1 Accounting revaluation regulations[1]

Our second example explores the ex ante impact of accounting asset revaluation policies on owners' investment decisions (and welfare) in an economy of, on average, price protected buyers. Prior to investment, an owner evaluates both investment prospects and the market for resale in the event the owner becomes liquidity stressed. The payoff from investment I is distributed uniformly and centered at $\hat{x} = \frac{\beta}{\alpha} I^{\alpha}$ where $\alpha, \beta > 0$ and $\alpha < 1$. Hence, support for investment payoff is $x : \hat{x} \pm f = [\underline{x}, \overline{x}]$. A potential problem with the resale market is the owner will have private information - knowledge of the asset value. However, since there is some positive probability the owner becomes distressed π (as in Dye [1985]) the market will not collapse. The equilibrium price is based on distressed sellers marketing potentially healthy assets combined with non-distressed sellers opportunistically marketing impaired assets. Regulators may choose to prop-up the price to aid distressed sellers by requiring certification of assets at cost k [2] with values below some cutoff x_c.[3] The owner's ex ante expected payoff from investment I and certification cutoff x_c is

$$\begin{aligned} E[V \mid I, x_c] &= \pi \frac{1}{2f}\left[\frac{1}{2}\left(x_c^2 - \underline{x}^2\right) - k\left(x_c - \underline{x}\right) + P\left(\overline{x} - x_c\right)\right] \\ &\quad + (1-\pi)\frac{1}{2f}\left[\frac{1}{2}\left(x_c^2 - \underline{x}^2\right) + P\left(P - x_c\right) + \frac{1}{2}\left(\overline{x}^2 - P^2\right)\right] - I \end{aligned}$$

The equilibrium uncertified asset price is

$$P = \frac{x_c + \sqrt{\pi}\,\overline{x}}{1 + \sqrt{\pi}}$$

This follows from the equilibrium condition

$$P = \frac{1}{4fq}\left[\pi\left(\overline{x}^2 - x_c^2\right) + (1-\pi)\left(P^2 - x_c^2\right)\right]$$

where

$$q = \frac{1}{2f}\left[\pi\left(\overline{x} - x_c\right) + (1-\pi)\left(P - x_c\right)\right]$$

is the probability that an uncertified asset is marketed. Further, the regulator may differentially weight the welfare $W(I, x_c)$ of distressed sellers compared with non-distressed sellers. Specifically, the regulator may value distressed seller's net gains dollar-for-dollar but value non-distressed seller's gains at a

[1] This example draws heavily from Demski, Lin, and Sappington [2008].

[2] This cost is incremental to normal audit cost. As such, even if audit fee data is available, k may be difficult for the analyst to observe.

[3] Owners never find it ex ante beneficial to commit to any certified revaluation because of the certification cost. We restrict attention to targeted certification but certification could be proportional rather than targeted (see Demski, et al [2008] for details). For simplicity, we explore only targeted certification.

fraction w on the dollar.

$$\begin{aligned} W(I, x_c) &= \pi \frac{1}{2f}\left[\frac{1}{2}\left(x_c^2 - \underline{x}^2\right) - k\left(x_c - \underline{x}\right) + P\left(\overline{x} - x_c\right)\right] \\ &\quad + w(1-\pi)\frac{1}{2f}\left[\frac{1}{2}\left(x_c^2 - \underline{x}^2\right) + P\left(P - x_c\right) + \frac{1}{2}\left(\overline{x}^2 - P^2\right)\right] \\ &\quad - I\left[\pi + (1-\pi) w\right] \end{aligned}$$

1.1 Numerical example

Consider the following parameters

$$\left\{\alpha = \frac{1}{2}, \beta = 10, \pi = 0.7, k = 2, f = 100\right\}$$

Then never certify ($x_c = \underline{x}$) results in investment $I = 100$, owner's expected payoff $E[V \mid I, x_c] = 100$, and equilibrium uncertified asset price $P \approx 191.1$. Owners will choose to never certify asset values. However, regulators may favor distressed sellers and require selective certification. Continuing with the same parameters, if regulators give zero consideration ($w = 0$) to the expected payoffs of non-distressed sellers, then the welfare maximizing certification cutoff $x_c = \overline{x} - \frac{(1+\sqrt{\pi})k}{(1-\sqrt{\pi})(1-w)} \approx 278.9$. This induces investment $I = \left[\frac{\beta(2f+\pi k)}{2f}\right]^{\frac{1}{1-\alpha}} \approx 101.4$, owner's expected payoff approximately equal to 98.8, and equilibrium uncertified asset price $P \approx 289.2$ (an uncertified price more favorable to distressed sellers). For various regulations and certification costs, some relevant investment choices and expected payoffs are summarized in the table below.

	$x_c = \underline{x}, k = 2$	$x_c = \underline{x}, k = 20$	$x_c = 200, k = 2$	$x_c = 200, k = 20$
π	0.7	0.7	0.7	0.7
I	100	100	101.4	114.5
P	191.1	191.1	246.2	251.9
$E[x-k]$	200	200	200.7	208
$E[V]$	100	100	99.3	93.5

Investment choice and payoffs for no certification and selective certification

	$x_c = \overline{x}, k = 2$	$x_c = \overline{x}, k = 20$	$x_c = \overline{x}, k = 2$	$x_c = \overline{x}, k = 20$
π	0.7	0.7	1.0	1.0
I	100	100	100	100
P	NA	NA	NA	NA
$E[x-k]$	198.6	186	198	180
$E[V]$	98.6	86	98	80

Investment choice and payoffs for full certification

2 Full certification

The base case involves full certification $x_c = \overline{x}$ and all owners market their assets, $\pi = 1$. This setting ensures outcome data availability (excluding investment cost) which may be an issue when we relax these conditions. There are two firm types: one with low mean certification costs $\widehat{k}^L = 2$ and the other with high mean certification costs $\widehat{k}^H = 20$. Suppose owners choose their investment levels anticipating selective certification with $x_c = 200$ and $\pi = 0.7$.[4] Optimal investment levels for a selective certification environment are $I^L = 101.4$ and $I^H = 114.5$, and expected asset values including certification costs are $E\left[x^L - k^L\right] = 199.4$ and $E\left[x^H - k^H\right] = 194$. Treatment (investment level) is chosen based on ex ante beliefs of selective certification. Treatment is binary, the analyst observes high or low investment but not the investment level,[5] and defined as $D = 1$ when $I^L = 101.4$ while non-treatment is defined as $D = 0$ when $I^H = 114.5$. Outcome is ex post value in an always certify environment $Y_j = x^j - k^j$. That is, the difference in realized value due to (ex ante) equilibrium investment choice is the treatment effect of interest.

The average treatment effect on the treated is

$$\begin{aligned} ATT &= E\left[Y_1 - Y_0 \mid D = 1\right] \\ &= E\left[x^L - k^L \mid D = 1\right] - E\left[x^H - k^L \mid D = 1\right] \\ &= E\left[x^L - x^H \mid D = 1\right] \\ &= 201.4 - 214 = -12.6 \end{aligned}$$

The average treatment effect on the untreated is

$$\begin{aligned} ATUT &= E\left[Y_1 - Y_0 \mid D = 0\right] \\ &= E\left[x^L - k^H \mid D = 0\right] - E\left[x^H - k^H \mid D = 0\right] \\ &= E\left[x^L - x^H \mid D = 0\right] \\ &= 201.4 - 214 = -12.6 \end{aligned}$$

Hence, outcome is homogeneous, $ATE = ATT = ATUT = -12.6$.[6] With no covariates and outcome not mean independent of treatment, OLS estimates

$$\begin{aligned} OLS &= E\left[Y_1 \mid D = 1\right] - E\left[Y_0 \mid D = 0\right] \\ &= E\left[x^L - k^L \mid D = 1\right] - E\left[x^H - k^H \mid D = 0\right] \\ &= 5.4 \end{aligned}$$

via β_1 in the following regression.[7]

$$E\left[Y \mid D\right] = \beta_0 + \beta_1 D$$

[4] As seen in the table, for full certification there is no variation in equilibrium investment level and our experiment is uninteresting.

[5] If the analyst observes the investment level, then outcome includes investment cost and we work with a more complete measure of the owner's welfare..

[6] If k is unobservable, then outcome Y may be measured by x only and treatment effects represent gross rather than net gains. Further, the treatment effect is the difference between equilibrium and off-equilibrium responses without adjusting the outcome measure. This is an apparent advantage relative to outcome $x - k$ as we must adjust $x - k$ outcomes to account for counterfactual cases $\left(I^L, k^H\right)$ or $\left(I^H, k^L\right)$. The drawback is that outcome x is an even more incomplete measure of welfare than $x - k$. Recall, the owner's welfare is an ex ante measure, $E\left[V\right]$, that depends on x, k, and I. In any case, we must exercise care in interpreting the treatment effect results because of limitations in our outcome measure.

[7] Notice the difference in the treatment effects and what is estimated via OLS is $k^L - k^H = -18 = -12.6 - 5.4$.

where $Y = D\left(x^L - k^L\right) + (1-D)\left(x^H - k^H\right)$ (ex post payoff), and

$$D = \begin{matrix} 1 & I^L = 101.4 \\ 0 & I^H = 114.5 \end{matrix}$$

We simulate 200 samples of 2,000 draws where

$$x^j \sim uniform\left(\widehat{x}^j - 100, \widehat{x}^j + 100\right)$$

$$k^j \sim uniform\left(\widehat{k}^j - 1, \widehat{k}^j + 1\right)$$

and

$$L - type \sim Bernoulli\,(0.5)$$

Simulation results for the above *OLS* model and sample treatment effect statistics are tabulated below.

statistics	β_0	$\beta_1\,(estATE)$	
mean	193.8	5.797	
median	193.7	5.805	
stand.dev.	1.831	2.684	
minimum	188.2	−1.778	
maximum	198.9	13.32	
OLS parameter estimates for $E[Y] = \beta_0 + \beta_1 D$ full certification setting			
statistics	*ATE*	*ATT*	*ATUT*
mean	−12.54	−12.49	−12.59
median	−12.55	−12.44	−12.68
stand.dev.	1.947	2.579	2.794
minimum	−17.62	−19.53	−21.53
maximum	−7.718	−6.014	−6.083
Average treatment effect sample statistics			

OLS clearly produces biased estimates of the treatment effect in this simple base case. This can be explained as low or high certification cost type is a perfect predictor of treatment. That is, $\Pr\left(D = 1 \mid k^L\right) = 1$ and $\Pr\left(D = 1 \mid k^H\right) = 0$. Therefore, the common support condition for identifying counterfactuals fails and standard approaches (ignorable treatment or even instrumental variables) don't identify treatment effects.[8]

[8] An alternative analysis tests the common support condition. Suppose everything remains as above except $k^H \sim uniform\,(1, 19)$ and sometimes the owners perceive certification cost to be low when it is high, hence $\Pr(D = 1 \mid \text{type} = H) = 0.1$. This setup implies observed outcome is

$$Y = D\left[(Y_1 \mid \text{type} = L) + (Y_1 \mid \text{type} = H)\right] + (1 - D)(Y_0 \mid \text{type} = H)$$

such that

$$E[Y] = 0.5E\left[x^L - k^L\right] + 0.5\left\{0.1E\left[x^L - k^H\right] + 0.9E\left[x^H - k^H\right]\right\}$$

However, from the above we can manipulate the outcome variable to identify the treatment effects via *OLS*. Observed outcome is

$$\begin{aligned} Y &= D\left(x^L - k^L\right) + (1-D)\left(x^H - k^H\right) \\ &= \left(x^H - k^H\right) + D\left(x^L - x^H\right) - D\left(k^L - k^H\right) \end{aligned}$$

The first term is captured via the regression intercept and the second term is the treatment effect. Therefore, if we add the last term $DE\left[k^L - k^H\right]$ to Y we can identify the treatment effect from the coefficient on D. If the analyst observes $k = Dk^L + (1-D)k^H$, then we can utilize a two-stage regression approach. The first stage is

$$E[k \mid D] = \alpha_0 + \alpha_1 D$$

where $\alpha_0 = E\left[k^H\right]$ and $\alpha_1 = E\left[k^L - k^H\right]$. Now, the second stage regression employs

$$\begin{aligned} Y' &= Y + D\widehat{\alpha}_1 \\ &= Y + D\left(\overline{k}^L - \overline{k}^H\right) \end{aligned}$$

and estimate the treatment effect via the analogous regression to the above[9]

$$E[Y' \mid D] = \beta_0 + \beta_1 D$$

Suppose the analyst ex post observes the actual certification cost type and let $T = 1$ if type $= L$. The common support condition is satisfied and the outcome mean is conditionally independent of treatment given T implies treatment is ignorable. *OLS* simulation results are tabulated below.

statistics	β_0	β_1	β_2 (*estATE*)
mean	196.9	7.667	−5.141
median	196.9	7.896	−5.223
stand.dev.	1.812	6.516	6.630
minimum	191.5	−10.62	−23.54
maximum	201.6	25.56	14.25

OLS parameter estimates for $E[Y \mid D] = \beta_0 + \beta_1 T + \beta_2 D$ full certification setting

statistics	*ATE*	*ATT*	*ATUT*
mean	−5.637	−5.522	−5.782
median	−5.792	−5.469	−5.832
stand.dev.	1.947	2.361	2.770
minimum	−9.930	−12.05	−12.12
maximum	0.118	0.182	0.983

Average treatment effect sample statistics

The estimate average treatment effect is slightly attenuated and has high variability that may compromise its finite sample utility. Nevertheless, the results are a dramatic departure and improvement from the results above where the common support condtion fails.

[9] This is similar to a regression discontinuity design (for example, see Angrist and Lavy [1999] and Angrist and Pischke [2009]). However, the jump in cost of certification k^j violates the regression continuity in X condition (assuming $k = Dk^L + (1-D)\,k^H$ is observed and included in X). If the support of k^L and k^H is adjacent, then the regression discontinuity design

$$E[Y \mid X, D] = \beta_0 + \beta_1 k + \beta_2 D$$

effectively identifies the treatment effects but fails with the current *DGP*. Typical results for

Simulation results for the revised *OLS* model and sample treatment effect statistics are tabulated below. As simulation allows us to observe both the factual data and counterfactual data in the experiment, the sample statistics describe the simulated draws in terms of the "observed" average treatment effects.

statistics	β_0	$\beta_1\,(estATE)$	
mean	193.8	-12.21	
median	193.7	-12.21	
stand.dev.	1.831	2.687	
minimum	188.2	-19.74	
maximum	198.9	-64.691	
OLS parameter estimates for $E[Y' \mid D] = \beta_0 + \beta_1 D$ full certification setting			
statistics	*ATE*	*ATT*	*ATUT*
mean	-12.54	-12.49	-12.59
median	-12.55	-12.44	-12.68
stand.dev.	1.947	2.579	2.794
minimum	-17.62	-19.53	-21.53
maximum	-7.718	-6.014	-6.083
Average treatment effect sample statistics			

With adjusted outcomes, *OLS* estimates correspond quite well with the treatment effects. Next, we explore *2SLS-IV* with a binary instrument to estimate *LATE*.

2.0.1 Binary instrument

Suppose we're interested in identifying the marginal treatment effect, *LATE*, from *IV* exclusion restrictions and have access to a binary instrument z.[10] From

the current *DGP* are

statistics	β_0	β_1	$\beta_2\,(estATE)$	*ATE*
mean	218.4	-1.210	-16.59	-12.54
median	213.2	-0.952	-12.88	-12.55
stand.dev.	42.35	2.119	38.26	1.947
minimum	122.9	-6.603	-115.4	-17.62
maximum	325.9	3.573	71.56	-7.718
OLS parameter estimates for $E[Y \mid D] = \beta_0 + \beta_1 k + \beta_2 D$ full certification setting				

The coefficient on D represents a biased and erratic estimate of the average treatment effect. Given the variability of the estimates, a regression discontinuity design has limited small sample utility for this *DGP*. However, we later return to regression discontinuity designs when modified *DGP*s are considered. For the current *DGP*, we employ the approach discussed above, which is essentially restricted least squares.

[10] For simulation purposes, the instrument is a noisy indicator of cost of certification k. That is, $z = \Im(k^L)\, z_1 + (1 - \Im(k^L))\, z_0$ where $\Im(k^L)$ is a binary indicator equal to one when $E[k] = 2$ and zero otherwise, $z_1 \sim Bernoulli\,(0.99)$ and $z_0 \sim Bernoulli\,(0.01)$.

Imbens and Angrist, β_1 identifies *LATE* for the subpopulation of compliers when estimated via *2SLS-IV*. Simulation results are tabulated below.

statistics	β_0	β_1 (*estLATE*)
mean	193.8	−12.21
median	193.7	−12.14
stand.dev.	1.845	2.698
minimum	187.6	−19.56
maximum	198.9	−3.985

2SLS-IV parameter estimates for
$E[Y' \mid D] = \beta_0 + \beta_1 D$
full certification setting

statistics	*LATE*	*ATE*	*ATT*	*ATUT*
mean	−12.51	−12.54	−12.49	−12.59
median	−12.41	−12.55	−12.44	−12.68
stand.dev.	2.599	1.947	2.579	2.794
minimum	−19.58	−17.62	−19.53	−21.53
maximum	−5.969	−7.718	−6.014	−6.083

Average treatment effect sample statistics

2SLS-IV effectively estimates the marginal treatment effect, *LATE*. Further, the evidence supports homogeneity, in other words, all treatments effects are the same, $ATE = ATT = ATUT = LATE$. Minor differences in the sample statistics stem simply from differences in support.

2.0.2 Propensity score

Since the data are conditionally mean independent (i.e., satisfy ignorability of treatment), the treatment effects can be estimated via the propensity score as discussed earlier in chapter 9. Propensity score is the estimated probability of treatment conditional on the regressors $m = \Pr(D = 1 \mid X)$. For simulation purposes, $X = z$ plus noise where the noise is standard normal random draws.

Estimators of some treatment effects via propensity score are

$$estATE = n^{-1} \sum_{j=1}^{n} \frac{(D_j - m_j) Y_j'}{m_j (1 - m_j)}$$

$$estATT = \frac{n^{-1} \sum_{j=1}^{n} \frac{(D_j - m_j) Y_j'}{(1 - m_j)}}{n^{-1} \sum_{j=1}^{n} D_j}$$

$$estATUT = \frac{n^{-1} \sum_{j=1}^{n} \frac{(D_j - m_j) Y_j'}{m_j}}{n^{-1} \sum_{j=1}^{n} (1 - D_j)}$$

Propensity score estimates and comparative sample statistics for the treatment effects are tabulated below.

statistics	*estATE*	*estATT*	*estATUT*
mean	−12.42	−13.96	−10.87
median	−12.50	−13.60	−11.40
stand.dev.	5.287	6.399	5.832
minimum	−31.83	−45.83	−25.61
maximum	−1.721	0.209	10.56
Propensity score treatment effect estimates full certification setting			
statistics	*ATE*	*ATT*	*ATUT*
mean	−12.54	−12.49	−12.59
median	−12.55	−12.44	−12.68
stand.dev.	1.947	2.579	2.794
minimum	−17.62	−19.53	−21.53
maximum	−7.718	−6.014	−6.083
Average treatment effect sample statistics			

The estimates are somewhat more variable than we would like but they are consistent with the sample statistics on average. Further, we cannot reject homogeneity even though the treatment effect means are not as similar as we might expect.

2.0.3 Propensity score matching

A simple and intuitively appealing variation on the above propensity score approach involves matching treated and untreated on propensity score then compute the average treatment effect based on the matched-pair differences. We follow Sekhon [2008] by employing the "Matching" library for **R.**[11] We find optimal matches of treated with untreated using replacement sampling. Simulation results for propensity score matching average treatment effects are tabulated

[11] We don't go into details regarding matching since we employ only one regressor in the propensity score model. Matching is a rich study in itself. For instance, Sekhon [2008] discusses a genetic matching algorithm. Heckman, Ichimura, and Todd [1998] discuss nonparametric kernel matching.

below.[12]

statistics	estATE	estATT	estATUT
mean	-12.46	-12.36	-12.56
median	-12.54	-12.34	-12.36
stand.dev.	3.530	4.256	4.138
minimum	-23.49	-24.18	-22.81
maximum	-3.409	-2.552	-0.659
Propensity score matching average treatment effect estimates full certification setting			
statistics	*ATE*	*ATT*	*ATUT*
mean	-12.54	-12.49	-12.59
median	-12.55	-12.44	-12.68
stand.dev.	1.947	2.579	2.794
minimum	-17.62	-19.53	-21.53
maximum	-7.718	-6.014	-6.083
Average treatment effect sample statistics			

The matched propensity score results correspond well with the sample statistics. In this setting, the matched propensity score estimates of causal effects are less variable than the previous propensity score results. Further, they are more uniform across (homogeneous) treatment effects.

2.1 Selective certification

Analysis of the ex post value net of certification cost treatment effect ($\pi = 0.7$ and $x_c = 200$ are known to the owners ex ante) is somewhat more challenging than the base case as there are multiple levels of mean effects for each investment level (or treatment). The average treatment effect on the treated, ATT, is

$$\begin{aligned}
&\pi \frac{1}{2f}\left[\frac{1}{2}\left(x_c^2 - \left(\underline{x}^L\right)^2\right) - k^L\left(x_c - \underline{x}^L\right) + P^L\left(\overline{x}^L - x_c\right)\right] \\
&+ (1-\pi)\frac{1}{2f}\left[\frac{1}{2}\left(x_c^2 - \left(\underline{x}^L\right)^2\right) + P^L\left(P^L - x_c\right) + \frac{1}{2}\left(\left(\overline{x}^L\right)^2 - \left(P^L\right)^2\right)\right] \\
&- \pi \frac{1}{2f}\left[\frac{1}{2}\left(x_c^2 - \left(\underline{x}^H\right)^2\right) - k^L\left(x_c - \underline{x}^H\right) + P^H\left(\overline{x}^H - x_c\right)\right] \\
&- (1-\pi)\frac{1}{2f}\left[\frac{1}{2}\left(x_c^2 - \left(\underline{x}^H\right)^2\right) + P^H\left(P^H - x_c\right) + \frac{1}{2}\left(\left(\overline{x}^H\right)^2 - \left(P^H\right)^2\right)\right]
\end{aligned}$$

[12] ATE, ATT, and $ATUT$ may be different because their regions of common support may differ. For example, ATT requires common support only in the region $D = 1$ and $ATUT$ requires common support only in the region $D = 0$.

The average treatment effect on the untreated, $ATUT$, is

$$\begin{aligned}
&\pi \frac{1}{2f}\left[\frac{1}{2}\left(x_c^2-\left(\underline{x}^L\right)^2\right)-k^H\left(x_c-\underline{x}^L\right)+P^L\left(\overline{x}^L-x_c\right)\right] \\
&+(1-\pi)\frac{1}{2f}\left[\frac{1}{2}\left(x_c^2-\left(\underline{x}^L\right)^2\right)+P^L\left(P^L-x_c\right)+\frac{1}{2}\left(\left(\overline{x}^L\right)^2-\left(P^L\right)^2\right)\right] \\
&-\pi \frac{1}{2f}\left[\frac{1}{2}\left(x_c^2-\left(\underline{x}^H\right)^2\right)-k^H\left(x_c-\underline{x}^H\right)+P^H\left(\overline{x}^H-x_c\right)\right] \\
&-(1-\pi)\frac{1}{2f}\left[\frac{1}{2}\left(x_c^2-\left(\underline{x}^H\right)^2\right)+P^H\left(P^H-x_c\right)+\frac{1}{2}\left(\left(\overline{x}^H\right)^2-\left(P^H\right)^2\right)\right]
\end{aligned}$$

But the OLS estimand is

$$\begin{aligned}
&\pi \frac{1}{2f}\left[\frac{1}{2}\left(x_c^2-\left(\underline{x}^L\right)^2\right)-k^L\left(x_c-\underline{x}^L\right)+P^L\left(\overline{x}^L-x_c\right)\right] \\
&+(1-\pi)\frac{1}{2f}\left[\frac{1}{2}\left(x_c^2-\left(\underline{x}^L\right)^2\right)+P^L\left(P^L-x_c\right)+\frac{1}{2}\left(\left(\overline{x}^L\right)^2-\left(P^L\right)^2\right)\right] \\
&-\pi \frac{1}{2f}\left[\frac{1}{2}\left(x_c^2-\left(\underline{x}^H\right)^2\right)-k^H\left(x_c-\underline{x}^H\right)+P^H\left(\overline{x}^H-x_c\right)\right] \\
&-(1-\pi)\frac{1}{2f}\left[\frac{1}{2}\left(x_c^2-\left(\underline{x}^H\right)^2\right)+P^H\left(P^H-x_c\right)+\frac{1}{2}\left(\left(\overline{x}^H\right)^2-\left(P^H\right)^2\right)\right]
\end{aligned}$$

As in the full certification setting, the key differences revolve around the costly certification terms. The costly certification term for ATT is

$$\begin{aligned}
&-\pi \frac{1}{2f}\left[k^L\left(x_c-\underline{x}^L\right)-k^L\left(x_c-\underline{x}^H\right)\right] \\
=\;&-\pi \frac{\left(\underline{x}^H-\underline{x}^L\right)}{2f}k^L
\end{aligned}$$

and for $ATUT$ is

$$\begin{aligned}
&-\pi \frac{1}{2f}\left[k^H\left(x_c-\underline{x}^L\right)-k^H\left(x_c-\underline{x}^H\right)\right] \\
=\;&-\pi \frac{\left(\underline{x}^H-\underline{x}^L\right)}{2f}k^H
\end{aligned}$$

While the costly certification term in the estimand for OLS is

$$-\pi \frac{1}{2f}\left[k^L\left(x_c-\underline{x}^L\right)-k^H\left(x_c-\underline{x}^H\right)\right]$$

Similar to our approach in the full certification setting, we eliminate the costly certification term for OLS by adding this OLS bias to observed outcomes

$$Y'=Y+D\pi \frac{1}{2f}\left[k^L\left(x_c-\underline{x}^L\right)-k^H\left(x_c-\underline{x}^H\right)\right]$$

However, now we add back the terms to recover the average treatment effects

$$ATT = E\left[Y_1^{'} - Y_0^{'} \mid D = 1\right] - \pi\frac{\left(\underline{x}^H - \underline{x}^L\right)}{2f}E\left[k^L\right]$$

$$ATUT = E\left[Y_1^{'} - Y_0^{'} \mid D = 0\right] - \pi\frac{\left(\underline{x}^H - \underline{x}^L\right)}{2f}E\left[k^H\right]$$

$$\begin{aligned} ATE &= \Pr(D=1)\,ATT + \Pr(D=1)\,ATUT \\ &= E\left[Y_1^{'} - Y_0^{'}\right] - \Pr(D=1)\,\pi\frac{\left(\underline{x}^H - \underline{x}^L\right)}{2f}E\left[k^L\right] \\ &\quad - \Pr(D=0)\,\pi\frac{\left(\underline{x}^H - \underline{x}^L\right)}{2f}E\left[k^H\right] \end{aligned}$$

and

$$\begin{aligned} LATE &= E\left[Y_1^{'} - Y_0^{'} \mid D_1 - D_0 = 1\right] \\ &\quad - \Pr(D_1=1)\,\pi\frac{\left(\underline{x}^H - \underline{x}^L\right)}{2f}E\left[k^L\right] \\ &\quad - \Pr(D_0=0)\,\pi\frac{\left(\underline{x}^H - \underline{x}^L\right)}{2f}E\left[k^H\right] \end{aligned}$$

These terms account for heterogeneity in this asset revaluation setting and are likely to be much smaller than the OLS bias.[13]

Initially, we assume outcomes x, k^j, and P^j for $j = L$ or H are fully observable to the analyst. Begin by letting P, the equilibrium price of uncertified, traded assets, be the reference level in the regression. Since investment choice (treatment) is observable to everyone, we have P^L and P^H. Let β_0 represent P^H and $\beta_0 + \beta_1$ represent P^L in the regression below. Then, changes in means due to ex post outcomes falling in different regions are captured via coefficients on indicator variables.

$$\begin{aligned} E\left[Y' \mid X\right] &= \beta_0 + \beta_1 D + \beta_2 \Im_c^H + \beta_3 \Im_c^L \\ &\quad + \beta_4 \Im_{ck}^H + \beta_5 \Im_{ck}^L + \beta_6 \Im_u^H + \beta_7 \Im_u^L \end{aligned}$$

where

$$Y' = Y + D\pi\left[\Im_{ck}^L k^L - \overline{\Im}_{ck}^H \overline{k}^H\right]$$

$\overline{\Im}_{ck}^H$ and $\overline{k}^H$ are sample averages taken from the $D = 0$ regime, X denotes the matrix of regressors including D, and

$$\begin{aligned} \Im_c^j = \quad & 1 \quad \text{certified range, } x < x_c \\ & 0 \quad \text{otherwise} \\ \Im_{ck}^j = \quad & 1 \quad \text{certified traded} \\ & 0 \quad \text{otherwise} \\ \Im_u^j \quad & 1 \quad \text{untraded asset, } x > P \qquad j \in \{L, H\} \\ & 0 \quad \text{otherwise} \end{aligned}$$

[13]In our running numerical example, the certification cost term for ATT is -0.0882 and for $ATUT$ is -0.882, while the OLS selection bias is 5.3298.

As discussed above, we adjust outcome by the incremental effect of certification cost to eliminate *OLS* selection bias. In this case, certification is stochastic not certain. Therefore, in expectation, we have

$$E\left[Y'\right] = E\left[Y\right] + D\left[\pi\frac{x_c - \underline{x}^L}{2f}\overline{k}^L - \pi\frac{x_c - \underline{x}^H}{2f}\overline{k}^H\right]$$

The incremental impact on mean value of assets in the certification region is reflected in β_2 for high investment and β_3 for low investment firms, while the mean incremental impact of costly certification of assets, k^j, is conveyed via β_4 and β_5 for high and low investment firms, respectively. Finally, the mean incremental impact of untraded assets with values greater than the equilibrium price are conveyed via β_6 and β_7 for high and low investment firms, respectively. The data generating process (DGP) can be written as follows with probabilities replacing the indicator functions and conditional expectations replacing the parameters in the above regression model[14]

$$\begin{aligned}
E\left[Y\right] &= 251.93\Pr\left(P^H\right) - 5.740D\Pr\left(P^L\right) - 94.93\Pr\left(\Im_c^H\right) \\
&\quad -95.49\Pr\left(\Im_c^L\right) - 20\Pr\left(\Im_{ck}^H\right) \\
&\quad -2\Pr\left(\Im_{ck}^L\right) + 31.03\Pr\left(\Im_u^H\right) + 27.60\Pr\left(\Im_u^L\right) \\
E\left[Y\right] &= 251.93\left(0.477\right) - 5.740D\left(0.424\right) - 94.93\left(0.129\right)\Im^H \\
&\quad -95.49\left(0.148\right)\Im^L - 20\left(0.301\right)\Im^H - 2\left(0.345\right)\Im^L \\
&\quad +31.03\left(0.093\right)\Im^H + 27.60\left(0.083\right)\Im^L
\end{aligned}$$

Based on the DGP,

$$\begin{aligned}
ATT &\equiv E\left[Y^L - Y^H \mid D = 1\right] = -12.7 \\
ATUT &\equiv E\left[Y^L - Y^H \mid D = 0\right] = -13.5
\end{aligned}$$

and

$$\begin{aligned}
ATE &\equiv E\left[Y^L - Y^H\right] \\
&= \Pr\left(D = 1\right)ATT + \Pr\left(D = 0\right)ATUT = -13.1
\end{aligned}$$

Hence, in the selective certification setting we encounter modest heterogeneity. Why don't we observe self-selection through the treatment effects? Remember, we have a limited outcome measure. In particular, the outcome excludes investment cost. If we include investment cost,

$$ATT = -12.7 - (101.4 - 114.5) = 0.4$$

and

$$ATUT = -13.5 - (101.4 - 114.5) = -0.4$$

[14] The probabilities reflect likelihood of the condition rather than incremental likelihood and hence sum to one for each investment level (treatment choice).

and self-selection is evidenced through the treatment effects.

Simulation results for the *OLS* model and sample treatment effect statistics are tabulated below.

statistics	β_0	β_1	β_2	β_3
mean	251.9	−11.78	−94.78	−93.81
median	251.9	−11.78	−94.70	−93.85
stand.dev.	0.000	0.157	2.251	2.414
minimum	251.9	−12.15	−102.7	−100.9
maximum	251.9	−11.41	−88.98	−86.90
statistics	β_4	β_5	β_6	β_7
mean	−20.12	−2.087	31.20	27.66
median	−20.15	−2.160	31.23	27.72
stand.dev.	2.697	2.849	1.723	1.896
minimum	−28.67	−9.747	26.91	22.44
maximum	−12.69	8.217	37.14	32.81
OLS parameter estimates for $E[Y' \mid X] = \beta_0 + \beta_1 D + \beta_2 \Im_c^H + \beta_3 \Im_c^L + \beta_4 \Im_{ck}^H + \beta_5 \Im_{ck}^L + \beta_6 \Im_u^H + \beta_7 \Im_u^L$ selective certification setting				
statistics	*estATE*	*estATT*	*estATUT*	
mean	−12.67	−12.29	−13.06	
median	−12.73	−12.33	−13.10	
stand.dev.	2.825	2.686	2.965	
minimum	−21.25	−20.45	−22.03	
maximum	−3.972	−3.960	−3.984	
Average treatment effect estimates				
statistics	*ATE*	*ATT*	*ATUT*	*OLS*
mean	−13.01	−12.57	−13.45	−6.861
median	−13.08	−12.53	−13.46	−6.933
stand.dev.	1.962	2.444	2.947	2.744
minimum	−17.90	−19.52	−22.46	−15.15
maximum	−8.695	−5.786	−6.247	1.466
Average treatment effect sample statistics				

Model-estimated treatment effects are derived in a non-standard manner as the regressors are treatment-type specific and we rely on sample evidence from each regime to estimate the probabilities associated with different ranges of support[15]

$$\begin{aligned} estATT &= \beta_1 - \beta_2 \overline{\Im}_c^H + \beta_3 \overline{\Im}_c^L - \beta_4 \overline{\Im}_{ck}^H + \beta_5 \overline{\Im}_{ck}^L - \beta_6 \overline{\Im}_u^H + \beta_7 \overline{\Im}_u^L \\ &\quad - \pi \overline{k}^L \left(\overline{\Im}_{ck}^L - \overline{\Im}_{ck}^H \right) \end{aligned}$$

[15] Expected value of indicator variables equals the event probability and probabilities vary by treatment. Since there is no common support (across regimes) for the regressors, we effectively assume the analyst can extrapolate to identify counterfactuals (that is, from observed treated to unobserved treated and from observed untreated to unobserved untreated).

$$
\begin{aligned}
estATUT &= \beta_1 - \beta_2 \overline{\Im}_c^H + \beta_3 \overline{\Im}_c^L - \beta_4 \overline{\Im}_{ck}^H + \beta_5 \overline{\Im}_{ck}^L - \beta_6 \overline{\Im}_u^H + \beta_7 \overline{\Im}_u^L \\
&\quad - \pi \overline{k}^H \left(\overline{\Im}_{ck}^L - \overline{\Im}_{ck}^H \right)
\end{aligned}
$$

and

$$
\begin{aligned}
estATE &= \beta_1 - \beta_2 \overline{\Im}_c^H + \beta_3 \overline{\Im}_c^L - \beta_4 \overline{\Im}_{ck}^H + \beta_5 \overline{\Im}_{ck}^L - \beta_6 \overline{\Im}_u^H + \beta_7 \overline{\Im}_u^L \\
&\quad - \overline{D} \pi \overline{k}^L \left(\overline{\Im}_{ck}^L - \overline{\Im}_{ck}^H \right) - \left(1 - \overline{D} \right) \pi \overline{k}^H \left(\overline{\Im}_{ck}^L - \overline{\Im}_{ck}^H \right)
\end{aligned}
$$

where

$$
\overline{\Im}_j^L = \frac{\sum Di \Im_{ji}^L}{\sum Di}
$$

and

$$
\overline{\Im}_j^H = \frac{\sum (1 - D_i) \Im_{ji}^H}{\sum (1 - D_i)}
$$

for indicator j.

OLS effectively estimates the average treatment effects (*ATE*, *ATT*, *ATUT*) in this (modestly heterogeneous) case. However, we're unlikely to be able to detect heterogeneity when the differences are this small. Note in this setting, while outcome is the ex post value net of certification cost, a random sample allows us to assess the owner's ex ante welfare excluding the cost of investment.[16]

We can say a bit more about conditional average treatment effects from the above analysis. On average, owners who select high investment and trade the assets at their equilibrium price sell the assets for 11.78 more than owners who select low investment. Owners who select high investment and retain their assets earn $31.20 - 27.66 = 3.54$ higher proceeds, on average, than owners who select low investment. On the other hand, owners who select high investment and are forced to certify and sell their assets receive lower net proceeds by $20.12 - 2.09 = 18.03$, on average, than owners who select low investment. Recall all outcomes exclude investment cost which, of course, is an important component of owner's welfare.

As we effectively randomize the indicator variables in the analysis and our focus is on identification and estimation of average treatment effects, the remaining analyses are explored without covariates, for simplicity. Next, we demonstrate the above randomization claim via *OLS* then we explore *2SLS-IV* with a binary instrument to estimate *LATE*.

Reduced OLS model We estimate the average treatment effects again via *OLS* but for a reduced model.

$$
E[Y' \mid D] = \beta_0 + \beta_1 D
$$

[16] Investment cost may also be observed or estimable by the analyst.

Results from the simulation are tabulated below.

statistics	β_0	β_1		
mean	207.7	-12.21		
median	207.50	-12.24		
stand.dev.	1.991	2.655		
minimum	202.8	-20.28		
maximum	212.8	-3.957		
OLS parameter estimates for $E[Y' \mid X] = \beta_0 + \beta_1 D$ selective certification setting				
statistics	*estATE*	*estATT*	*estATUT*	
mean	-12.67	-12.29	-13.06	
median	-12.73	-12.33	-13.10	
stand.dev.	2.825	2.686	2.965	
minimum	-21.25	-20.45	-22.03	
maximum	-3.972	-3.960	-3.984	
Average treatment effect estimates				
statistics	*ATE*	*ATT*	*ATUT*	*OLS*
mean	-13.01	-12.57	-13.45	-6.861
median	-13.08	-12.53	-13.46	-6.933
stand.dev.	1.962	2.444	2.947	2.744
minimum	-17.90	-19.52	-22.46	-15.15
maximum	-8.695	-5.786	-6.247	1.466
Average treatment effect sample statistics				

Binary instrument Further, simulation results for *LATE* estimated via *2SLS-IV* are tabulated below.

statistics	β_0	β_1		
mean	207.7	−12.25		
median	207.7	−12.23		
stand.dev.	2.005	2.698		
minimum	202.4	−20.49		
maximum	213.0	−3.128		
2SLS-IV parameter estimates for $E[Y' \mid X] = \beta_0 + \beta_1 D$ selective certification setting				
statistics	*estATE*	*estATT*	*estATUT*	*estLATE*
mean	−12.71	−12.33	−13.09	−12.43
median	−12.73	−12.32	−13.11	−12.72
stand.dev.	2.863	2.727	3.000	2.864
minimum	−21.46	−20.66	−22.25	−22.25
maximum	−3.143	−3.131	−3.155	−3.120
Average treatment effect estimates				
statistics	*ATE*	*ATT*	*ATUT*	*LATE*
mean	−13.01	−12.57	−13.45	−12.59
median	−13.08	−12.53	−13.46	−12.55
stand.dev.	1.962	2.444	2.947	2.468
minimum	−17.90	−19.52	−22.46	−19.63
maximum	−8.695	−5.786	−6.247	−5.842
Average treatment effect sample statistics				

2SLS-IV effectively estimates the treatment effect. Again, it's unlikely the evidence rejects outcome homogeneity even though there are modest differences in the treatments effects.

2.1.1 Propensity score

Propensity score estimates and comparative sample statistics for the treatment effects in the selective certification setting are tabulated below.

statistics	*estATE*	*estATT*	*estATUT*
mean	−12.84	−14.18	−11.47
median	−13.09	−13.71	−11.87
stand.dev.	5.680	6.862	6.262
minimum	−33.93	−49.88	−25.06
maximum	−0.213	1.378	13.80
Propensity score average treatment effect estimates selective certification setting			
statistics	*ATE*	*ATT*	*ATUT*
mean	−13.01	−12.57	−13.45
median	−13.08	−12.53	−13.46
stand.dev.	1.962	2.444	2.947
minimum	−17.90	−19.52	−22.46
maximum	−8.695	−5.786	−6.247
Average treatment effect sample statistics			

As in the full certification setting, the estimates are more variable than we prefer but, on average, correspond with the sample statistics. Again, homogeneity cannot be rejected but estimated differences in treatment effects do not correspond well with the sample statistics (e.g., estimated ATT is the largest in absolute value but ATT is the smallest sample statistic).

2.1.2 Propensity score matching

Simulation results for propensity score matching estimates of average treatment effects in the selective certification setting are tabulated below.

statistics	*estATE*	*estATT*	*estATUT*
mean	−12.90	−12.54	−13.27
median	−13.20	−12.89	−13.09
stand.dev.	3.702	4.478	4.335
minimum	−25.87	−25.54	−26.20
maximum	−4.622	−2.431	−2.532
Propensity score matching average treatment effect estimates selective certification setting			
statistics	*ATE*	*ATT*	*ATUT*
mean	−13.01	−12.57	−13.45
median	−13.08	−12.53	−13.46
stand.dev.	1.962	2.444	2.947
minimum	−17.90	−19.52	−22.46
maximum	−8.695	−5.786	−6.247
Average treatment effect sample statistics			

Again, propensity score matching results correspond well with the sample statistics, are less variable than the propensity score approach above, and cannot reject outcome homogeneity.

2.2 Outcomes measured by value x only

Now, we revisit selective certification when the analyst cannot observe the incremental cost of certification, k, but only asset value, x. Consequently, outcomes and therefore treatment effects reflect only $Y = x$. For instance, the DGP now yields

$$\begin{aligned}
ATT &= E\left[Y^L - Y^H \mid D = 1\right] \\
&= E\left[x^L - x^H \mid D = 1\right] \\
&= 201.4 - 214 = -12.6 \\
ATUT &= E\left[Y^L - Y^H \mid D = 0\right] \\
&= E\left[x^L - x^H \mid D = 0\right] \\
&= 201.4 - 214 = -12.6 \\
ATE &= E\left[Y^L - Y^H\right] \\
&= E\left[x^L - x^H\right] \\
&= \Pr(D = 1)\, ATT + \Pr(D = 0)\, ATUT \\
&= 201.4 - 214 = -12.6 \\
OLS &= E\left[Y^L \mid D = 1\right] - E\left[Y^H \mid D = 0\right] \\
&= E\left[x^L \mid D = 1\right] - E\left[x^H \mid D = 0\right] \\
&= 201.4 - 214 = -12.6
\end{aligned}$$

The apparent advantage to high investment is even more distorted because not only are investment costs excluded but now also the incremental certification costs are excluded. In other words, we have a more limited outcome measure. Next, we quickly summarize similar treatment effect analyses as those reported above but for the alternative, presumably data limited, outcome measure $Y = x$.

2.2.1 OLS results

Simulation results for the *OLS* model and sample treatment effect statistics are tabulated below.

statistics	β_0	$\beta_1\,(estATE)$	
mean	214.0	-12.70	
median	214.1	-12.70	
stand.dev.	1.594	2.355	
minimum	209.3	-18.5	
maximum	218.11	-5.430	
OLS parameter estimates for $E[Y] = \beta_0 + \beta_1 D$ $Y = x$ and selective certification			
statistics	*ATE*	*ATT*	*ATUT*
mean	-12.73	-12.72	-12.75
median	-12.86	-12.78	-12.62
stand.dev.	1.735	2.418	2.384
minimum	-17.26	-19.02	-18.96
maximum	-7.924	-5.563	-6.636
Average treatment effect sample statistics			

OLS effectively estimates the treatment effects and outcome homogeneity is supported.

2.2.2 Binary instrument

Simulation results for *2SLS-IV* with a binary instrument are tabulated below.

statistics	β_0	$\beta_1\,(estLATE)$		
mean	214.1	-12.74		
median	214.2	-12.69		
stand.dev.	1.617	2.428		
minimum	209.2	-18.71		
maximum	217.9	-6.056		
2SLS-IV parameter and average treatment effect estimates for $E[Y] = \beta_0 + \beta_1 D$ $Y = x$ and selection certification				
statistics	*ATE*	*ATT*	*ATUT*	*LATE*
mean	-12.73	-12.72	-12.75	-12.73
median	-12.86	-12.78	-12.62	-12.68
stand.dev.	1.735	2.418	2.384	2.492
minimum	-17.26	-19.02	-18.96	-19.05
maximum	-7.924	-5.563	-6.636	-5.550
Average treatment effect sample statistics				

Again, *2SLS-IV* effectively estimates the treatment effects and the evidence supports outcome homogeneity.

2.2.3 Propensity score

Propensity score estimates and comparative sample statistics for the treatment effects are tabulated below.

statistics	*estATE*	*estATT*	*estATUT*
mean	−13.02	−14.18	−11.86
median	−13.49	−13.96	−11.20
stand.dev.	5.058	5.764	5..680
minimum	−27.00	−34.39	−24.25
maximum	2.451	0.263	7.621
Propensity score average treatment effect estimates for $Y = x$ and selective certification			
statistics	*ATE*	*ATT*	*ATUT*
mean	−12.73	−12.72	−12.75
median	−12.86	−12.78	−12.62
stand.dev.	1.735	2.418	2.384
minimum	−17.26	−19.02	−18.96
maximum	−7.924	−5.563	−6.636
Average treatment effect sample statistics			

As before, the propensity score results are more variable than we'd like but generally consistent with other results.

2.2.4 Propensity score matching

Propensity score matching simulation results are tabulated below.

statistics	*estATE*	*estATT*	*estATUT*
mean	−12.61	−12.43	−12.76
median	−12.83	−12.40	−13.10
stand.dev.	3.239	3.727	4.090
minimum	−20.57	−21.79	−24.24
maximum	−4.025	0.558	−1.800
Propensity score matching average treatment effect estimates for $Y = x$ and selective certification			
statistics	*ATE*	*ATT*	*ATUT*
mean	−12.73	−12.72	−12.75
median	−12.86	−12.78	−12.62
stand.dev.	1.735	2.418	2.384
minimum	−17.26	−19.02	−18.96
maximum	−7.924	−5.563	−6.636
Average treatment effect sample statistics			

Propensity score matching results are generally consistent with other results in supporting homogeneous outcome (for $Y = x$) and, similar to results for the alternative outcome measure, are less variable than the previous propensity score results. Outcome $Y = x$ is more limited than $Y = x - k$, hence for the remaining discussion of this asset revaluation example we refer to the broader outcome measure $Y = x - k$.

2.3 Selective certification with missing "factual" data

It is likely the analyst will not have access to ex post values when the assets are not traded. The only outcome data observed is when assets are certified or when traded at the equilibrium price. In addition to not observing counterfactuals, we now face missing factual data. Missing outcome data produces a challenging treatment effect identification problem. The treatment effects are the same as the above observed data case but require some creative data augmentation to recover. We begin our exploration by examining model-based estimates if we ignore the missing data problem.

If we ignore missing data and estimate the model via *OLS* we find the following simulation results.

statistics	β_0	β_1
mean	207.2	−9.992
median	207.2	−9.811
stand.dev.	2.459	3.255
minimum	200.9	−18.30
maximum	213.2	−2.627

OLS parameter estimates for
$E[Y' \mid X] = \beta_0 + \beta_1 D$
selective certification, ignoring missing data

statistics	*estATE*	*estATT*	*estATUT*
mean	−10.45	−10.07	−10.81
median	−9.871	−5.270	−14.92
stand.dev.	3.423	3.285	3.561
minimum	−19.11	−18.44	−19.75
maximum	−2.700	−2.640	−2.762

Average treatment effect estimates ignoring missing data

statistics	*ATE*	*ATT*	*ATUT*	*OLS*
mean	−13.01	−12.57	−13.45	−6.861
median	−13.08	−12.53	−13.46	−6.933
stand.dev.	1.962	2.444	2.947	2.744
minimum	−17.90	−19.52	−22.46	−15.15
maximum	−8.695	−5.786	−6.247	1.466

Average treatment effect sample statistics

The average model-estimated treatment effects are biased toward zero due to the missing outcome data.

2.3.1 Binary instrument

Using the binary instrument z discussed earlier we estimate *LATE* via *2SLS-IV*. Simulation results are tabulated below.

statistics	β_0	β_1		*estLATE*
mean	207.2	-10.06		-11.10
median	207.3	-9.901		-11.06
stand.dev.	2.466	3.236		3.261
minimum	201.4	-18.71		-19.66
maximum	213.5	-2.636		-0.078
2SLS-IV parameter estimates for $E[Y' \mid X] = \beta_0 + \beta_1 D$ selective certification, ignoring missing data				
statistics	*ATE*	*ATT*	*ATUT*	*LATE*
mean	-13.01	-12.57	-13.45	-12.59
median	-13.08	-12.53	-13.46	-12.55
stand.dev.	1.962	2.444	2.947	2.468
minimum	-17.90	-19.52	-22.46	-19.63
maximum	-8.695	-5.786	-6.247	-5.842
Average treatment effect sample statistics ignoring missing data				

Like the *OLS* results, the *IV* treatment effect estimates are biased towards zero.

2.3.2 Data augmentation

The above results suggest attending to the missing data. The observed data may not, in general, be representative of the missing factual data. We might attempt to model the missing data process and augment the observed data. Though in this setting, augmented data might generate poorer estimates of the treatment effects as the missing data produce offsetting effects as described below. The observed data are

$$Y_1^o = \Im_{ck}^L \left(x^L - k^L\right) + \Im_p^L P^L$$

and

$$Y_0^o = \Im_{ck}^H \left(x^H - k^H\right) + \Im_p^H P^H$$

where

$$\Im_p^j = \begin{array}{cc} 1 & \text{asset traded at uncertified, equilibrium price for choice } j \\ 0 & \text{otherwise} \end{array}$$

and $\Im_{ck}^j$ refers to assets certified and traded for choice j, as before.

For the region $x < x_c$, we have data for firms forced to sell $x^j - k^j$, we are missing data x^j. Therefore, if we know k^j or can estimate it, we can model the missing data for this region. Since $I^H > I^L$, $E\left[x^H \mid x^H < x_c\right] > E\left[x^L \mid x^L < x_c\right]$ but $\Pr\left(x^L < x_c\right) > \Pr\left(x^H < x_c\right)$. Hence, the impact on the average treatment effect for this region is ambiguous. We adjust the effect estimated from the data on assets forced sold and certified by the certification cost

$$\begin{aligned} & \frac{(1-\pi)}{\pi}\left(\frac{\sum \Im^{L}_{ck,i}\left(x^{L}_{i}-k^{L}_{i}\right)}{\sum D_{i}}-\frac{\sum \Im^{H}_{ck,i}\left(x^{H}_{i}-k^{H}_{i}\right)}{\sum (1-D_{i})}\right) \\ & +\frac{(1-\pi)}{\pi}\left(\frac{\sum \Im^{L}_{ck,i}k^{L}_{i}}{\sum D_{i}}-\frac{\sum \Im^{H}_{ck,i}k^{H}_{i}}{\sum (1-D_{i})}\right) \\ = \quad & \frac{(1-\pi)}{\pi}\left[\frac{\sum \Im^{L}_{ck,i}x^{L}_{i}}{\sum D_{i}}-\frac{\sum \Im^{H}_{ck,i}x^{H}_{i}}{\sum (1-D_{i})}\right] \end{aligned}$$

Based on the DGP for our continuing example, this quantity is $22.289-20.253 = 2.036$.

The other region, $x^j > P^j$ for untraded assets, is more delicate as we have no direct evidence and the conditional expectation over this region differs by investment choice. Since $P^H > P^L$ it is likely $E\left[x^H \mid x^H > P^H\right] > E\left[x^L \mid x^L > P^L\right]$. Based on the DGP for our continuing example, this quantity is $22.674-26.345 = -3.671$. How do we model missing data in this region?

This is not a typical censoring problem as we don't observe the sample size for either missing data region. Missing samples make estimating the probability of each mean level more problematic — recall this is important for estimating average treatment effects in the data observed, selective certification case.[17] Conditional expectations and probabilities of mean levels are almost surely related which implies any augmentation errors will be amplified in the treatment effect estimate.

We cannot infer the probability distribution for x by nonparametric methods since x is unobserved. To see this, recall the equilibrium pricing of uncertified assets satisfies

$$\begin{aligned} P \quad = \quad & \frac{\pi \Pr\left(x_c < x < \overline{x}\right) E\left[x \mid x_c < x < \overline{x}\right]}{\pi \Pr\left(x_c < x < \overline{x}\right) + (1-\pi)\Pr\left(x_c < x < P\right)} \\ & +\frac{(1-\pi)\Pr\left(x_c < x < P\right) E\left[x \mid x_c < x < P\right]}{\pi \Pr\left(x_c < x < \overline{x}\right) + (1-\pi)\Pr\left(x_c < x < P\right)} \end{aligned}$$

For instance, if all the probability mass in these intervals for x is associated with P, then the equilibrium condition is satisfied. But the equilibrium condition is satisfied for other varieties of distributions for x as well. Hence, the distribution for x cannot be inferred when x is unobserved. If π is known we can estimate $\Pr\left(x_c < x < \overline{x}\right)$ from certification frequency scaled by π. However, this still leaves much of the missing factual data process unidentified when x is unobserved or the distribution for x is unknown. On the other hand, if the distribution for x is known then π can be inferred from observable data, P and x_c, as well as the support for x: $\underline{x} < x < \overline{x}$. Further, if the distribution of

[17] As is typical, identification and estimation of treatment effects in this setting is more delicate than identification and estimation of model parameters.

x is known we are able to model the DGP for the missing factual data. In particular, when the distribution for x is known we can infer π and identify $\Pr(\underline{x} < x < x_c)$, $E[x \mid \underline{x} < x < x_c]$, $\Pr(P < x < \overline{x})$, and $E[x \mid P < x < \overline{x}]$.

To model missing factual data we assume π is known and k^j is observed, $\Pr(P < x < \overline{x}) = \Pr(x_c < x < P)$, and $E[x \mid P < x < \overline{x}] = P + \frac{P - x_c}{2} = \frac{3P - x_c}{2}$. Based on the model for missing factual data, estimated average treatment effects (TE) are

$$\begin{aligned} estTE &= \text{above } TE \text{ estimate} \\ &\quad + \frac{(1-\pi)}{\pi}\left(\frac{\sum \Im^L_{ck,i}\left(x^L_i - k^L_i\right)}{\sum D_i} - \frac{\sum \Im^H_{ck,i}\left(x^H_i - k^H_i\right)}{\sum (1 - D_i)}\right) \\ &\quad + \frac{(1-\pi)}{\pi}\left(\frac{\sum \Im^L_{ck,i} k^L_i}{\sum D_i} - \frac{\sum \Im^H_{ck,i} k^H_i}{\sum (1 - D_i)}\right) \\ &\quad + \frac{(1-\pi)}{1+\pi}\left[\frac{3P^L - x_c}{2}\frac{\sum \Im^L_{P,i}}{\sum D_i} - \frac{3P^H - x_c}{2}\frac{\sum \Im^H_{P,i}}{\sum (1 - D_i)}\right] \end{aligned}$$

Results adjusted by the augmented factual missing data based on the previous OLS parameter estimates are tabulated below.

statistics	*estATE*	*estATT*	*estATUT*	
mean	−11.80	−11.43	−12.18	
median	−11.76	−11.36	−12.06	
stand.dev.	3.165	3.041	3.290	
minimum	−20.37	−19.58	−21.15	
maximum	−2.375	−2.467	−2.280	
Treatment effect *OLS* model estimates based on augmented missing data selective certification				
statistics	*ATE*	*ATT*	*ATUT*	*OLS*
mean	−13.01	−12.57	−13.45	−6.861
median	−13.08	−12.53	−13.46	−6.933
stand.dev.	1.962	2.444	2.947	2.744
minimum	−17.90	−19.52	−22.46	−15.15
maximum	−8.695	−5.786	−6.247	1.466
Average treatment effect sample statistics				

These augmented-OLS results are less biased, on average, than results that ignore missing factual data. Thus, it appears data augmentation has modestly aided our analysis of this asset revaluation setting.

Simulation results for the augmented factual missing data based on *2SLS-IV* parameter estimates are tabulated below.

statistics	*estLATE*	*LATE*
mean	-11.79	-12.57
median	-11.48	-12.53
stand.dev.	1.739	2.444
minimum	-20.31	-19.52
maximum	-1.737	-5.786

Average treatment effect sample statistics and *2SLS-IV* model estimates based on the augmented missing data, selective certification

Like the augmented *OLS* results, the augmented *2SLS-IV* results are also less biased than those that ignore missing factual data (this is not surprising as data augmentation has the same impact on both analyses). However, we should bear in mind there is risk in modeling missing data, when the effect of missing data on the estimands of interest is small we might inject more error by attempting to model the missing data. If our probability and outcome assumptions are badly flawed we expect poor treatment effect estimates.

2.4 Sharp regression discontinuity design

Suppose the *DGP* is altered only in that

$$k^L \sim uniform\,(1,3)$$

and

$$k^H \sim uniform\,(3,37)$$

The means for k remain 2 and 20 but we have adjacent support. There is a crisp break at $k=3$ but the regression function excluding the treatment effect (the regression as a function of k) is continuous. That is, the treatment effect fully accounts for the discontinuity in the regression function. This is a classic "sharp" regression discontinuity design (Trochim [1984] and Angrist and Pischke [2009]) where β_2 estimates the average treatment effect via *OLS*.

$$E\,[Y \mid D] = \beta_0 + \beta_1 k + \beta_2 D$$

With the previous *DGP*, there was discontinuity as a function of both the regressor k and treatment D. This creates a problem for the regression as least squares is unable to distinguish the treatment effect from the jump in the outcome regression and leads to poor estimation results. In this revised setting, we anticipate substantially improved (finite sample) results.

2.4.1 Full certification setting

Simulation results for the revised *DGP* in the full certification setting are tabulated below.

statistics	β_0	β_1	$\beta_2\,(estATE)$
mean	214.2	-1.007	-12.93
median	214.5	-1.019	-13.04
stand.dev.	4.198	0.190	4.519
minimum	203.4	-1.503	-26.18
maximum	226.3	-0.539	-1.959
Sharp *RD* *OLS* parameter estimates for $E[Y \mid D] = \beta_0 + \beta_1 k + \beta_2 D$ full certification setting			
statistics	*ATE*	*ATT*	*ATUT*
mean	-12.54	-12.49	-12.59
median	-12.55	-12.44	-12.68
stand.dev.	1.947	2.579	2.794
minimum	-17.62	-19.53	-21.53
maximum	-7.718	-6.014	-6.083
Average treatment effect sample statistics			

Unlike the previous *DGP*, sharp regression discontinuity (*RD*) design effectively identifies the average treatment effect and *OLS* produces reliable estimates for the (simple) full certification setting. Next, we re-evaluate *RD* with the same adjacent support *DGP* but in the (more challenging) selective certification setting.

2.4.2 Selective certification setting

To satisfy the continuity condition for the regression, cost of certification $k = Dk^L + (1 - D)\,k^H$ is always observed whether assets are certified or not in the regression discontinuity analysis of selective certification. Simulation results for the revised *DGP* in the selective certification setting are tabulated below.[18]

[18] We report results only for the reduced model. If the analyst knows where support changes (i.e., can identify the indicator variables) for the full model, the results are similar and the estimates have greater precision.

statistics	β_0	β_1	$\beta_2\,(estATE)$
mean	214.2	−0.299	−13.00
median	214.5	−0.324	−12.89
stand.dev.	4.273	0.197	4.546
minimum	202.0	−0.788	−25.81
maximum	225.5	0.226	−1.886
Sharp *RD OLS* parameter estimates for $E[Y \mid D] = \beta_0 + \beta_1 k + \beta_2 D$ selective certification setting			
statistics	*ATE*	*ATT*	*ATUT*
mean	−12.54	−12.49	−12.59
median	−12.55	−12.44	−12.68
stand.dev.	1.947	2.579	2.794
minimum	−17.62	−19.53	−21.53
maximum	−7.718	−6.014	−6.083
Average treatment effect sample statistics			

In the selective certification setting, *RD* again identifies the average treatment effect and *OLS* provides effective estimates. Next, we employ *RD* in the missing factual data setting.

2.4.3 Missing factual data

If some outcome data are unobserved by the analyst, it may be imprudent to ignore the issue. We employ the same missing data modeling assumptions as before and estimate the average treatment effect ignoring missing outcome data (β_2) and the average treatment effect adjusted for missing outcome data ($\beta_2^{'}$). Simulation results for the revised *DGP* (with adjacent support) in the selective certification setting with missing outcome data are tabulated below.

statistics	β_0	β_1	β_2	$\beta_2^{'}\,(estATE)$
mean	214.4	−0.342	−11.35	−12.22
median	214.5	−0.336	−11.50	−12.47
stand.dev.	4.800	0.232	5.408	5.237
minimum	201.3	−0.928	−25.92	−26.50
maximum	227.9	0.325	2.542	1.383
Sharp *RD OLS* parameter estimates for $E[Y \mid D] = \beta_0 + \beta_1 k + \beta_2 D$ selective certification, missing data				
statistics		*ATE*	*ATT*	*ATUT*
mean		−12.54	−12.49	−12.59
median		−12.55	−12.44	−12.68
stand.dev.		1.947	2.579	2.794
minimum		−17.62	−19.53	−21.53
maximum		−7.718	−6.014	−6.083
Average treatment effect sample statistics				

2.5 Fuzzy regression discontinuity design

Now, suppose the DGP is altered only in that support is overlapping as follows:

$$k^L \sim uniform\,(1,3)$$

and

$$k^H \sim uniform\,(1,39)$$

The means for k remain 2 and 20 but we have overlapping support. There is a crisp break in $E\,[D \mid k]$ at $k = 3$ but the regression function excluding the treatment effect (the regression as a function of k) is continuous. This leads to a fuzzy discontinuity regression design (van der Klaauw [2002]). Angrist and Lavy [1999] argue that *2SLS-IV* consistently estimates a local average treatment effect in such case where

$$T = \begin{array}{ll} 1 & k \leq 3 \\ 0 & k > 3 \end{array}$$

serves as an instrument for treatment. In the first stage, we estimate the propensity score[19]

$$\widehat{D} \equiv E\,[D \mid k, T] = \gamma_0 + \gamma_1 k + \gamma_2 T$$

The second stage is then

$$E\,[Y \mid k, D] = \gamma_0 + \gamma_1 k + \gamma_2 \widehat{D}$$

2.5.1 Full certification setting

First, we estimate *RD* via *OLS* then we employ *2SLS-IV*. Simulation results for the overlapping support *DGP* in the full certification setting are tabulated below.

statistics	β_0	β_1	$\beta_2\,(estATE)$
mean	214.3	−1.012	−12.79
median	214.2	−1.011	−12.56
stand.dev.	3.634	0.163	3.769
minimum	204.9	−1.415	−23.51
maximum	222.5	−0.625	−3.001

Fuzzy *RD OLS* parameter estimates for
$E\,[Y \mid D] = \beta_0 + \beta_1 k + \beta_2 D$
full certification setting

statistics	*ATE*	*ATT*	*ATUT*
mean	−12.54	−12.49	−12.59
median	−12.55	−12.44	−12.68
stand.dev.	1.947	2.579	2.794
minimum	−17.62	−19.53	−21.53
maximum	−7.718	−6.014	−6.083

Average treatment effect sample statistics

[19] In this asset revaluation setting, the relations are linear. More generally, high order polynomial or nonparametric regressions are employed to accommodate nonlinearities (see Angrist and Pischke [2009]).

Perhaps surprisingly, *OLS* effectively estimates the average treatment effect in this fuzzy *RD* setting. Recall the selection bias is entirely due to the expected difference in certification cost, $E\left[k^H - k^L\right]$. *RD* models outcome as a (regression) function of k, $E[Y \mid k]$; hence, the selection bias is eliminated from the treatment effect. Next, we use *2SLS-IV* to estimate *LATE*.

Binary instrument Now, we utilize T as a binary instrument. Simulation results for the overlapping support *DGP* in the full certification setting are tabulated below.

statistics	β_0	β_1	$\beta_2\,(estLATE)$
mean	214.5	-1.020	-13.07
median	214.6	-1.021	-13.27
stand.dev.	4.139	0.181	4.456
minimum	202.7	-1.461	-27.60
maximum	226.0	-0.630	-1.669

Fuzzy *RD* *2SLS-IV* parameter estimates for
$E[Y \mid D] = \beta_0 + \beta_1 k + \beta_2 D$
full certification setting

statistics	*ATE*	*ATT*	*ATUT*	*LATE*
mean	-12.54	-12.49	-12.59	-12.51
median	-12.55	-12.44	-12.68	-12.41
stand.dev.	1.947	2.579	2.794	2.599
minimum	-17.62	-19.53	-21.53	-19.58
maximum	-7.718	-6.014	-6.083	-5.969

Average treatment effect sample statistics

As expected, *2SLS-IV* effectively identifies *LATE* in this fuzzy *RD*, full certification setting. Next, we revisit selective certification with this overlapping support *DGP*.

2.6 Selective certification setting

First, we estimate *RD* via *OLS* then we employ *2SLS-IV*. Simulation results for the overlapping support *DGP* in the selective certification setting are tabulated below.

statistics	β_0	β_1	β_2 (*estATE*)
mean	214.3	−0.315	−12.93
median	214.1	−0.311	−12.73
stand.dev.	3.896	0.179	3.950
minimum	202.5	−0.758	−24.54
maximum	223.3	0.078	−3.201

Fuzzy *RD OLS* parameter estimates for
$E[Y \mid D] = \beta_0 + \beta_1 k + \beta_2 D$
selective certification setting

statistics	*ATE*	*ATT*	*ATUT*
mean	−12.54	−12.49	−12.59
median	−12.55	−12.44	−12.68
stand.dev.	1.947	2.579	2.794
minimum	−17.62	−19.53	−21.53
maximum	−7.718	−6.014	−6.083

Average treatment effect sample statistics

Since *RD* effectively controls the selection bias (as discussed above), *OLS* effectively estimate the average treatment effect.

Binary instrument Using T as a binary instrument, *2SLS-IV* simulation results for the overlapping support *DGP* in the selective certification setting are tabulated below.

statistics	β_0	β_1	β_2 (*estLATE*)
mean	214.4	−0.321	−13.09
median	214.5	−0.317	−13.03
stand.dev.	4.438	0.200	4.631
minimum	201.1	−0.805	−27.23
maximum	225.6	−0.131	1.742

Fuzzy *RD 2SLS-IV* parameter estimates for
$E[Y \mid D] = \beta_0 + \beta_1 k + \beta_2 D$
selective certification setting

statistics	*ATE*	*ATT*	*ATUT*	*LATE*
mean	−12.54	−12.49	−12.59	−12.51
median	−12.55	−12.44	−12.68	−12.41
stand.dev.	1.947	2.579	2.794	2.599
minimum	−17.62	−19.53	−21.53	−19.58
maximum	−7.718	−6.014	−6.083	−5.969

Average treatment effect sample statistics

In the selective certification setting, *2SLS-IV* effectively estimates *LATE*, as anticipated.

Missing factual data Continue with the overlapping support *DGP* and employ the same missing data modeling assumptions as before to address unobserved outcomes (by the analyst) when the assets are untraded. First, we report *OLS* simulation results then we tabulate *2SLS-IV* simulation results where β_2 is the estimated for average treatment effect ignoring missing outcome data and $\beta_2^{'}$ is the average treatment effect adjusted for missing outcome data.

$statistics$	β_0	β_1	β_2	$\beta_2^{'}\,(estATE)$
$mean$	215.9	−0.426	−12.74	−13.60
$median$	216.2	−0.424	−12.63	−13.52
$stand.dev.$	4.765	0.223	4.792	4.612
$minimum$	201.9	−1.132	−24.20	−23.85
$maximum$	226.3	0.117	0.119	−0.817
Fuzzy *RD OLS* parameter estimates for $E[Y \mid D] = \beta_0 + \beta_1 k + \beta_2 D$ selective certification, missing data				
$statistics$		ATE	ATT	$ATUT$
$mean$		−12.54	−12.49	−12.59
$median$		−12.55	−12.44	−12.68
$stand.dev.$		1.947	2.579	2.794
$minimum$		−17.62	−19.53	−21.53
$maximum$		−7.718	−6.014	−6.083
Average treatment effect sample statistics				

This *OLS RD* model for missing outcome data does not offer any clear advantages. Rather, the results seem to be slightly better without the missing data adjustments.

2SLS-IV with T as a binary instrument and missing outcome data adjustments are considered next. Simulation results for the overlapping support *DGP* in the selective certification, missing outcome data setting are tabulated below.

$statistics$	β_0	β_1	β_2	$\beta_2^{'}\,(estLATE)$
$mean$	217.7	−0.428	−12.80	−13.67
$median$	214.8	−0.425	−13.12	−14.30
$stand.dev.$	25.50	0.256	5.919	5.773
$minimum$	139.2	−1.147	−25.24	−25.97
$maximum$	293.9	0.212	6.808	6.010
Fuzzy *RD 2SLS-IV* parameter estimates for $E[Y \mid D] = \beta_0 + \beta_1 k + \beta_2 D$ selective certification, missing data				
$statistics$	ATE	ATT	$ATUT$	$LATE$
$mean$	−12.54	−12.49	−12.59	−12.51
$median$	−12.55	−12.44	−12.68	−12.41
$stand.dev.$	1.947	2.579	2.794	2.599
$minimum$	−17.62	−19.53	−21.53	−19.58
$maximum$	−7.718	−6.014	−6.083	−5.969
Average treatment effect sample statistics				

Again, modeling the missing outcome data offers no apparent advantage in this fuzzy *RD*, *2SLS-IV* setting. In summary, when we have adjacent or overlapping support, sharp or fuzzy regression discontinuity designs appear to be very effective for controlling selection bias and identifying average treatment effects in this asset revaluation setting.

2.7 Common support

Standard identification conditions associated with ignorable treatment (and IV approaches as well) except for regression discontinuity designs include common support $0 < \Pr(D = 1 \mid X) < 1$. As indicated earlier, this condition fails in the asset revaluation setting as certification cost type is a perfect predictor of treatment $\Pr(D = 1 \mid T = 1) = 1$ and $\Pr(D = 1 \mid T = 0) = 0$ where $T = 1$ if type is L and zero otherwise. The foregoing discussion has addressed this issue in two ways. First, we employed an ad hoc adjustment of outcome to eliminate selection bias. This may be difficult or impractical to implement. Second, we employed a regression discontinuity design. The second approach may be unsatisfactory as the analyst needs full support access to the regressor k whether assets are certified or not.

However, if there is some noise in the relation between certification cost type and treatment (perhaps, due to nonpecuniary cost or benefit), then a third option may have some utility. We briefly illustrate this third possibility for the full certification setting.

Suppose everything remains as in the original full certification setting except $k^H \sim uniform(1, 19)$ and some owners select treatment (lower investment) when certification cost is high, hence $\Pr(D = 1 \mid \text{type} = H) = 0.1$. This setup implies observed outcome is

$$Y = D[(Y_1 \mid T = 1) + (Y_1 \mid T = 0)] + (1 - D)(Y_0 \mid T = 0)$$

such that

$$E[Y] = 0.5E\left[x^L - k^L\right] + 0.5\left\{0.1E\left[x^L - k^H\right] + 0.9E\left[x^H - k^H\right]\right\}$$

Suppose the analyst ex post observes the actual certification cost type. The common support condition is satisfied as $0 < \Pr(D = 1 \mid T = 0) < 1$ and if outcomes are conditionally mean independent of treatment given T then treatment is ignorable. The intuition is type T controls the selection bias and allows D to capture the treatment effect. This involves a delicate balance as T and D must be closely but imperfectly related.

OLS simulation results are tabulated below.

statistics	β_0	β_1	$\beta_2\,(estATE)$
mean	196.9	7.667	−5.141
median	196.9	7.896	−5.223
stand.dev.	1.812	6.516	6.630
minimum	191.5	−10.62	−23.54
maximum	201.6	25.56	14.25
OLS parameter estimates for $E[Y \mid D] = \beta_0 + \beta_1 T + \beta_2 D$ full certification setting			

statistics	*ATE*	*ATT*	*ATUT*
mean	−5.637	−5.522	−5.782
median	−5.792	−5.469	−5.832
stand.dev.	1.947	2.361	2.770
minimum	−9.930	−12.05	−12.12
maximum	0.118	0.182	0.983
Average treatment effect sample statistics			

The estimated average treatment effect is slightly attenuated and has high variability that may compromise its finite sample utility. Nevertheless, the results are a dramatic departure and improvement from the results above where the common support condtion fails and is ignored.

2.8 Summary

Outcomes at our disposal in this asset revaluation setting limit our ability to assess welfare implications for the owners. Nonetheless, the example effectively points to the importance of recognizing differences in data available to the analyst compared with information in the hands of the economic agents whose actions and welfare is the subject of study. To wit, treatment effects in this setting are uniformly negative. This is a product of comparing net gains associated with equilibrium investment levels, but net gains exclude investment cost. The benefits of higher investment when certification costs are low are not sufficient to overcome the cost of investment but this latter feature is not reflected in our outcome measure. Hence, if care is not exercised in interpreting the results we might draw erroneous conclusions from the data.