able to upward report for all possible levels of reserves R_{t-1}. On the other hand, if upward reporting is possible only over a portion of the support for reserves then $p_t^u(z_t)$ and $p_t^0(z_t)$ are nonzero reflecting positive support of their respective regions. If the lower bound of reserve support based on investors' prior beliefs would result in current reserves below $\underline{r}$, then $\underline{R}_{t-1}$ is revised upward to eliminate the contradiction in investors' prior beliefs.[2]

Given investors' expectations and the entrepreneur's inability to credibly signal otherwise, the entrepreneur's equilibrium reporting strategy is to upwardly mis-report the maximum whenever possible, downwardly mis-report the maximum when reserves are exhausted, and partially upward mis-report to $\underline{r}$ when results lie in the band between. Hence, evaluated over the support for R_{t-1} the equilibrium price for the firm following a report of z_t is[3]

$$\begin{aligned} P_t &= E\left[\widetilde{d}_{t+1} \mid \widetilde{z}_t = z_t, \underline{R}_{t-1} \leq R_{t-1} \leq \overline{R}_{t-1}\right] \\ P_t &= p_t^u(z_t)(z_t - 0.5\Delta) + p_t^d(z_t)(z_t + 5\Delta) \\ &\quad + p_t^0(z_t)\left(z_t - 0.5\Delta\left[\alpha_t - \frac{(1-\alpha_t)\underline{R}_{t-1}^0}{\overline{R}_{t-1}^0 - \underline{R}_{t-1}^0}\right]\right) \\ &\quad - \frac{0.5\Delta(1-\alpha_t)}{\overline{R}_{t-1}^0 - \underline{R}_{t-1}^0}\left(\mu_t^0 p_t^0(z_t) - \left(\sigma_t^0\right)^2\left[\phi\left(\overline{w}_t^0\right) - \phi\left(\underline{w}_t^0\right)\right]\right) \end{aligned}$$

where $\overline{w}_t^0 = \frac{\overline{R}_{t-1}^0 - \mu_t^0}{\sigma_t^0}$ and $\underline{w}_t^0 = \frac{\underline{R}_{t-1}^0 - \mu_t^0}{\sigma_t^0}$. Price is no longer a linear function of reported accruals.

Consider the following simulation to illustrate. Let $\Delta = 1$, $\underline{r} = 0$, $\sigma = 1$, the entrepreneur knows beginning reserves $R_0 = 2$, and investors' perceive R_0

[2] There exists a contradiction as the implied level of y_t^p based on z_t and downward mis-reporting would allow upward mis-reporting - a contradiction. Based on the contradiction, investors infer a smaller region of reserve support.

[3] The expression for the partial upward mis-reporting interior region involves a standard truncated (at both ends) expected value for a normal random variable except that the function of the random variable R_{t-1} is a normal density kernel but not normalized for σ_t^0. Consequently, we have standard normal ordinates multiplied by $\left(\sigma_t^0\right)^2$ in the truncated expected value expression rather than the customary σ_t^0.

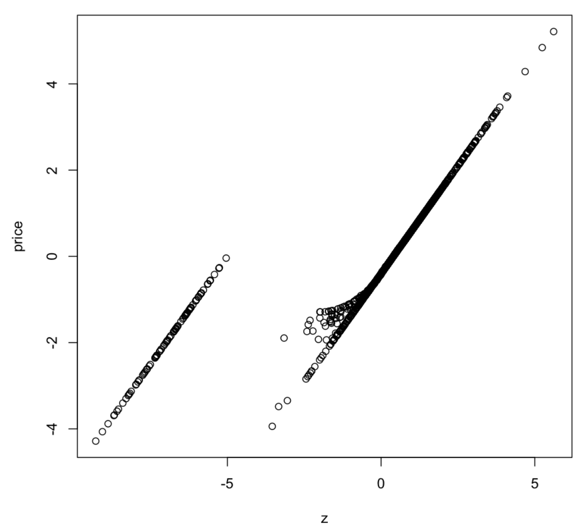
to be uniformly distributed over the interval $[0, 5]$. A panel data simulation is employed with $T = 10$ periods for each firm and $\frac{n}{T}$ firms in the sample. For sample size $n = 1,000$, $T = 10$, (100 firms) and 200 simulated samples, the regression of price on reported accruals is

$$P_t = \beta_0 + \beta_1 z_t$$

where

$$\begin{aligned}
z_t &= I\left(\theta_t^u\right)\left(\frac{1}{2}y_t^p + \frac{1}{2}\Delta\right) + I\left(\theta_t^d\right)\left(\frac{1}{2}y_t^p - 5\Delta\right) + I\left(\theta_t^0\right)\left(\frac{1}{2}y_t^p + \frac{\gamma_t}{2}\Delta\right) \\
I\left(\theta_t^u\right) &= 1 \quad \text{if } \underline{r} \leq R_{t-1} + \frac{1}{2}y_t^p - \frac{1}{2}\Delta \qquad (\theta_t^u) \\
&= 0 \quad \text{otherwise} \\
I\left(\theta_t^d\right) &= 1 \quad \text{if } \underline{r} \geq R_{t-1} + \frac{1}{2}y_t^p \qquad (\theta_t^d) \\
&= 0 \quad \text{otherwise} \\
I\left(\theta_t^0\right) &= 1 - I\left(\theta_t^u\right) - I\left(\theta_t^d\right) \\
&= 1 \quad \text{if } \underline{r} - R_{t-1} < \frac{1}{2}y_t^p < \underline{r} - R_{t-1} + \frac{1}{2}\Delta \qquad (\theta_t^0) \\
&= 0 \quad \text{otherwise}
\end{aligned}$$

A typical plot of the sampled data, price versus reported accruals is below. There is a piece-wise nonlinear pattern (again, there is a break in the middle à la Burgstahler and Dichev[1997]) in the data.



Price versus reported accruals (z)

Sample statistics for the regression estimates are tabulated below. The estimates of β_1 are substantially biased downward and the interval does not include $\beta_1 = 1$.

statistic	β_0	β_1
mean	0.085	0.424
median	0.086	0.423
standard deviation	0.015	0.009
minimum	0.049	0.398
maximum	0.133	0.450

An alternative, more appropriate specification involves a regression of price on propensity to upwardly mis-report, $\Pr\left(I_t^u \mid \widetilde{z}_t = z_t\right) \equiv p_t^u\left(z_t\right)$, propensity to downwardly mis-report, $\Pr\left(I_t^d \mid \widetilde{z}_t = z_t\right) \equiv p_t^d\left(z_t\right)$, and propensity for partial upwardly mis-report, $\Pr\left(I_t^0 \mid \widetilde{z}_t = z_t\right) \equiv p_t^0\left(z_t\right)$, each multiplied by reported accruals z_t, upward and downward propensities to mis-report, and a control function for the truncated mean associated with partial upward reporting,

$$\begin{aligned} \lambda_t^0 &= p_t^0\left(z_t\right) 0.5\Delta \left[\alpha_t - \frac{\left(1-\alpha_t\right) \underline{R}_{t-1}^0}{\overline{R}_{t-1}^0 - \underline{R}_{t-1}^0}\right] \\ &\quad + \frac{0.5\Delta\left(1-\alpha_t\right)}{\overline{R}_{t-1}^0 - \underline{R}_{t-1}^0} \left(\mu_t^0 p_t^0\left(z_t\right) - \left(\sigma_t^0\right)^2 \left[\phi\left(\overline{w}_t^0\right) - \phi\left(\underline{w}_t^0\right)\right]\right) \end{aligned}$$

The regression then is

$$\begin{aligned} P_t &= \beta_1 p_t^u\left(z_t\right) + \beta_2 p_t^u\left(z_t\right) * z_t + \beta_3 p_t^d\left(z_t\right) + \beta_4 p_t^d\left(z_t\right) * z_t \\ &\quad + \beta_5 \lambda_t^0 + \beta_6 p_t^0\left(z_t\right) * z_t \end{aligned}$$

The *DGP* implies $\beta_1 = -0.5$, $\beta_2 = 1$, $\beta_3 = -5$, $\beta_4 = 1$, $\beta_5 = -0.5$, and $\beta_6 = 1$. The average perceived bias associated with partial upward mis-reporting can be recovered from the structural relations as the average of $\beta_1 p_t^u\left(z_t\right) + \beta_5 \lambda_t^0$ over the support of partial upward mis-reporting. We denote this quantity $\overline{\lambda}$. Simulation results are tabulated below.

statistic	β_1	β_2	β_3	β_4	β_5	β_6	$\overline{\lambda}$
mean	−0.500	1.000	5.000	1.000	−0.500	1.000	−0.284
median	−0.500	1.000	5.000	1.000	−0.500	1.000	−0.285
std. dev.	0.000	0.000	0.000	0.000	0.000	0.000	0.031
minimum	−0.500	1.000	5.000	1.000	−0.500	1.000	−0.367
maximum	−0.500	1.000	5.000	1.000	−0.500	1.000	−0.200

Now, assume ex post the analyst can identify the entrepreneur's reporting behavior (upward, downward, or partial upward mis-report). A regression utilizing indicators of the entrepreneur's actual report strategy (upward, downward, or partial upward mis-report) plus each multiplied by reported accruals

$$P_t = \beta_1 I_t^u + \beta_2 I_t^0 + \beta_3 I_t^d + \beta_4 I_t^u * z_t + \beta_5 I_t^0 * z_t + \beta_6 I_t^0 * z_t$$

produces less satisfactory results than those above. The posterior intervals do not contain the *DGP* parameter values for $\beta_1 = -0.5\Delta = -0.5$, or $\beta_4 = \beta_5 = 1$. Although random coefficients are challenging, the propensity scores combined with the control function above is likely to more effectively mitigate any bias than this alternative (naive) specification.[4]

statistic	β_1	β_2	β_3	β_4	β_5	β_6
mean	−0.450	−0.436	5.000	0.967	0.479	1.000
median	−0.451	−0.434	5.000	0.968	0.477	1.000
std. dev.	0.010	0.044	0.000	0.008	0.071	0.000
minimum	−0.474	−0.616	5.000	0.943	0.236	1.000
maximum	−0.419	−0.340	5.000	0.986	0.708	1.000

1 Propensity score estimation

We next explore ordered probit and nonparametric estimation of the propensity to upward, downward, or partially upward mis-report as estimated propensity score inputs to a linear regression.

1.1 Ordered parametric estimation of propensity scores

1.1.1 Accruals as the sole ordered probit propensity score regressor

Simulation results (based on 200 samples of size $N = 1,000$, $T = 10$) for the linear regression of price on observed accruals, ordered probit estimation of propensity to upward mis-report, $\Pr(I_t^u \mid \widetilde{z}_t = z_t) \equiv \widehat{p}_t^u(z_t)$, estimation of propensity for partial upward mis-reporting, $\Pr\left(I_t^0 \mid \widetilde{z}_t = z_t\right) \equiv \widehat{p}_t^0(z_t)$, both conditional on reported accruals, and the estimated control function, $\widehat{\lambda}_t^0$ is the same as λ_t^0 except $\widehat{p}_t^0(z_t)$ replaces $p_t^0(z_t)$.

$$\begin{aligned} P_t &= \beta_1\widehat{p}_t^u(z_t) + \beta_2\widehat{p}_t^u(z_t) * z_t + \beta_3\widehat{p}_t^d(z_t) + \beta_4\widehat{p}_t^d(z_t) * z_t \\ &\quad + \beta_5\widehat{\lambda}_t^0 + \beta_6\widehat{p}_t^0(z_t) * z_t \end{aligned}$$

are tabulated below. Again, the average bias in partial upward mis-reporting $\overline{\overline{\lambda}}$ equals the average of $\beta_1\widehat{p}_t^u(z_t) + \beta_5\widehat{\lambda}_t^0$ over the support for partial upward mis-reporting. Parameters β_3, β_4, β_5, and β_6 are not contained in their estimated intervals.

statistic	β_1	β_2	β_3	β_4	β_5	β_6	$\overline{\overline{\lambda}}$
mean	−0.499	0.996	7.669	1.344	−0.426	1.664	−0.294
median	−0.500	0.996	7.625	1.340	−0.422	1.652	−0.288
std. dev.	0.014	0.008	0.580	0.076	0.112	0.157	0.093
minimum	−0.540	0.972	6.480	1.191	−0.734	1.269	−0.615
maximum	−0.460	1.019	9.523	1.599	−0.121	2.146	−0.075

[4] This specification is "naive" in that it assumes investors have information in setting prices to which they are not privy.

1.1.2 Accruals as the sole ordered logit propensity score regressor

The above analysis is repeated with a logit link function replacing probit. Analogous quantities are employed, for example, ordered logit estimation of propensity to upward mis-report is $\Pr\left(I_t^u \mid \widetilde{z}_t = z_t\right) \equiv \widehat{q}_t^u\left(z_t\right)$, and estimation of propensity for partial upward mis-report is $\Pr\left(I_t^0 \mid \widetilde{z}_t = z_t\right) \equiv \widehat{q}_t^0\left(z_t\right)$ both conditional on reported accruals

$$
\begin{aligned}
P_t &= \beta_1 \widehat{q}_t^u\left(z_t\right) + \beta_2 \widehat{q}_t^u\left(z_t\right) * z_t + \beta_3 \widehat{q}_t^d\left(z_t\right) + \beta_4 \widehat{q}_t^d\left(z_t\right) * z_t \\
&\quad + \beta_5 \widehat{\lambda}_t^0\left(z_t\right) + \beta_6 \widehat{q}_t^0\left(z_t\right) * z_t
\end{aligned}
$$

Simulation results for the linear regression of price on the regressors are tabulated below. Results are similar to those based on ordered probit propensity scores though noisier; while β_3, β_4, β_5, and β_6 are not contained in their estimated intervals.

statistic	β_1	β_2	β_3	β_4	β_5	β_6	$\overline{\overline{\lambda}}$
mean	-0.505	0.996	7.386	1.305	-0.438	1.555	-0.302
median	-0.505	0.996	7.367	1.303	-0.435	1.542	-0.294
std. dev.	0.014	0.007	0.482	0.063	0.100	0.116	0.087
minimum	-0.548	0.975	6.340	1.171	-0.722	1.255	-0.612
maximum	-0.469	1.014	8.866	1.506	-0.174	1.860	-0.110

1.1.3 Accruals and indicators of perceived mis-reporting as propensity score regressors

Ordered probit propensity score estimation. The above inadequacies suggest mis-specification in the propensity score estimation. In response, we add regressors to the propensity score estimation. In particular, an indicator for perceived (by investors) to be upwardly mis-reported PI_t^u and an indicator for perceived to be either upwardly or partially upward mis-reported PI_t^{u0} are added as potential omitted, correlated variables. In other words, investors' perception of support for reserves determines the support for mis-reporting strategies as reflected in PI_t^u and PI_t^{u0}. Simulation results for the linear regression of price on observed accruals, ordered probit estimation of propensity to upward mis-report,

$$
\Pr\left(I_t^u \mid \widetilde{z}_t = z_t, PI_t^u, PI_t^{u0}\right) \equiv \widetilde{p}_t^u\left(z_t, PI_t^u, PI_t^{u0}\right)
$$

and estimation of propensity for partial upward mis-reporting,

$$
\Pr\left(I_t^0 \mid \widetilde{z}_t = z_t, P_t^u, P_t^{u0}\right) \equiv \widetilde{p}_t^0\left(z_t, PI_t^u, PI_t^{u0}\right)
$$

both conditional on reported accruals and perceived mis-reporting

$$
\begin{aligned}
P_t &= \beta_1 \widetilde{p}_t^u\left(z_t, PI_t^u, PI_t^{u0}\right) + \beta_2 \widetilde{p}_t^u\left(z_t, PI_t^u, PI_t^{u0}\right) * z_t \\
&\quad + \beta_3 \widetilde{p}_t^d\left(z_t, PI_t^u, PI_t^{u0}\right) + \beta_4 \widetilde{p}_t^d\left(z_t, PI_t^u, PI_t^{u0}\right) * z_t \\
&\quad + \beta_5 \widetilde{\lambda}_t^0 + \beta_6 \widetilde{p}_t^0\left(z_t, PI_t^u, PI_t^{u0}\right) * z_t + \beta_7 \lambda_t^0
\end{aligned}
$$

are tabulated below. The results suggest the propensity score estimates and control function are more effective in estimating upward and downward mis-reporting effects as all parameters are contained in their estimated intervals.

statistic	β_1	β_2	β_3	β_4	β_5	β_6	$\overline{\overline{\lambda}}$
mean	−0.507	1.003	5.000	1.000	−0.363	0.961	−0.163
median	−0.507	1.003	5.000	1.000	−0.359	0.964	−0.158
std. dev.	0.003	0.002	0.000	0.000	0.131	0.225	0.063
minimum	−0.515	0.996	5.000	1.000	−0.813	0.351	−0.373
maximum	−0.496	1.008	5.003	1.000	−0.021	1.689	−0.010

Ordered logit propensity score estimation. Simulation results for the linear regression of price on observed accruals, ordered logit estimation of propensity to upward mis-report,

$$\Pr\left(I_t^u \mid \widetilde{z}_t = z_t, PI_t^u, PI_t^{u0}\right) \equiv \widetilde{q}_t^u\left(z_t, PI_t^u, PI_t^{u0}\right)$$

and estimation of propensity for no mis-report,

$$\Pr\left(I_t^0 \mid \widetilde{z}_t = z_t, P_t^u, P_t^{u0}\right) \equiv \widetilde{q}_t^0\left(z_t, PI_t^u, PI_t^{u0}\right)$$

both conditional on reported accruals and perceived mis-reporting

$$\begin{aligned} P_t &= \beta_1 \widetilde{q}_t^u\left(z_t, PI_t^u, PI_t^{u0}\right) + \beta_2 \widetilde{q}_t^u\left(z_t, PI_t^u, PI_t^{u0}\right) * z_t \\ &\quad + \beta_3 \widetilde{q}_t^d\left(z_t, PI_t^u, PI_t^{u0}\right) + \beta_4 \widetilde{q}_t^d\left(z_t, PI_t^u, PI_t^{u0}\right) * z_t \\ &\quad + \beta_5 \widetilde{\overline{\lambda}}_t^0 + \beta_6 \widetilde{q}_t^0\left(z_t, PI_t^u, PI_t^{u0}\right) * z_t \end{aligned}$$

are tabulated below. The logit results are very similar to the probit results above. The additional regressors employed in the estimation of the propensity scores reduce the noise in the estimates in the price regression.

statistic	β_1	β_2	β_3	β_4	β_5	β_6	β_7
mean	−0.507	1.003	5.001	1.000	−0.350	0.944	−0.156
median	−0.507	1.003	5.000	1.000	−0.344	0.939	−0.149
std. dev.	0.003	0.002	0.001	0.000	0.133	0.227	0.063
minimum	−0.515	0.996	5.000	1.000	−0.808	0.331	−0.373
maximum	−0.496	1.009	5.004	1.000	−0.001	1.685	−0.000

2 Summary

This example involves a random coefficients model. In this sense it is similar to Heckman and Vytlacil's [2005, 2007] marginal treatment effects (*MTE*) model. Our approach to this stage has been to address the determinants of the random parameters as *observable* (in the form of λ_t^0) while Heckman and Vytlacil focus on unobservable heterogeneity. As with *MTE*, unobservable heterogeneity in the present setting is challenging as conditional mean independence is not likely to be satisfied by the data. We address *unobservable heterogeneity* in a later example.

Key features of this example are once again the role of propensity scores and the importance of its specification but here complemented with a control function to account for the expected values of truncated random variables. Although the data are normally distributed, this setting does not lend itself to a recipe-approach say based on Heckman's inverse-Mills ratio. Even though the interior region of partial upward mis-reporting involves elements similar to Heckman's standard model, blind application is inadequate. As Heckman [2001] emphasizes no algorithm for econometric analysis exists nor is one likely to be identified in the future. Rather, we must heed the theory and data in hand.

2.1 Nonparametric propensity score estimation

To re-evaluate the above parametric analysis, we now employ nonparametric regression to estimate the propensity to upward or partially upward mis-report. Nonparametric estimation of multinomial data (with K levels) proceeds as with binomial data except that $K-1$ regressions are estimated each employing its own observed outcome I_t^k ($k \in \{u, d, 0\}$) and all regressions using the same bandwidth.[5] Simulation results (based on 200 samples of size $N = 1,000$, $T = 10$) for the linear regression of price on observed accruals, nonparametric estimation of propensity to upward mis-report, $\Pr(I_t^u \mid \widetilde{z}_t = z_t) \equiv m_t^u(z_t)$, nonparametric estimation of propensity for partial upward mis-reporting, $\Pr\left(I_t^0 \mid \widetilde{z}_t = z_t\right) \equiv m_t^0(z_t)$ both conditional on reported accruals, and the control function, λ_t^0,

$$\begin{aligned} P_t &= \beta_1 m_t^u(z_t) + \beta_2 m_t^u(z_t) * z_t + \beta_3 m_t^d(z_t) + \beta_4 m_t^d(z_t) * z_t \\ &\quad + \beta_5 m_t^0(z_t) + \beta_6 m_t^0(z_t) * z_t + \beta_7 \lambda_t^0 \end{aligned}$$

are tabulated below. Nonparametric propensity scores based on accruals alone produce similar estimated intervals to the ordered parametric models. As expected, the estimates are somewhat noisy but perhaps less noisy than those based on parametric propensity scores employing accruals alone. Again, estimates of the random coefficients are quite noisy and $\beta_7 = 1$ is not contained in its estimated interval.

statistic	β_1	β_2	β_3	β_4	β_5	β_6	β_7
mean	-0.396	0.951	4.981	0.997	-1.635	-0.104	-0.926
median	-0.400	0.952	5.000	1.000	-1.630	-0.112	-0.908
std. dev.	0.055	0.027	0.012	0.193	0.668	0.320	0.049
minimum	-0.523	0.880	4.143	0.877	-3.546	-1.240	-1.500
maximum	-0.264	1.008	5.080	1.011	0.642	1.050	0.202

Before leaving this example, we revisit the nonparametric propensity scores by adding perceived mis-reporting to the regression. Simulation results for the linear regression of price on observed accruals, nonparametric estimation of propensity to upward mis-report,

$$\Pr\left(I_t^u \mid \widetilde{z}_t = z_t, PI_t^u, PI_t^{u0}\right) \equiv \widetilde{m}_t^u\left(z_t, PI_t^u, PI_t^{u0}\right)$$

and estimation of propensity for partial upward mis-reporting,

$$\Pr\left(I_t^0 \mid \widetilde{z}_t = z_t, P_t^u, P_t^{u0}\right) \equiv \widetilde{m}_t^0\left(z_t, PI_t^u, PI_t^{u0}\right)$$

both conditional on reported accruals and perceived mis-reporting

$$\begin{aligned} P_t &= \beta_1 \widetilde{m}_t^u\left(z_t, PI_t^u, PI_t^{u0}\right) + \beta_2 \widetilde{m}_t^u\left(z_t, PI_t^u, PI_t^{u0}\right) * z_t \\ &\quad + \beta_3 \widetilde{m}_t^d\left(z_t, PI_t^u, PI_t^{u0}\right) + \beta_4 \widetilde{m}_t^d\left(z_t, PI_t^u, PI_t^{u0}\right) * z_t \\ &\quad + \beta_5 \widetilde{m}_t^0\left(z_t, PI_t^u, PI_t^{u0}\right) + \beta_6 \widetilde{m}_t^0\left(z_t, PI_t^u, PI_t^{u0}\right) * z_t + \beta_7 \lambda_t^0 \end{aligned}$$

[5] Nonparametric regression estimates of the propensity score utilize a bandwidth of $h = 0.2$ throughout.

are tabulated below. The nonparametric propensity score estimates based on accruals and perceived mis-reporting produce similar results to the parametric discrete choice models. $\beta_1 = -0.5$, $\beta_1 = 1$, $\beta_2 = 5$, and $\beta_3 = 1$ are contained in their estimated intervals. The random coefficients are contained in their estimated intervals as well, though they are again quite noisy. Collectively, the results support the inclusion of perceived mis-reporting as regressors in the propensity score estimation, the importance of propensity score specification, and the effectiveness of including propensity scores and the control function as otherwise omitted, correlated variables in the regression of price on (endogenous) reported accruals.

statistic	β_1	β_2	β_3	β_4	β_5	β_6	β_7
mean	-0.485	0.991	4.981	0.997	-0.314	0.215	0.032
median	-0.487	0.992	5.000	1.000	-0.295	0.234	0.078
std. dev.	0.011	0.008	0.117	0.017	0.168	0.182	0.261
minimum	-0.501	0.963	3.958	0.850	-0.840	-0.237	-0.467
maximum	-0.436	1.003	5.053	1.007	0.332	0.815	1.087

3 Summary

Key features of this example are once again the role of propensity scores and the importance of its specification but now complemented with a control function to account for the expected values of truncated random variables. Although the data are normally distributed, this setting does not lend itself to a recipe-approach say based on Heckman's inverse-Mills ratio. Even though the interior region of partial upward mis-reporting involves elements similar to Heckman's standard model, blind application is nonsensical. As Heckman [2001] indicates no algorithm for econometric analysis exists nor is one likely to be identified in the future. Rather, we must heed the theory and data in hand.