Ratios of Changes:
How Real Estate Shocks Did Not Affect Corporate Investment

Ivo Welch

2020-08-04

Abstract

Real-estate price shocks did not positively associate with corporate investment from 1992 to 2008, as suggested by Chaney, Sraer, and Thesmar (2012). Most of their coefficients explain changes in their firm-scale normalization (PP&E), not changes in their variables of interest (real-estate and investment). This problem cannot be cured by an additive 1/PP&E control. By working with "ratios of changes" instead of "changes of ratios," denominator-caused scaling volatility can be purged, and their estimated coefficients then turn negative. My paper also discusses the lack of shock identification and the effects of time trends. A sample extended to 2018 also shows no effect.

JEL D22, G31, R30.
https://ssrn.com/abstract=3599280

I think Thomas Chaney, David Sraer, and David Thesmar for sharing their data and for helpful conversations. Their response (in https://ssrn.com/abstract=3599277) to my earlier draft helped me improve this paper greatly. The reader should be aware that my paper points out problems in Chaney, Sraer, and Thesmar (2012) that ultimately arose from their adoption of a specification that is standard in the literature. Thus, the same problems could be important in other papers, too. (However, until reexamined, it is not clear whether they matter elsewhere.) All remaining mistakes are my own. I would especially welcome comments if my analysis here is wrong. (I would also welcome pointers if my verbiage anywhere is offensive rather than objective. And I would welcome pointers to relevant literature that I should have cited but did not.)

With the financial crisis of 2008, real estate moved to the center of attention for financial economists. Widely considered seminal, Chaney, Sraer, and Thesmar (2012), henceforth sometimes also abbreviated as CST12, suggested that real-estate holdings facilitated collateral that in turn facilitated corporate investment. Their paper provides a powerful narrative for one aspect of the origins of the decline in corporate capital spending in the Great Recession. Policy-wise, it implies that support of the real-estate market through lower mortgage rates could have a secondary positive effect on corporate investment, too.

Their hypothesis can be viewed in levels (with the key objects being real-estate value and investment) or in changes (with the key objects being real-estate *price appreciation* and *incremental investments*). Although the empirical implementation in CST12 examines *level investments* and *level real-estate values*, the specification becomes similar to a test in changes through their inclusion of firm dummies. Chaney, Sraer, and Thesmar (2012, p.2389) explain their main specification as follows:

> ...for firm i, at date t, with headquarters in some location l (state or MSA), investment is given by
>
> $$\text{INV}_{it} = a_i + g_t + b \cdot \text{REValue}_{it} + s \cdot P_t^l + \text{controls}_{it} + e_{it} \ , \qquad (1)$$
>
> where INV is the ratio of investment to lagged PPE, REValue_{it} is the ratio of the market value of real-estate assets in year t to lagged PPE, and P_{it} controls for the level of prices in location l in year t.

More precisely, "INV" is capital expenditure divided by lagged PPE (Property, Plant, and Equipment). "REValue" is the firm's real-estate holdings value multiplied by the local real-estate price index and also divided by lagged PPE. The set of controls often contains cash earnings (also divided by lagged PPE), the lagged market/book ratio of the firm (winsorized to prevent negative values), the real-estate price index series itself, and the full set of year and firm dummies. In Section II.B., titled *Main Results*, Chaney, Sraer, and Thesmar (2012) present Table 5, their key estimated empirical results:

1. In their first six columns, they regress

$$\frac{\text{capex}_{i,y}}{\text{PPE}_{i,y-1}} = b \cdot \frac{\text{RE (dollar value)}_{i,y}}{\text{PPE}_{i,y-1}} + s \cdot \text{repi.*}_{(i,)y} + \text{controls}_{i,y} + \text{FE}_i + \text{FE}_y + e_{i,y} \ , \tag{2}$$

where 'FE' means fixed effect, and 'repi' is the local real-estate price index applicable to firm i in year y. The dollar value of real-estate holdings is initially obtained from a snapshot value from a now-discontinued variable on Compustat at incept time (typically 1992). Changes over time are obtained by using repi as the multiplier on the incept real-estate dollar value.

[Insert Table 1 here: **Glossary of Variable Names Used Repeatedly**]

Because my paper needs to change some variables, Table 1 shows a glossary of revised variable names as used in my paper. With the variable definitions in Table 1, their estimation equation (2) becomes

$$\text{capex/lagPPE} = b \cdot \text{redol/lagPPE} + s \cdot \text{repi} + a_i \cdot \{\text{firmids}\} + g_y \cdot \{\text{yearids}\} + e \ , \tag{3}$$

where {...} denote a full set of dummies

2. In their last two columns, they replace the independent variable redol/lagPPE with a dummy for whether the firm has real estate (REisPOS) multiplied by the (time-moving) repi real-estate price series.

$$\text{capex/lagPPE} = b \cdot \text{REisPOS*repi.offc} + s \cdot \text{repi.offc} + \cdots \ . \tag{4}$$

(Recall that firms have no time-variation in whether they own or do not own real estate.) Holding repi constant, this coefficient thus measures whether the specific set of firms with real estate have a higher response coefficient to real-estate price index changes. Their specifications also add many other controls.

Explaining their estimate in the first column of Table 5 (*Main Results*), they write

> Column 1 starts with the simplest estimation of equation (1) without any additional controls. Land-holding firms increase their investment more than non–land-holding firms when real-estate prices increase. The baseline coefficient is 0.074, so that each additional $1 of real-estate collateral increases investment by $0.074. The coefficient is significant at the 1 percent confidence level.

Their abstract emphasizes the magnitude of this and other findings: "Over the 1993-2007 period, the representative US corporation invests $0.06 out of each $1 of collateral."

The above should be an uncontroversial summary of Chaney, Sraer, and Thesmar (2012). My own paper offers a different view on their evidence (and that in Chaney, Sraer, and Thesmar (2020)).

Most of their results are due to a common but faulty scaling control. Their coefficients explain shared changes and trends in their normalizing denominator, the *time-varying* (lagged) property, plant, and equipment (lagPPE). This is especially the case in 29 out of 33 specifications (as in equation 3). These use the same scaling not only in their dependent variable (capex, capital expenditures), but also in their independent variable (redol, real-estate dollar holdings). In this case, the effect is nearly mechanical.

In their defense, Chaney, Sraer, and Thesmar (2012) did not invent such shared-scaling specifications. Instead, they have been common fixtures in the literature—including in such seminal corporate-finance papers as Fazzari, Hubbard, and Petersen (2000), Baker, Stein, and Wurgler (2003), Almeida, Campello, and Weisbach (2004), Rauh (2006), and many others. (It is however unknown whether scaling causes problems of material consequence in any other papers.)

My paper first shows that due to volatility in their denominator, the same empirical inference strategy as that in CST12 can provide support for absurd hypotheses about other numerators—such as the claim that the constant number 1.0 influences corporate investment. Instead, the correct interpretation is that the coefficient on 1/lagPPE explaining capex/lagPPE with fixed effects merely measures shared variation in denominators.

In response, Chaney, Sraer, and Thesmar (2020) suggest an additive control term (for 1/lagPPE). They show that their real-estate coefficient declines but remains strong. However, my updated paper here shows that this additive control term cannot correct the problem. Even under simple placebo null processes, the real-estate coefficient remains robustly positive despite inclusion of the 1/lagPPE control. Moreover, my paper shows that lagPPE scaling also introduced non-linearities into their capital expenditures measure that were picked up by their real-estate measure.

Yet there is a better approach that eliminates the problem of unwanted volatility in the scaling denominator altogether. Instead of using level regressions with fixed effects, this better approach is easiest to understand and implement by switching to differences regressions (Mundlak (1978), Angrist and Pischke (2008, Chapter 5)). The proposed scaling improvement then involves switching from "changes of ratios" (CoRs, as in $\Delta[v/d]$) to "ratios of changes" (RoCs, as in $[\Delta v]/d$). In panel regressions with RoC inputs, the denominator returns to its original intended role of scaler (as in the weighted-least-squares sense), and the coefficient estimates become largely immune to volatility in the denominator. This reduces undesirable noise and bias. In Chaney, Sraer, and Thesmar (2012), all CoR specifications yield positive coefficients; all RoC alternatives yield negative coefficients.

In their remaining 4 out of 33 regressions without mechanical correlation (as in equation 4), Chaney, Sraer, and Thesmar (2012) introduce a new independent variable: a dummy for whether the firm owned real estate multiplied by the real-estate price series (REisPOS∗repi), holding the real-estate price series (repi) constant. Their dependent corporate investment variable continues to be a level ratio. With fixed effects in the regression, it acts in effect like a CoR. Thus, the 1/lagPPE denominator in their dependent capex/lagPPE variable can and indeed does correlate with their new independent variable, too. Although the positive correlation here is not mechanical, it is still spurious. And again, the coefficients are negative when the regressions are specified in terms of RoC differences.

In sum, the empirical evidence in Chaney, Sraer, and Thesmar (2012) is not sufficient to support the view that real-estate collateral increases had positively associated with corporate investment. Moreover, neither is the new evidence in Chaney, Sraer, and Thesmar (2020).

In the last section of my paper, I explore the association between real estate and capital expenditures beyond 2007. This includes the period of the Great Recession and subsequent recovery. There is no evidence of a positive association in this expanded sample, either, nor does there seem to be a time-series (non-cross-sectional) association in the aggregate.

The appendices point out two further issues. First, real-estate prices were too persistent to allow empirical identification of shocks. Second, there were *differential* investment trends of large firms with real-estate vs. small firms without real-estate that were not picked up by the fixed effects. The remedy requires only one extra fixed-effect control (and no real-estate price changes)—and turns the critical coefficient negative again. Real-estate price changes had played only the role of capturing the fixed-trends differential.

I Descriptive Statistics

[Insert Figure 1 here: **Time-Series (Not Cross-Section) of Aggregate Real-Estate Price Indexes and lagPPE-adjusted Capital Expenditures**]

Before the formal analysis, a graphical perspective about the time trends in firms with versus firms without real-estate is helpful. Unlike the formal analysis, these plots are not principally cross-sectional but time-series.

In the top-left graph of Figure 1, the dashed lines show that average real-estate prices mostly rose throughout the original CST sample. They only declined after their sample ended (from 2008 to 2012). The solid lines show that the capital-stock (P) normalized capital expenditures of firms (capex/P) without real estate started out higher but declined over time. In contrast, firms with real estate had roughly stable investments throughout the sample. This was not related to survivorship bias, as the dotted lines suggest.

The next two plots (TR and BL) show that changes in real-estate values did not line up well with changes in capex, the first plot presenting equal-weighted, the second value-weighted averages. The bottom right plot attempts to fit a relationship between changes in real-estate prices and RoC in investment. The Chaney, Sraer, and Thesmar (2012) hypothesis is that the blue slope should be steeper than the red slope. Instead, the linear line slopes are 0.08 for firms with real estate and 0.12 for firms without real estate. The 0.04 slope difference is not significant.

[Insert Table 2 here: **Descriptive Statistics for Common Variables**]

Table 2 shows descriptive statistics for the key variables used in the paper. The panel contains about 27,000 firm-years in total, about 3,000 firms over 20 years. Thus, most of the variation in level ratios is heterogeneity across firm types. Panel A (Row A1) shows that firms had capital stock (lagged property, plant, and equipment) of about $500 million on average, with a standard deviation of $2.5 billion (a long right tail). Firms without real estate were about 1-2 orders of magnitude smaller than firms with real-estate. For convenience, I shall often abbreviate this capital-stock variable either lagPPE or even more simply P.

Unfortunately, there are firms with very low P, which makes it difficult to assess means and standard deviations. Winsorizing ratios, both level and change ratios, is unavoidable. CST ratio variables remain winsorized as they were. The CoR and RoC variables are winsorized at –200% and +200%.[1]

Panel B shows that when firms owned real-estate, its value tended to be larger than their lagged property, plant, and equipment (B1). Firms made annual investments about 20-35% as large as their capital stocks (lagPPE), means-medians (B2). Firms without real-estate also had higher investment ratios. The inverse of lagPPE had more heterogeneity than redol/lagPPE and capex/lagPPE. Again, magnitudes are generally different for firms with vs. firms without real estate (B3).

[1]Winsorizations can distort the univariate impressions. However, the results in the panel regressions reported below are not materially changed by reasonable alternative winsorization thresholds (from ±1 to ±10).

Panel C is about a major focus of my paper, changes in variables. Implied real-estate value changes had modest variation (C1). Heterogeneity in other variables is more difficult to assess. In medians and interquartile ranges, the differences between CoRs and RoCs seem small (C2-C5). The volatility in real-estate values is smaller than that of capital expenditures. The volatility in percent changes in capital-stock seems larger than the volatility in both (C6).

II Numerator and Denominator-Caused Variation

The two important problems in our empirical data analysis are:

1. Unobservable but constant unit-specific effects that need to be neutralized.
2. Scale differences among units that need to be neutralized.

Scaling is practically a necessity in empirical work with firms that often differ by four or five orders of magnitude in size. Without scaling, just a few of the largest firms would often become too influential and effectively determine all slope coefficients.

A Scaling in Chaney, Sraer, and Thesmar (2012)

Chaney, Sraer, and Thesmar (2012) explain their choice of scaling in their footnote 9:

> ...normalization by PPE is standard in the investment literature (see, e.g., Kaplan and Zingales 1997 or Almeida, Campello, and Weisbach 2004). It provides typically a median investment ratio of 0.21. An alternative specification is to normalize all variables by lagged asset value (item No. 6), as in Rauh (2006) for instance, which delivers notably lower ratios.

Chaney, Sraer, and Thesmar (2012) note their scaling of other variables by PPE in all tables but do not discuss it further. They use this same scaling on both the dependent and independent variables in the first six out of eight specifications in Tables 5 and 6 and in all specifications in Tables 7, 8, and 9. In total, they use this shared-scaling approach in 29 of 33 specifications.

In Chaney, Sraer, and Thesmar (2012) as in many other papers, it is often the case that scaling is not of primary interest to the economic hypothesis. If the scaler itself also varies over time, it introduces undesirable contamination into firm-fixed-effect regressions. In addition to noise, scaling can also introduce bias, especially if there are time trends. In the case of Chaney, Sraer, and Thesmar (2012), lagged PP&E (P) not only varied considerably year by year but generally trended upwards, although not necessarily at the same rate as inflation or their other variables. Small growth firms without real estate tended to start out with higher but declining volatility and growth.

Volatility and trends in the denominator can induce spurious associations between variables, here real-estate value and investment. The bias can be particularly insidious when *both* the dependent variable (typically capex in CST12) *and* the independent variable (typically redol) are scaled with the same denominator. This will be examined next.

B Denominator-Induced Covariation in CST

[Insert Table 3 here: **Reproduction and Associations with 1/P Quasi-Constants**]

The most prominent headline finding of Chaney, Sraer, and Thesmar (2012) appears in their Section *Main Results*, Table 5, Column 1. Row A1 of my Table 3 reproduces their specification (as in eq. 3). Following the original, my panel regressions always also include the real-estate price index and all fixed year and firm effects. However, my tables do not report those coefficients. They report only the interesting coefficient, here (and mostly) the coefficient on real estate, redol/P. With the same specification and data, the results are identical. The coefficient is 0.074 with an OLS T-statistic of 19.09. The T-statistic drops to 8.98 when adjusted for two-way clustered standard errors.[2]

[2]The OLS T-statistics are different on the second digit after the decimal point due to a degree-of-freedom correction. The T-statistics based on clustered standard errors are typically much smaller than the OLS T-statistics, but they never shed doubt on the statistical significance of the reported coefficients in CST12. Nevertheless, simulations suggest that control for heteroskedasticity is important (in other cases).

The first indication that volatility in the denominator is a problem comes from a specification that mimicks their fixed-effects regression, but uses the normalized constant 1/P either as the dependent variable (A2) or as the independent variable (A3). These could be viewed *ad-absurdum* specifications in the sense that we would not want to interpret the numerator (the number 1) to influence investment, analogous to how CST12 interpret their numerator (redol) to influence investment. Instead, the coefficients provide measures of the components of variation that are associated with the denominator.

In Row A2, capex/P is regressed on 1/P. The statistical significance is even higher than it was for redol/P, with an OLS T-statistic of 45.96 (13.75 with cluster correction). In Row A3, the OLS T-statistic explaining 1/P with redol/P is 23.76 (8.62 with cluster correction).

These coefficient estimates raise doubt about the validity of the test design for interpreting coefficients as being only between numerators. Instead, they suggest that volatility in the denominator plays a role.

[Insert Table 4 here: **Adding 1/P as a Control Variable**]

Chaney, Sraer, and Thesmar (2020) respond with a control for the unwanted denominator-based covariation. They include an additive 1/P term into the regression. Table 4 confirms their finding that the coefficient on redol/P diminishes but remains comfortably large and statistically significant.

Although the next section will show that this control is insufficient, we agree that including the denominator as a regressor—both as a substitute and as a complement—is a useful diagnostic in research question concerning numerators when there are time-varying scaling denominators.

III Cor and RoC Ratios

Our goal remains the elimination not only of unobservable fixed effects but also scale volatility effects. Fortunately, there is a simple approach that does both. This approach eliminates the effect of denominator volatility from the variables themselves instead of relying on the OLS machinery to control for its effect (via an added $1/P$ as in CST20).[3]

A Definitions

It is well known that the level regressions with fixed effects

$$(\text{capex}/\text{P}) = b \cdot (\text{redol}/\text{P}) + \text{Firm Fixed Effects} + \ldots \tag{5}$$

is a close kin of the first-differences regression

$$\begin{array}{c}\underline{\textbf{Change of Ratio (CoR)}:}\\ \Delta(\,\text{capex}/\text{P}\,) = b \cdot \Delta(\,\text{redol}/\text{P}\,) + \ldots\,. \end{array} \tag{6}$$

Both are designed to control for unobservable but constant firm-specific effects. The two specifications even give the same numerical coefficients for two periods. The differenced specification is generally held to be more efficient than the fixed-effects specification when there is reason to believe that residual deviations are persistent. (It seems likely that this would be relevant in our application.)

From our perspective, the key advantage of the in-differences regression approach is that it makes it easy to contrast the results using changes in ratios with results using changes that are only then normalized. The latter kind of changes effectively removes the undesirable (co-)variation in the scaling denominator.[4]

$$\begin{array}{c}\underline{\textbf{Ratio of (over) Change (RoC)}:}\\ (\Delta\text{capex})\,/\text{P} = b \cdot (\Delta\text{redol})\,/\text{P} + \ldots\,. \end{array} \tag{7}$$

[3] It is possible to adapt a fixed-effects regression to reduce the effect of denominator volatility, too. However, this is both less intuitive and slower than the in-differences approach proposed here.

[4] For purposes of eliminating the effects of changes in the denominator, one could divide the difference not by the lagged P, but by the average PPE at time t and $t-1$.

Although this simple change of definition seems almost mundane, it has important consequences. Because the ROC variables use P as a fixed scaling scalar, year-to-year *changes* in 1/P no longer influence the variable in a regression. Only the P *levels* do. As a scaler, its year-to-year change is usually of lesser importance.

The use of RoCs is already common. For example, a rate of return is an RoC, not a CoR. Thus, any regression explaining rates of return is in effect an RoC regression. Yet, the advantages of RoC variables over CoR variables seem not to have been widely and appropriately appreciated.[5]

The CoR variable in (6) and RoC variable in (7) are nearly the same, except for a scale correction on the subtractive term:

$$\begin{aligned} \text{CoR:} \quad \Delta(\mathrm{v}/\mathrm{P}) \quad &\equiv \quad \frac{\mathrm{v}_t}{\mathrm{P}_t} \quad - \quad \frac{\mathrm{v}_{t-1}}{\mathrm{P}_{t-1}} \\ \text{RoC}: \quad (\Delta \mathrm{v})/\mathrm{P} \quad &\equiv \quad \frac{\mathrm{v}_t}{\mathrm{P}_t} \quad - \quad \frac{\mathrm{v}_{t-1}}{\mathrm{P}_{t-1}} \quad \times \quad \overbrace{\left(\frac{\mathrm{P}_{t-1}}{\mathrm{P}_t}\right)}^{\text{adjustment factor}} \quad . \end{aligned} \tag{8}$$

The adjustment factor removes the effects of (undesirable) volatility in the denominator $(1/\mathrm{P}_t - 1/\mathrm{P}_{t-1})$, keeping only the (desirable) time variation in the numerator (v). The CoR and RoC variables are different by a factor

$$\Delta(\mathrm{v}/\mathrm{P}) - (\Delta \mathrm{v})/\mathrm{P} = (v_t \cdot P_t)/P_{t-1}^2 \ . \tag{9}$$

If there is no variance in the growth rate of the denominator or no covariance between v_t and P_t, the CoR and RoC variables are the same.

B Assessing CoR Slope Effects

Although the effect of variation in the denominator on individual variables is obvious, the effect on the slope between them is not. Consider three consecutive firm-years that fix two data points in a difference regression. The slope (without an intercept) between

[5] An informal survey of the literature suggests that the CoR approach is more common. An informal verbal survey of empiricists by myself furthermore suggested that most authors could not even remember whether their own published papers had worked with CoRs or RoCs. They had to go back and check first. Chaney, Sraer, and Thesmar (2012) are just one example where the standard CoR approach probably did not merit a second thought.

two CoR sample points is $\hat{b} \equiv (\text{capex}_t/\text{P}_t - \text{capex}_{t-1}/\text{P}_{t-1})/(\text{redol}_t/\text{P}_t - \text{redol}_{t-1}/\text{P}_{t-1}) = \Delta(\text{capex}/\text{P})/\Delta(\text{redol}/\text{P})$. How is this two-point slope affected by volatility in its three components?

Table 2 showed that the volatility of capex changes (sd of 0.4) is higher than the volatility of real-estate changes (0.2-0.3), and both are less than the volatility of the 1/P changes (0.64). The low volatility of real-estate changes is not surprising. They represent mostly year-to-year area-wide real-estate price changes.

[Insert Table 5 here: **Sign Changes of CoR and RoC Variables**]

Table 5 shows how sign changes in the investment and real-estate CoR variables come about. Panel A shows that in about 86.5% of firm-years, the CoR Δ(capex/P) shares the sign with changes in its implied numerator capex. When capital expenditures increase, the CoR measure typically but not always increases.

Such shared direction in signs is less common for real-estate. It is only in 32.0% of firm-years that the sign of the CoR Δ(redol/P) is positive when the real-estate collateral increased, and vice-versa. In 24.6% of the cases, the two moved in opposite directions. An RoC regression input could therefore measure a potentially large loss in collateral (a negative value) despite an increase in the value of the underlying real-estate. This is mathematically sensible, because the *relative* collateral declined. Yet this is not necessarily an intuitive notion of the effect of real-estate collateral on investment.

Panel B shows that a linear decomposition suggests that while most of the sign changes of the CoR Δ(capex/P) can be attributed to its numerator, the CoR Δ(redol/P) variable is nearly unrelated to its numerator. Most of its sign changes are attributed to sign changes in its denominator (1/P).

Returning to the slope $\hat{b}$, small changes in the CoR denominator (Δ(redol/P)) can have large effects on $\hat{b}$, especially when Δ(redol/P)'s values are small and near zero. This is often the case here. The median Δ(redol/P) is exactly zero and the interquartile range is –0.0012 to +0.0178. This renders their estimated $\hat{b}$ slopes in CoR difference regressions (or nealy equivalently fixed-effects level regressions) less stable. The regression standard errors provide some relief by effectively deemphasizing such near-zero large sample points in $\Delta($redol/P$)$, at least to the extent that they fit the standard OLS assumptions and heteroskedasticity (and potentially 1/P) are adequately controlled for.

IV Reanalysis of CST Results

[Insert Table 6 here: **CoR and RoC Regressions**]

Table 6 proceeds to in-differences regressions with various combinations of CoR and RoC investment and real-estate variables. The regressions always include year fixed effects and *changes* in the real-estate price index (repi), but not firm fixed effects.[6]

Row A shows that the CoR-CoR specification (i.e., both $\Delta($capex/P$)$ and $\Delta($redol/P$)$) not only retains the strong positive coefficient of the fixed-effects level regression from Table 3, but increases it. The clustered T-statistic remains almost the same (8.30 instead of 8.98).

Row B shows that the positive coefficient remains but is only half as strong if only the dependent capital-expense variable is defined in RoC terms.

Row C shows that the coefficient becomes practically zero once real-estate is defined in ROC terms.

Row D shows that the coefficient turns negative (–0.021) and becomes insignificant when we remove the denominator volatility in both capital expenditures and real estate.[7] The same inference holds in year-by-year regressions (see Appendix Table $\mathscr{B}$.1).

[6]The regressions do not include a firm scale control, such as 1/P, or firm fixed effects (which would reintroduce denominator volatility), but this turns out to matter little. When added, the coefficients remain similar.

[7]The change variables were winsorized at –200% and +200%. The equivalents are –0.039 when not winsorized, –0.023 when winsorized at –300% and +300%, and –0.030 when winsorized at –100% and +100%.

This evidence suggests that the positive CST12 coefficient of 0.074 in the fixed-effects regression was primarily due to the (mechanical) effect of shared changes in the two variables' capital-stock denominators.[8]

A Performance Analysis of Scale Control: Time-Series

There are now two approaches to the scale problem: one is the control for 1/P in the fixed-effects regression proposed in Chaney, Sraer, and Thesmar (2020); the other is the RoC approach proposed here. The former employs the estimation machinery of least-squares estimator, the latter employs direct neutralization in the input variables. How do they perform? And can we reconcile the fact that the former 1/P approach suggests a strong positive association while the latter RoC approach does not?

There are too many parameters in the actual Chaney, Sraer, and Thesmar (2012) setting (and no closed-form solutions) to make it easy to gain intuition about the differences between the two approaches. Thus, we need to rely on examining the effects in placebos.

It is useful first to construct a minimalist example. Table 7 contains the results of simulating a time-series with a (degenerate) panel of one firm and no association between x and y. The data-generating processes are

$$\tilde{x}_t = \pi_t \cdot (A + \tilde{r}_t) \ , \qquad \tilde{y}_t = \pi_t \cdot (A + \tilde{s}_t) \ , \qquad \tilde{d}_t = \pi_t \cdot \exp(\tilde{e}/D) \ , \tag{10}$$

where $\tilde{r}_t$ and $\tilde{s}_t$ are iid unit-standard random walks, $\tilde{e}$ is a unit-normal iid draw, $\pi_t \equiv (1+2\%)^t$ is an inflation-like trend, and A and D are large constants (set to $A{=}T{=}20, D{=}10$). The ingredients were carefully chosen: [1] A ensures that x and y both remain almost always in the positive quadrant; [2] π_t induces a modest shared trend link between all three variables; and [3] D allows for a (modest but not overwhelming) "zig-zag" time-series

[8] Appendix B discusses an earlier placebo from my previous draft in more detail. It used constant scaling (e.g., incept-year P) in order to avoid shared covariation. Chaney, Sraer, and Thesmar (2020) correctly point out that this makes for a weaker test from the perspective of their theory. My current manuscript introduces the much improved RoC alternative. However, their perspective seems too lax in deemphasizing Type-1 error in favor of lowering Type-2 error. In the extreme, under the theory, the null hypothesis is false and the theory is true.

volatility in the denominator.[9] The 25-line R program is in Table $\mathscr{D}$.1 and can clarify any remaining textual ambiguities—or be used to explore different parameters.

[Insert Table 7 here: **Univariate Time-Series Placebo**]

Table 7 shows the results for 100,000 draws of these degenerate panels, i.e., time-series of 20 periods for one firm. Panel A confirms that the generated regression inputs seem reasonable. The simulated x and y tend to wander between about 20 and 30 (with typical within single-draw spreads of 3.5), never beyond 10 and 40; d averages are about 1.25 with typical within-draw spreads of 0.190, and $1/d$ averages are about 0.8 with typical within-draw spreads of 0.125.

The fixed-effects regression "FE" here is just the plain OLS regression with intercept, explaining y/d with x/d. Row B1 shows that the OLS beta has a positively biased mean of 0.58. The average T-statistic of 3.8 is also too confident (having a standard deviation of about 2.8 instead of 1.0). Row B2 shows that control for $1/d$ remedies this only modestly. The coefficient is still positively biased with its mean of 0.47, though now even more confident (standard deviation of 3.39).

Row B3 shows that the CoR specification on both x and y has even worse bias, with its mean of 0.89 and T-statistic of about 8.3.

Rows B4-B6 show that using an RoC ratio in at least one of the two variables eliminates the shared influence of the denominator. This gives the estimated coefficients and T-statistics good behavior, rendering the coefficients nearly unbiased and the T-statistics standard deviations nearly one.

[9] It is possible to gain some intuition about differential biases by comparing $\left\{\operatorname{cov}\left(\frac{y_{t-1}}{d_{t-1}}\frac{d_{t-1}}{d_t}\ ,\ \frac{x_{t-1}}{d_{t-1}}\frac{d_{t-1}}{d_t}\right)\right\}$ and $\left\{\operatorname{cov}\left(\frac{y_{t-1}}{d_{t-1}}\ ,\ \frac{x_{t-1}}{d_{t-1}}\right)\right\}$ under admittedly invalid assumptions. This suggests that the bias is affected not only by the variance of d_t, but also by keeping x and y in the same quadrant. If $\mathrm{E}\left(\frac{x_{t-1}}{d_{t-1}}\right)\cdot\mathrm{E}\left(\frac{y_{t-1}}{d_{t-1}}\right)\cdot\operatorname{var}\left(\frac{d_{t-1}}{d_t}\right)$ is 0, then CoR-based estimates are as unbiased as RoC-based estimates.

Not shown, experimentation suggests that the positive x and y means and the time-series volatility in the denominator are important for the inferior performance of the CoR and uncontrolled FE specifications. The π trend is responsible for the poor performance of the controlled 1/P regression. Removing the π trend renders B2 a good estimator.

B Performance Analysis of Scale Control: Full CST-Like Panel

Unlike the minimalist time-series example, it is too difficult to understand the full panel. There are effects of initial conditions, trends, and positivity of both capex/P and redol/P, and a dichotomy between firms with vs. firms without real-estate and winsorization. Thus, we now consider a second placebo, which is designed to offer a somewhat more realistic panel process that is closer to CST12. The generating process matches some principal features of their data. Although the distributional moments were chosen to be similar to their empirical counterparts, the statistical distributions themselves were not. Instead, normal distributions were chosen for simplicity of illustration. In brief:

- Each of 100,000 Monte-Carlo draws is for 3,000 firms and 15 years. (The following description applies individually to each firm.)

- The first (dollar) lagged PP&E (P_0) is an exponential function of a random Gaussian normal with mean 3 and standard deviation 4. P_t then grows by a random rate drawn from a normal with a mean of 3% and a standard deviation of 20%.

$$\begin{aligned} \mathrm{P}_0 &= \exp[\,\mathcal{N}(\mu{=}3.0, \sigma{=}4.0)\,]\,, \\ \mathrm{P}_t &= \mathrm{P}_{t-1} \cdot \big[1 + \mathcal{N}(\mu{=}0.03, \sigma{=}0.20)\big]\,, \end{aligned} \tag{11}$$

  where $\mathcal{N}(\mu, \sigma)$ is a draw from a standard normal distribution.

- The first (dollar) corporate investment (capex_0) is a product of the first P_0 with a random normal, with mean 21% and standard deviation 40%. capex_t then grows by a random rate *of P* drawn from a normal with a mean of 1% and a standard deviation of 16%.

$$\begin{aligned} \mathrm{capex}_0 &= \mathrm{P}_0 \cdot \mathcal{N}(\mu{=}0.21, \sigma{=}0.40)\;, \\ \mathrm{capex}_t &= \mathrm{capex}_{t-1} + \big[\mathrm{P}_{t-1} \cdot \mathcal{N}(\mu{=}0.01, \sigma{=}0.16)\big]\,. \end{aligned} \tag{12}$$

capex is not winsorized and can turn negative. Note that this equation links firm scale with capital expenditures.

- The relevant repi series appreciates independently for each firm with a growth rate of 5% and standard deviation of 5%.

$$\begin{aligned} \text{repi}_0 &= 1.0 \ , \\ \text{repi}_t &= \text{repi}_{t-1} \cdot \left[1 + \mathcal{N}(\mu{=}0.05, \sigma{=}0.05) \right] \ , \end{aligned} \tag{13}$$

The first (dollar) real-estate holding (redol_0) is 0 for 40% of the firms. For the remaining 60%, it is a random normal with mean 100% and standard deviation 150%, again multiplied by P_0. redol_t then grows with repi.

$$\begin{aligned} \text{redol}_0 &= \text{P}_0 \cdot \mathcal{N}(\mu{=}1.0, \sigma{=}1.5) \cdot \mathcal{B}(60\%) \ . \\ \text{redol}_t &= \text{redol}_0 \cdot \text{repi}_t \ , \end{aligned} \tag{14}$$

where $\mathcal{B}(p{=}60\%)$ is a draw from a Bernoulli distribution with probability $p{=}60\%$ of 1, and probability $1{-}p$ of 0. Note that real-estate values evolve without consideration of firm scale.

Winsorizations in each period prevent P and repi from turning negative. By design, there is no correlation between initial level real-estate holdings and initial capital expenditures, except that induced by the common scale at incept time. Larger firms have higher capex and higher redol than smaller firms, but they do not have higher ratios. Changes are independent, except for the scaling link between capex and lagged firm size.

Once generated, the dollar quantities of capex and redol are transformed into ratios for purposes of testing empirical specifications. In line with the empirical winsorizations, capex/P is winsorized to lie between –150% and +200%. redol/P is winsorized to lie between 0 and 5.9 times. 1/P is winsorized to lie between 0 and 6. The R code consumes 1.5 pages.

Although more realistic than the simple time-series example, this simulation still favors simplicity over realism. It remains a test-bed for illustration. It does remain a true placebo in the sense that the experiment does not generate any association between redol and capex, except that induced by their common scale. And it remains simple enough to be easy to understand.[10]

[Insert Table 8 here: **Panel Regression Placebo, T-statistics**]

Panel A of Table 8 shows the distribution of firm-clustered T-statistics of the fixed-effects specifications, as in the original CST12 panel regression with redol/P as the sole regressor. The average clustered T-statistic here is 8.27—surprisingly close to that in the data. There were no draws in which the T-statistic did not suggest strong positive statistical significance. Panel B shows the same statistics for 1/P as the sole regressor. Again, the average coefficient estimate is not 0, but 7.26. (Not shown, it is smaller when not winsorized.) Most importantly, Panel C shows both T-statistics when both redol/P and 1/P regressors are included. The inclusion of 1/P moderates the bias on redol/P (from a typical T-statistic of about 8.27 down to 7.02) but it does not eliminate it. The mechanical correlation continues nearly unabated.

Panel D shows the performance of difference specifications. The CoR approach again suffers from strong positive bias, with a mean (year-clustered) T-statistic of 17.86. Using RoC variants either for the dependent or independent variable greatly reduces but does not eliminate the bias or overconfidence (to 2.66 and 0.88, respectively). Using RoC variants for both investment and real-estate offers the least noisy and best estimates (with an average T-statistic of 0.15).

C Least-Squares Control?

The placebos suggest that the RoC approach is superior. Although there may be cases in which control for 1/P suffices, the suggestion in Chaney, Sraer, and Thesmar (2020)—that control for 1/P suffices in their real-estate and corporate investment application—seems misplaced.

[10] Experimentation suggests that heterogeneity in firm-scale is important. More experimentation suggests that other placebo processes can produce similar results.

There is a related important advantage of the RoC specification approach. It is largely immune to *any* misspecification in the denominator. It has removed its volatility from consideration altogether. In contrast, the 1/P control relies on the *exact* specification of the functional form of 1/P over the entire domain to allow the least-squares estimator to neutralize it. If the dependent or independent CoR ratios contain non-linearities in P, the 1/P control would remain affected, but the RoC approach would not.

[Insert Table 9 here: **Adding log(1/P) as a Control Variable**]

Table 9 shows that this problem applies to Chaney, Sraer, and Thesmar (2012). Rows A and B repeat Table 4. Rows C and D show that including the logarithm of P negates the sign of the redol/P coefficient.[11] It seems that the introduction of the time-varying 1/P denominator into the equation had also inadvertently introduced non-linearities (especially but not only into the dependent variable). These were then picked up by the real-estate variable with its shared time-varying denominator. Rows E and F repeat the final regression from Table 6 with and without a few more P related controls. They show that the RoC approach remains unaffected by these non-linearities, never having allowed 1/P volatility into the estimations in the first place.

[11] Dividing the sample into 10 groups based on size (the 1/P ratio), either equally or using exp(−7) to exp(+1), also negates the sign of redol/P.

V The REisPOS$*$repi Independent Variable

In four specifications in Chaney, Sraer, and Thesmar (2012, Tables 5&6), lagged capital stock normalizes only the dependent investment variable. The independent variable is a dummy for real-estate ownership (REisPOS) multiplied by the real-estate price index (repi).

$$\text{capex/P} = \theta_1 \cdot (\text{REisPOS*repi}_t) + \theta_2 \cdot (\text{repi}_t) + \dots \; . \tag{15}$$

By including repi and firm fixed effects at the same time, the coefficient effectively assesses whether firms with real-estate holdings had a different response coefficient to real-estate price index (repi) changes.

A Reanalysis of CST Results

These specifications do not suffer from the mechanical correlation induced by the denominator discussed in Section IV. Nevertheless, the dependent variable continues to be an amalgam of the capex numerator and the P denominator, potentially affected by noise and trends in P.

[Insert Table 10 here: **CST12 Table 5, Column 7 Analog: REisPOS*repi.offc as Independent Variable**]

Table 10 Panel A replicate the original CST12 regression and then mimick the 1/P regression from Table 3 and the regression with 1/P and log(P) control from Table 9. Row A1 shows that the REisPOS*repi variable is also strongly correlated with 1/P. Again, if anything, the clustered T-statistic of 3.30 is higher when 1/P is the dependent variable than when capex/P is the dependent variable (T_{clstr}=2.27). Although the REisPOS*repi variable does not mechanically contain 1/P, it still explains some of the variation in the denominator rather than the numerator of the dependent variable. Interpreting the real-estate coefficient to relate only to the numerator capex and not the denominator 1/P would seem incorrect. Row A2 shows that including $\log(1/P)$ renders REisPOS*repi insignificant (T_{clstr}=0.72).

Panel B shows that the positive coefficient in the fixed-effects specification does not survive either in a CoR or in an RoC regression. The coefficient estimate in the CoR specification is –0.046 with a year-clustered T-statistic of –0.22. In the RoC specification, the coefficient estimate is –0.348 with a T-statistic of –1.69. Table $\mathscr{B}$.1 shows that there is no pattern in year-by-year regressions, either.

VI Extended Evidence Through the Great Recession

The Chaney, Sraer, and Thesmar (2012) sample ended in 2007. Thus, they could not have included the *Financial Crisis of 2008* and its aftermath. In December 2019, I updated the Home Price Index series from FHFA up to 2018.[12,13] Figure 1 already used this data.

[Insert Table 11 here: **Evidence Extended to 2018**]

Table 11 repeats the analyses of earlier sections within this extended sample. Panel A shows that the real-estate coefficient from the CST fixed-effects base specification remains nearly the same. Panel B shows that 1/P still works nearly as well as a predictor and that including the logarithm of 1/P still reverses the sign of the key variable redol/P. Row B6 contains a modest surprise. REisPOS*repi now turns positive and statistically significant. Appendix $\mathscr{C}$ shows that this turns out to be related to shared trends. Firms with real-estate were different from firms without real-estate. Including year dummies interacted with a dummy of whether a firm did or did not start with real-estate (i.e., not using any real-estate pricing information) reverses the sign of this coefficient.

[12] My reanalysis of the CST12 data was possible only for the state and not the MSA real-estate index. The FHFA has changed the set of largest MSA between the original data and now. The FHFA also revises its series ex-post, even many years later, after repeat house sales have occurred, thus making it impossible to exactly match old data with new data. However, all rebuilt variables had more than 99.5% correlations with the original CST12 variables where they overlapped. An earlier draft shows that a switch from repi.off to repi.state is innocuous (increasing the coefficient from 0.213 to 0.242 in the 2008 sample) in the REisPOS*repi regression, as is the removal of the additional controls (increasing the coefficient from 0.242 to 0.345). The extended sample has practically the same coefficient estimate, 0.346.

[13] An appropriate objection to adding more years at the end of the CST12 sample is that the number of firms halves around every 10 years (e.g., from about 2,792 in 1993 to 1,298 in 2003). This could push the evidence from a reasonable 15 years to an unreasonable 25 years. Yet, similar results also obtain with shorter modestly extended samples, which still include parts or all the Great Recession and recovery.

Panel C shows the extended RoC regressions. Again, the positive association between real-estate and investment disappears.

VII Conclusion

The theoretical conjecture that real-estate prices could influence capital expenditures through collateral remains appealing. But research is never perfect. My paper does not claim that it is impossible that future research may reinvestigate this conjecture and find support with better measures—especially real-estate holdings and price innovation measures (though Campello et al. (2020) finds even less support), controls, and specifications. My paper only claims that the tests in Chaney, Sraer, and Thesmar (2012) and Chaney, Sraer, and Thesmar (2020) have not provided adequate empirical support for this hypothesis. Their tests seem uninformative. Their inference was based on covariation attributable to volatility in their PP&E scaling denominator (P) and not on covariation between their two numerators, real estate (redol or REisPOS) and corporate investment (capex). Tests with improved scaling control using ratios of changes (RoCs) suggest no positive association between real-estate shocks and corporate investment either in their sample or in an extended sample reaching into the Great Recession.

More generally, the problem of volatility in a scaling variable and its potential remedy could be of interest in other work, too. I conclude with some thoughts of the potential relevance. If a theory is primarily interested in the association between numerators and if it uses denominators primarily as (nuisance) scale controls, the RoC approach is superior. It may seem mundane, but it effortlessly reduces measurement noise and biases, including that induced by non-linear scaling effects.

If the theory is specific in *both* the numerator *and* the denominator and/or prescribes a very specific denominator that *must* be allowed to time-vary, it may be the case that this interest applies only either to the dependent or independent variable. Neutralizing the effect of the denominator in the unaffected variable via an RoC ratio can still provide better estimation properties than either a level regressions with fixed effects or a CoR-CoR difference regression. For example, ratios that divide two stock variables (as in a book value and a market value) may be less flexible than ratios that divide a flow variable by a stock variable (as in a capital-expenditure ratio).

If the theory is adamant that both numerator and denominator must be as specified and denominator volatility must be allowed to enter the estimation—i.e., they are intricate components of the predictions of the theory—then the usefulness of RoC approach has its limits. Of course, empirical tests of such theories—in which even arbitrary volatility in the scaling denominator could create similar associations—would be quite weak. In this case, the RoC approach could still be used as a diagnostic. It can inform whether variation in the numerator and the denominator are both important or whether the variation works through the denominator alone.

I recommend that readers also consult Chaney, Sraer, and Thesmar (2020) (a response to my first [and worse] draft of January 2020). The version here is however much different and thus obsoletes much of my earlier draft and thus their response. The temporary appendix retains some of the original material for this reason.

References

Almeida, Heitor, Murillo Campello, and Michael S. Weisbach (2004). "The cash flow sensitivity of cash". In: *The Journal of Finance* 59.4, pp. 1777–1804.

Angrist, Joshua D. and Jorn-Steffen Pischke (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Baker, Malcolm, Jeremy C Stein, and Jeffrey Wurgler (2003). "When does the market matter? Stock prices and the investment of equity-dependent firms". In: *The Quarterly Journal of Economics* 118.3, pp. 969–1005.

Campello, Murillo, Robert A. Connolly, Gaurav Kankanhalli, and Eva Steiner (May 2020). *Does Real Estate Value Boost Corporate Borrowing? Evidence from Contract-Level Data*. Tech. rep. Cornell University and University of North Carolina.

Chaney, Thomas, David Sraer, and David Thesmar (2012). "The collateral channel: How real estate shocks affect corporate investment". In: *American Economic Review* 102.6, pp. 2381–2409.

— (May 2020). *Real Estate Collateral Does Affect Corporate Investment. Response to Welch (2020)*.

Davidoff, Thomas (2016). "Supply constraints are not valid instrumental variables for home prices because they are correlated with many demand factors". In: *Critical Finance Review* 5.2, pp. 177–206.

Fazzari, Steven M, R Glenn Hubbard, and Bruce C Petersen (2000). "Investment-cash flow sensitivities are useful: A comment on Kaplan and Zingales". In: *The Quarterly Journal of Economics* 115.2, pp. 695–705.

Grieder, Timothy and Hashmat Khan (Dec. 2017). *Do Real Estate Shocks Affect Corporate Investment?* Tech. rep. Carleton University.

Kaplan, Steven N and Luigi Zingales (1997). "Do investment-cash flow sensitivities provide useful measures of financing constraints?" In: *The quarterly journal of economics* 112.1, pp. 169–215.

Mundlak, Yair (Jan. 1978). "On the Pooling of Time Series and Cross Section Data". In: *Econometrica* 46.1.

Rauh, Joshua D (2006). "Investment and financing constraints: Evidence from the funding of corporate pension plans". In: *The Journal of Finance* 61.1, pp. 33–71.

Stambaugh, Robert (1999). "Predictive Regressions". In: *Journal of Financial Economics* 54, pp. 375–421.

Tables and Figures

Table 1: Glossary of Variable Names Used Repeatedly

Real-Estate Variables

Variable	Chaney, Sraer, and Thesmar (2012) Eqn (1)	Data Set	Description
repi.state	P_t / State R. Prices	index_state	Real-Estate Price Index, Home Prices, State
repi.off	P_t / MSA O. Prices	offprice	Real-Estate Price Index, Offices and Commercial Real-Estate Prices, MSA
redol			Real-Estate Holdings in Dollars by the Firm, calculated as initial value grossed up by repi.
REisPOS		REAL_ESTATE0	Dummy whether firm owns real estate. Unchanging for each firm over the sample.

Corporate Investment and Capital Stock Variables (Compustat)

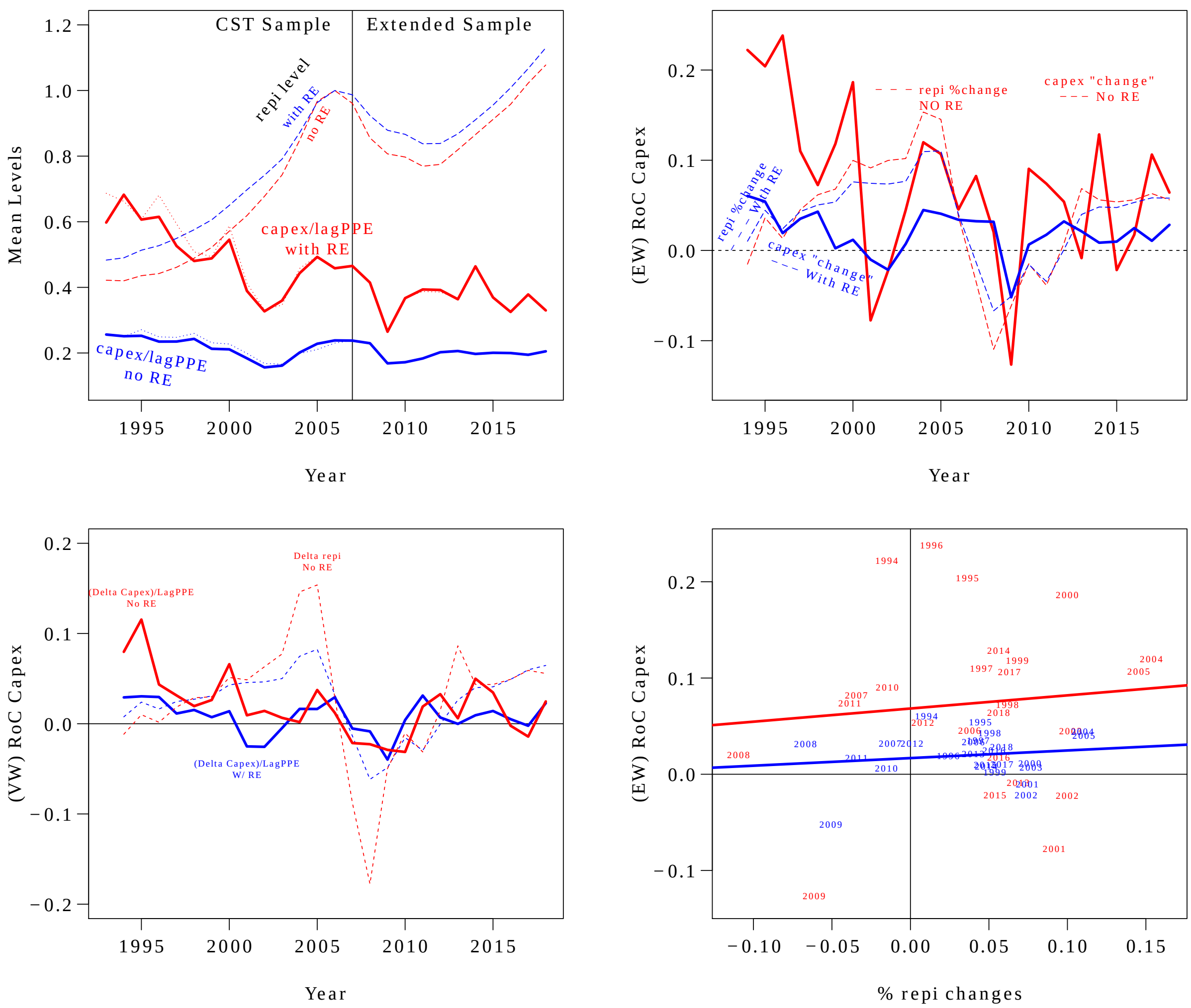
Variable	Eqn (1)	Data Set	Description
capex			Capital Expenditures, Compustat 'capx'
lagPPE ≡ P		PPEm	One-Year Lagged PP&E (net), Compustat 'ppent'
1/lagPPE ≡ /P			1/P, winsorized at 6.0.

Regression Inputs

Variable	Eqn (1)	Data Set	Description
capex/P	INV_{it}	inv	Dependent Variable, capex/P.
redol/P	REValue_{it}	RE_value	Independent Variable, CST12 Table 5 Column 1
REisPOS*repi.offc	RE_OWNER × MSA O(ffice) Prices		Independent Variable, CST12 Table 5 Column 7
REisPOS*repi.state			Incremental response coefficient to real-estate prices for firms with real-estate holdings
REisPOS*yrtrend			Differential (by REisPOS) but deterministic year trends, defined as REisPos × (Year-1992)/1000
REisPOS*years			Complete set of differential (by REisPOS) year effects (cross-dummies)

Explanation: CST graciously provided their data, which begins in 1992 and ends (mostly) in 2007. The original real-estate data vendor was *Global Real Analytics*. The original Chaney, Sraer, and Thesmar (2012) provides more details about their variables. My paper uses different naming conventions.

Figure 1: Time-Series (Not Cross-Section) of Aggregate Real-Estate Price Indexes and lagPPE-adjusted Capital Expenditures



Explanation: The top left plot (TL) shows the time-series behavior of 25 years of equal-weighted averaged real-estate price indexes (repi.state, in solid lines) and of the averaged capex/P (in dashed lines). Blue lines are (typically older bigger) firms with real estate, red lines are (typically younger smaller) firms without real estate. (The dotted lines are firms that survived until 2008.) The top right plot (TR) shows the same in averaged changes, where changes for investment are $(\Delta\text{capex})\ /\text{P}$. The bottom left plot (BL) shows a value-weighted version. The bottom right plot (BR) shows the (equal-weighted) associations of repi.state changes and capex changes, then normalized by lagPPE.

Interpretation: TL: repi.state declined from 2007 to 2012. The stark real-estate price drop in 2008 was followed by a stark capex/P drop only with two years delay. TR and BL: The fat blue line (with RE) does not respond more to changes in the blue dashed line than the equivalent red lines (without RE). However, the fat red line starts out higher than the fat blue line and converges towards it over time. BR: The red association is not less than the blue association.

Source: 3plots.R: (timeplotlvl.pdf, timeplotcht.pdf, fig3-agg.pdf, xy.pdf).. July 16, 2020

Table 2: Descriptive Statistics for Common Variables

Panel A: Base Variables

		N	Mean	SD	All Firms 0%	25%	50%	75%	100%	NO RE Mean	NO RE SD	WITH RE Mean	WITH RE SD
A1.	lagPPE (≡P), in-m$	27,201	488	2,538	0.001*	2.7	20.5	160.7	88,440	44.0	784	807	3,221
A2.	log(P)	27,201	3.0	3.0	–6.9*	1.0	3.0	5.1	11.4	0.8	2.2	4.6	2.2
A3.	redol, in-m$	27,009	563	3,701	z	z	8.1	123.4	140,960	z	z	973	4,821
A4.	capex, in-m$	27,201	99	713	–35.6	0.6	4.7	31.8	33,143	12.2	157.6	161	920

Panel B: Level Ratio Variables

		N	Mean	SD	All Firms 0%	25%	50%	75%	100%	NO RE Mean	NO RE SD	WITH RE Mean	WITH RE SD
B1.	redol/P	27,201	0.86	1.36	z	z	0.29	1.12	5.89*	z	z	1.48	1.51
B2.	capex/P	27,201	0.35	0.40	–1.50	0.11	0.21	0.41	1.95*	0.52	0.52	0.22	0.22
B3.	one/P	27,201	0.66	1.48	0.00	0.01	0.05	0.38	6.00*	1.42	1.97	0.11	0.48

Panel C: Changes (PP&E-normalized Ratios)

		N	Mean	SD	All Firms 0%	25%	50%	75%	100%	NO RE Mean	NO RE SD	WITH RE Mean	WITH RE SD
C1.	Δrepi.state	24,177	0.033	0.03	–0.05	0.013	0.027	0.045	0.58	0.033	0.038	0.034	0.030
C2.	CoR Δ(redol/P)	24,177	0.017	0.29	–2.00*	–0.001	z	0.018	2.00*	z	z	0.03	0.38
C3.	RoC (Δredol) /P	24,003	0.058	0.20	–1.79	z	0.004	0.044	2.00*	z	z	0.10	0.26
C4.	CoR Δ(capex/P)	24,177	–0.017	0.42	–2.00*	–0.095	–0.001	0.077	1.95*	–0.03	0.60	–0.01	0.22
C5.	RoC (Δcapex) /P	24,177	0.070	0.38	–2.00*	–0.058	0.010	0.102	2.00*	0.13	0.54	0.03	0.20
C6.	RoC P/ppe -1	24,177	0.208	0.82	–0.86*	–0.075	0.034	0.217	6.39*	0.36	1.13	0.10	0.47
C7.	log(P/ppe)	24,177	0.072	0.45	–2.00*	–0.078	0.034	0.196	2.00*	0.11	0.60	0.05	0.30
C8.	CoR Δ(1/P)	24,177	–0.026	0.64	–5.99*	–0.007	0.000	0.001	5.99*	–0.07	0.93	0.01	0.29

Explanation: The sample were publicly traded firms on Compustat from 1992 to 2007 with available lagPPE, capex, and REisPOS. 'z' indicates the masspoint of firms without real estate. Raised stars indicate winsorized values. Grieder and Khan (2017) also replicate the Chaney, Sraer, and Thesmar (2012) results and find that winsorization can increase their point estimates. (Winsorization in real estate would tend to bias coefficient estimates upward, while winsorization in capital expenditures would tend to bias them downward.) Not shown, 1,645 firms owned real estate, 1,359 firms did not. Of firms with real estate, about half held real-estate worth more than their property, plant, and equipment.

Source: 1descriptive.Rout, Aug 1, 2020.

Table 3: Reproduction and Associations with 1/P Quasi-Constants

			On (key) x			
	y Dep ← Indep (key) x + ···	Coef	T_{ols}	T_{clstr}	Obs	Firms/Yrs
A1.	capex/P ← redol/P + ···	0.074	19.09	8.98	27,201	3,024/15
A2.	capex/P ← 1/P + ···	0.127	45.96	13.75	27,201	3,024/15
A3.	1/P ← redol/P + ···	0.204	23.76	8.62	27,543	3,033/15

Explanation: This table reports the key coefficients and T-statistics from fixed-effects regressions that are analogous to those in Chaney, Sraer, and Thesmar (2012, Table 5 Column 1). The regressions include all firm and year fixed effects and a real-estate price index (repi.state), but do not report their coefficients. The variables are described in Table 1, descriptive statistics are in Table 2. capex are capital expenditures, redol is real-estate dollar value, P is (time-varying) lagged net property, plant, and equipment. Row A1 shows the (perfect) reproduction. A2 and A3 replace the independent and dependent variable with 1/P, respectively. T_{ols} is a plain OLS T-statistic. T_{clstr} is based on clustered standard errors according to the fixed-effects—here two-ways.

Interpretation: A2 and A3 suggest that variation in the denominator contributes generously to the reported coefficient explaining capital expenditures with real estate (as in A1).

Source: 2feregs-2007.R. July 14, 2020

Table 4: Adding 1/P as a Control Variable

		on redol/P			on 1/P		
		Coef1	T_{ols}	T_{clstr}	Coef	T_{ols}	T_{clstr}
A.	CST12 Regression	0.074	19.15	9.01			
B.	CST20 Regression	0.049	13.06	6.91	0.122	43.79	13.07

Explanation: The data panel contains 27,201 observations for 3,025 firms over 15 years. The dependent variable is capex/P. The specifications include but do not report the coefficients on repi and on the fixed firm and year effects. The CST20 regression first appeared in Chaney, Sraer, and Thesmar (2020).

Interpretation: This table confirms Chaney, Sraer, and Thesmar (2020). The additive control alone does not greatly diminish the capex/P effect.

Source: 2feregs-2007.R. July 28, 2020

Table 5: Sign Changes of CoR and RoC Variables

Panel A: Implied Sign Retentions and Reversals

(Δv) (RoC)	CoR: Same Signs	CoR: Opposite Signs	(Zero Involved)
capex	86.5%	12.4%	1.1%
redol	32.0%	24.6%	43.4%

Panel B: Linear Regression Decomposition

Estimated Linear Regressions							R^2
sign(cor.capex)	=	–0.05	+	0.74· sign(Δcapex)	–	0.25· sign(ΔlagPPE)	63.0%
	=	–0.10	+	0.76· sign(Δcapex)			56.7%
	=	0.03	+		–	0.30· sign(ΔlagPPE)	8.6%
sign(cor.redol)	=	0.03	+	0.17· sign(Δredol)	–	0.40· sign(ΔlagPPE)	29.9%
	=	–0.02	+	0.15· sign(Δredol)			1.4%
	=	0.12	+		–	0.40· sign(ΔlagPPE)	28.1%

Explanation: Panel A shows the frequency with which "changes of ratios" (CoRs) reverse the sign of changes in their implied numerator. (The "ratio of changes" [RoCs] always shares the sign with its numerator.) In Panel B, the sign function yields –1 if the sign is negative, 0 if zero, and +1 if positive. The panels are based on 24,000 firm-years with complete data.

Interpretation: The sign of the real-estate CoR ratio Δ(redol/P) is almost unrelated to whether the underlying real-estate appreciated or depreciated (Δrepi or Δ redol). Instead, the sign is primarily determined by whether the firm's lagged PP&E appreciated or depreciated.

Source: 1descriptives.R. July 31, 2020

Table 6: CoR and RoC Regressions

Variables in Changes	capex	redol	on redol FD Coef	T_{ols}	T_{clst}	Years	Firms
A.	CoR	CoR	0.199	21.18	8.30	24,177	14
B.	RoC	CoR	0.097	6.99	4.28	24,003	14
C.	CoR	RoC	0.006	0.70	0.44	24,177	14
D.	RoC	RoC	–0.021	–1.68	–0.88	24,003	14

Explanation: The independent and dependent variables here are in changes. The panel regressions (in differences) explain either CoRs ("$\Delta(v/d)$"), as in equation 6, or RoCs ("$(\Delta v)/d$"), as in equation 7. All change variables are winsorized at –200% and +200%. d.repi.state and year fixed effects (no firm fixed effects) are always included, but their coefficients are not reported. Only the coefficient on the variable of interest (i.e., changes in real-estate values) is shown. Row A is a close kin of the CST12 *level regression with firm-fixed effects*.

Interpretation: The association is strongly positive only in the CoR-CoR regression (A) with shared denominator volatility. It disappears in (C) and (D).

Source: 2feregs-2007.R. July 14, 2020

Table 7: Univariate Time-Series Placebo

Panel A: Simulated Time-Series of Inputs

Variable	Within Simul	Mean	SD	(Over 100,000 Simulations)				
				0%	10%	50%	90%	100%
A1. x	Mean	24.79	3.25	11.10	20.63	24.79	28.97	39.69
	SD	3.52	1.65	0.55	1.57	3.27	5.81	11.79
A2. y	Mean	24.78	3.24	11.71	20.61	24.79	28.92	39.71
	SD	3.51	1.65	0.47	1.55	3.27	5.81	12.16
A3. d	Mean	1.25	0.03	1.13	1.21	1.25	1.28	1.37
	SD	0.19	0.03	0.07	0.16	0.19	0.23	0.32
A4. 1/d	Mean	0.82	0.02	0.75	0.80	0.82	0.85	0.90
	SD	0.13	0.02	0.05	0.10	0.13	0.15	0.23

Panel B: Estimation Results Assessing Correlation of x and y

Method	Stat	Mean	SD	(Over 100,000 Simulations)				
				0%	10%	50%	90%	100%
B1. FE (OLS)	Coef	0.58	0.36	–2.02	0.11	0.61	1.01	2.56
	T-stat	3.834	2.831	–5.624	0.532	3.534	7.523	23.404
B2. FE w/ 1/PPE	Coef	0.47	0.62	–4.70	–0.23	0.47	1.20	4.76
report stats only on x	T-stat	2.884	3.394	–17.573	–1.122	2.668	7.143	31.851
B3. CoR y, CoR x	Coef	0.89	0.19	0.12	0.66	0.88	1.14	2.40
	T-stat	8.361	2.709	0.446	5.225	8.034	11.892	29.567
B4. CoR y, RoC x	Coef	0.00	0.25	–1.44	–0.32	0.00	0.32	1.29
	T-stat	0.005	1.082	–5.587	–1.355	0.005	1.367	5.413
B5. RoC y, CoR x	Coef	–0.01	0.09	–0.61	–0.12	–0.01	0.10	0.57
	T-stat	–0.112	1.078	–6.251	–1.464	–0.114	1.237	5.539
B6. RoC y, RoC x	Coef	–0.08	0.78	–4.26	–1.04	–0.08	0.88	4.37
	T-stat	–0.114	1.083	–5.953	–1.468	–0.111	1.243	5.197

Explanation: Each placebo draw is one time-series for a firm with T=20 years of data. The processes are

$$\tilde{x}_t = \pi_t \cdot (20 + \tilde{r}_t)\ , \qquad y_t = \pi_t \cdot (20 + \tilde{s}_t)\ , \qquad \tilde{d}_t = \pi_t \cdot \exp(\tilde{e}/10)\ ,$$

where π_t is 1.02^t, $t \in [1...T]$, $A = T$, $\tilde{r}_t$ $\tilde{s}_t$ are unit random walks and $\tilde{e}$ is a random normal. The reported statistics are based on 100,000 simulations. The top panel shows the distribution of input variables into the regressions. The bottom panel shows estimation results with various techniques. The implementing (28-line) R source code is in Appendix Table $\mathcal{D}$.1.

Explanation: Only regressions with RoC components (B4-B6) perform well.

Source: simul-simple.R. July 14, 2020

Table 8: Panel Regression Placebo, T-statistics

Panel A: Only redol/P, as in Table 4, Specification A

	Moments		Percentiles							Positive
	Mean	SD	0%	1%	25%	50%	75%	99%	100%	#>0
redol/P	8.27	1.12	3.28	5.73	7.50	8.25	9.01	10.97	13.65	All

Panel B: Only 1/P, no redol/P

	Mean	SD	0%	1%	25%	50%	75%	99%	100%	#>0
1/P	7.26	1.21	2.71	4.59	6.42	7.22	8.05	10.25	13.27	All

Panel C: redol/P with 1/P Control, as in Table 4, Specification B

	Mean	SD	0%	1%	25%	50%	75%	99%	100%	#>0
redol/P	7.02	1.09	1.99	4.55	6.27	7.00	7.74	9.63	11.98	All
1/P	5.49	1.14	1.23	2.96	4.71	5.47	6.24	8.28	11.45	All

Panel D: Differences, as in Table 6

capex	redol	Mean	SD	0%	1%	25%	50%	75%	99%	100%	#>0
CoR	CoR	17.86	4.91	4.90	9.34	14.42	17.16	20.53	32.65	68.57	All
RoC	CoR	0.88	1.11	–5.93	–1.65	0.16	0.85	1.57	3.77	9.34	0.80
CoR	RoC	2.66	1.49	–3.58	–0.21	1.65	2.50	3.48	7.04	14.47	0.98
RoC	RoC	0.15	1.13	–8.02	–2.58	–0.56	0.14	0.85	2.95	7.20	0.56

Explanation: Each placebo draw is a panel with 3,000 firms and 15 years. The reported T-statistics are from 100,000 panel regressions (with clustering) obtained from various estimations techniques. Panels A-C include fixed effects for both firms and years, Panel D only for years. The placebo generation and estimation is described starting on Page 17. The implementing (80-line 1.5 page) R source code is in Appendix Table $\mathscr{D}$.2.

Interpretation: Panel C shows that including 1/P as an additive control does not solve the redol/P bias problem. Only the RoC-based regressions in Panel D do.

Source: simul-cst.R. July 17, 2020

Table 9: Adding log(1/P) as a Control Variable

		on redol/P		on 1/P		on log(1/P)	
		Coef	T_{ols}	Coef	T_{ols}	Coef	T_{ols}
A.	CST12 Regression	0.074	9.01				
B.	CST20 (2020) Regression	0.049	6.91	0.122	13.07		
C.	With Log(1/P)	–0.031	–7.55			0.185	54.08
D.	With 1/P and Log(1/P)	–0.020	–5.64	0.045	12.42	0.149	33.07
E.	RoC Regression	–0.021	–1.68				
F.	RoC Regression with 1/P, log(P), plus Δlog(P)	–0.015	–1.21	0.033	8.54	0.010	4.91

Explanation: This table adds rows C and D to rows A and B from Table 4. Not shown, the results are similar with 10 group dummies based on 1/P.

Interpretation: Dividing by lagged PP&E introduced non-linearities into the investment ratio. The real-estate variable accidentally captured this association.

Source: 2feregs-2007.Rout, but handentered.. July 26, 2020

Table 10: CST12 Table 5, Column 7 Analog: REisPOS*repi.offc as Independent Variable

Panel A: CST12 and Associations With Quasi-Constant Denominator 1/P

	y Dep ← Indep (key) x + ⋯	On (key) ΔREisPOS*repi Coef	T_{ols}	T_{clstr}	Obs	Firms/Yrs
A0.	capex/P ← REisPOS*repi.offc + ⋯	0.213	3.39	2.27	18,031	2,258/14
A1.	1/P ← REisPOS*repi.offc + ⋯	0.823	5.88	3.30	18,218	2,263/14
A2.	add log(1/P)	0.062	1.03	0.72	18,031	2,258/14

Panel B: CoR and RoC Specifications

	y Dep ← Indep (key) x + ⋯	On (key) ΔREisPOS*repi Coef	T_{ols}	T_{clstr}	Obs	Firms/Yrs
B1. CoR	Δ(capex/P) ← ΔREisPOS*repi + ⋯	–0.046	–0.32	–0.22	16,243	2,161/13
B2. RoC	(Δcapex) /P ← ΔREisPOS*repi + ⋯	–0.348	–2.24	–1.69	16,243	2,161/13

Explanation: This table reports the key coefficients from panel regressions that are analogous to those in Chaney, Sraer, and Thesmar (2012, Table 5 Column 7). They use as the key independent variable the interacted dummy variable whether the firm has real estate (REisPOS*repi). (In line with CST12, the regressions also include repi.offc itself and more controls, such as the cash earnings and market-to-book ratio, age, SIC index, ROA, assets, and some cross-dummies.)

Interpretation: Control for log(1/P or the denominator eliminates the positive inference.

Source: 2feregs-2007.Rout. July 14, 2020

Table 11: Evidence Extended to 2018

Panel A: Analogs of Table 3 (CST12 Table 5, Column 1): Shock Identification

	y Dep ← Indep (key) x + ···	On (key) redol/P Coef	T_{ols}	T_{clstr}	Obs	Firms/Yrs
A1. –2007	capex/P ← redol/P + ···	0.074	19.10	8.98	27,201	3,024/15
A2. –2018	capex/P ← redol/P + ···	0.063	22.00	9.38	31,655	2,710/26

Panel B: Analogs of Table 3 and Table 10 B: Denominator Association

	y Dep ← Indep (key) x + ···	Coef	On (key) x T_{ols}	T_{clstr}	Obs	Firms/Yrs
B1.	capex/P ← 1/P + ···	0.127	55.23	17.23	34,298	2,988/26
B2.	1/P ← redol/P + ···	0.230	34.99	9.35	34,130	2,712/26
B3.	capex/P ← redol/P+log(1/P) + ···	–0.014	–4.27	–2.38	31,655	2,710/26
B4	capex/P ← REisPOS*repi.state + ···	0.046	16.40	8.75	31,655	2,710/26
B5.	1/P ← REisPOS*repi.state + ···	0.432	7.85	1.56	37,508	2,990/26
B6.	capex/P ← REisPOS*repi+log(1/P) + ···	0.215	10.20	3.88	34,298	2,988/26

Panel C: RoC

	y Dep ← Indep (key) x + ···	Coef	On (key) x T_{ols}	T_{clstr}	Obs	Firms/Yrs
C1. Tbl 6, D	(Δcapex) /P ← (Δredol) /P + ···	–0.017	–1.83	–0.85	28,859	-/25
C2. Tbl 10, B	(Δcapex) /P ← ΔREisPOS*repi.state + ···	–0.938	–10.90	–3.39	31,240	-/25

Explanation: This table repeats earlier analyses for the sample extended to 2018.

Interpretation: The inference to 2018 is the same as that to 2008.

Source: 2feregs-2018.R. August 1, 2020

APPENDICES

$\mathscr{A}$. The Real-Estate Shock Aspect

$\mathscr{B}$. Year-by-Year Associations

$\mathscr{C}$. Differential Trends

$\mathscr{D}$. Placebo Source Code

$\mathscr{E}$. Temporary Appendices: Earlier Placebos. Earlier Static Incept Denominators. Further Challenges and Investment Theory (Levels vs. Changes).

$\mathscr{A}$ The Real-Estate Shock Aspect

This appendix shows that despite the title suggesting an analysis of real-estate shocks, the independent variable in Chaney, Sraer, and Thesmar (2012) (real-estate) is too persistent to allow for the empirical identification of shocks. If anything, the data itself is more inclined to support a *reductio ad absurdum* association between *future* (rather than *contemporaneous*) real-estate values and corporate investment. It is only their theory-based prior that anchors the association of real estate innovations to contemporaneous capital expenditures.

[Insert Table $\mathscr{A}$.1 here: **"Shock" Identification**]

Panel A of Table $\mathscr{A}$.1 repeats the original specification and then explores the quality of the time identification—the shocks in "how real estate shocks do affect corporate investment." The regressions show that future real-estate variables have stronger associations with corporate investment than current real-estate variables—yet presumably firms did not make corporate investments in anticipation of *future* collateral shocks. The farther into the future, the higher the coefficient estimates become. The 1-year through 5-year ahead coefficient estimates are 0.074, 0.073, 0.074, 0.078 and 0.083, respectively. (The T-statistics drop with the number of observations.) Panel B shows the same pattern for the REisPOS*repi variable.

The evidence suggests the time identification in CST12 is not empirical. That is, the data itself cannot suggest that it is contemporaneous real-estate value shocks that matter. It is only the theoretical a-priori identification that can.[14]

[14] Chaney, Sraer, and Thesmar (2020) retort that the future capex relationships are not absurd. After all, the redol/P variables are highly persistent. We agree on the persistence. We differ on the implication. Their conclusion is that the persistence is an unfortunate aspect of the data—if the reader has the prior that the theory is correct, then the data is not inconsistent with the hypothesis that it is contemporary real-estate values that matter. My conclusion is that it effectively voids the shock identification.

Table 𝒜.1: "Shock" Identification

Panel A: Future Real-Estate Values as Alternative (Solo) Independent Variables

	y Dep ← Indep (key) x + ···	On (key) redol/P Coef	T_{ols}	T_{clstr}	Obs	Firms/Yrs
A0.	capex/P ← redol/P + ···	0.074	19.09	8.98	27,201	3,024/15
A1.	capex/P ← lead(redol,1)/P + ···	0.074	17.62	7.83	24,059	2,898/14
A2.	capex/P ← lead(redol,2)/P + ···	0.073	16.10	7.49	21,178	2,694/13
A3.	capex/P ← lead(redol,3)/P + ···	0.074	15.00	7.34	18,500	2,520/12
A4.	capex/P ← lead(redol,4)/P + ···	0.078	14.31	6.71	16,003	2,328/11
A5.	capex/P ← lead(redol,5)/P + ···	0.083	13.77	6.89	13,685	2,096/10

Panel B: Shock Identification: Future Real-Estate Interactions as Independent Variables

	y Dep ← Indep (key) x + ···	On (key) REisPOS*repi.offc Coef	T_{ols}	T_{clstr}	Obs	Firms/Yrs
B0.	capex/P ← REisPOS*repi.offc + ···	0.213	3.39	2.27	18,031	2,258/14
B1.	capex/P ← lead(REisPOS*repi.offc,2) + ···	0.189	3.19	2.00	13,867	1,992/12
B2.	capex/P ← lead(REisPOS*repi.offc,3) + ···	0.215	3.44	2.27	11,971	1,845/11
B3.	capex/P ← lead(REisPOS*repi.offc,4) + ···	0.182	2.53	1.85	10,216	1,666/10

Explanation: This table reports the key coefficients and T-statistics from fixed-effects regressions that are analogous to those in Chaney, Sraer, and Thesmar (2012, Table 5 Column 1). The regressions include all firm and year fixed effects and a real-estate price index (repi.state), but do not report their coefficients. The variables are described in Table 1, descriptive statistics are in Table 2. capex are capital expenditures, redol is real-estate dollar value, P is (time-varying) lagged net property, plant, and equipment. Panel A shows the lead relationships for the redol/P variable, Panel B for the REisPOS*repi variable. The T_{ols} statistic is the standard OLS T-statistic; the T_{clstr} statistic clusters standard errors according to the included fixed effects (here, two-way clustering for firms and years).

Interpretation: To the extent that the time identification is empirical, it does not point to the contemporaneous real estate, but (absurdly) to future real estate.

Source: 2feregs-2007.R. July 14, 2020

tbl:year-by-year in `cst-appendix.tex`, *2020-08-04*

ℬ Year-By-Year Associations

Table ℬ.1: Year by Year Coefficients

		Yearly Coefficients							Mean
CoR redol/P	1994-2000	0.33*	0.21*	0.29*	0.17*	0.18*	0.16*	0.20*	
	2001-2007	0.20*	0.15*	0.09*	0.10*	0.09*	0.29*	0.14*	0.18
RoC redol/P	1994-2000	–0.12	–0.46	–1.27*	–0.33*	–0.06	–0.02	–0.09*	
	2001-2007	0.11	–0.08	0.03	0.05	0.05	–0.09	0.10	–0.16
CoR REisPos*repi	1994-2000	0.37	1.31	–1.01	1.47*	0.75*	–0.60	–0.43	
	2001-2007	–2.31*	–0.86*	0.44	–0.55	–0.42	0.33		–0.12
RoC REisPos*repi	1994-2000	–1.47*	–1.41	–3.89*	–1.23*	–0.29	–1.46*	–1.31*	
	2001-2007	–1.80*	–0.35	0.04	–1.22*	–0.81	–0.08		–1.18

Explanation: This table reports year-by-year cross-sectional coefficients for the panel regression equivalents from Table 6. The independent variable in the first two parts is based on redol/p, in the second two parts on REisPOS*repi. The dependent variable is always based on capital expenditures. The change variables were winsorized at –200% and +200%. Not shown, differences in repi.state are always appropriately included, too. Starred values have absolute T-statistics of 2.0 or more.

Interpretation: The only stable positive year-by-year relationship appears in the CoR regressions with $\Delta($ redol/P $)$.

Table ℬ.1 shows that the CoR specification, with its covariation in both $\Delta($ capex/P $)$ and $\Delta($ redol/P $)$, has a strong positive association in each and every year. The year-averaged coefficient of 0.18 is similar to the coefficient of 0.198 in the panel regression. In contrast, the RoC specification with $((\Delta$capex$)\,/$P and $((\Delta$redol$)\,/$P$)$ does not have a reliable positive association. The associations of REisPOS*repi are never solidly positive, even if the dependent variable is the CoR $\Delta($ capex/P $)$.

$\mathscr{C}$ Differential Trends (REisPOS*repi)

The positive relationship between Δ(capex/P) and ΔREisPOS*repi relies on another feature of the data, especially the surprising positive coefficient in Table 11 Row B6. This section discusses how firms without real estate started out *observably* different. They were smaller and had higher initial capex. These firms experienced subsequent age-related growth reversion towards the investment rates of slower-growing firms. This happened during a period of generally increasing real-estate prices. This appendix shows that merely allowing all firms without real estate just one shared deterministic *average* investment trend explains their positive coefficients in CST12. Real-estate prices played no role.

[Insert Figure $\mathscr{C}$.1 here: **Investment Ratios by Firm Size *in Incept Year***]

Figure $\mathscr{C}$.1 shows that firms without real estate started out smaller than firms with real estate and had higher investment. Most large firms owned real estate. Many small firms did not.

[Insert Table $\mathscr{C}$.1 here: **Descriptive Statistics for Firms With vs. Firms Without Real Estate**]

Panel A of Table $\mathscr{C}$.1 shows that firms without real-estate were different at the start of the sample: The average firm without real estate had assets of only \$35 million and PP&E of \$10 million, while the average firm with real estate had assets of \$1.1 billion and PP&E of \$420 million. Moreover, the average firm without real estate spent 67% of its PP&E on capex, while the average firm with real estate spent 26%.

Then, over time, as small high-investment firms aged, their investment ratios converged about halfway towards the investment ratios of bigger firms. Panel B of Table $\mathscr{C}$.1 shows that the capex ratios of zero-real-estate firms—smaller firms having started with higher investment ratios—subsequently slowed their investment spending growth. In the first year, they increased capex as a fraction of P by 12%; in the last year, only by 7%. In contrast, firms with real estate increased capex by a fairly stable 2-3%.

Thus, firms with real estate started and ended with stable investment ratios. Firms without real estate started at about 0.68 and declined to about 0.41, halfway towards the levels of firms with real estate (0.26 to 0.22).

These numbers should be taken with a grain of salt, because they do not reflect differential survivorship bias. They only reflect the particular sample used in CST12.

Second, the reason why these differential trends in investment growth induced a positive force on the fixed-effects-held-constant coefficient is that the real-estate price index repi also trended up in most markets throughout the Chaney, Sraer, and Thesmar (2012) sample period (by about 5.3% and 5.8% for firms with vs. without real estate).

[Insert Table $\mathscr{C}$.2 here: **Deterministic Trends For Firms With vs. Firms Without Real Estate**]

The following conjecture is testable:

> *As real-estate prices crept up, the investment ratios of large stable firms with real estate tended to stay the same. This was better than the declining investment ratios of small fast growing firms.*

It requires adding an interacted differential trend control (REisPOS*yrtrend) as a control-variable, where yrtrend is a simple year counter. Table $\mathscr{C}$.2 shows that this single extra control variable is enough to reverse the sign of the CST12 REisPOS*repi variable. The coefficient on REisPOS*repi.offc drops from +0.213 to –0.199. I emphasize that the only information that is needed to construct the REisPOS*yrtrend control variable is whether the firm did or did not have real estate at incept time. This was known at the outset of the sample. The variable incorporated no subsequent real-estate value or pricing (or investment or size) information.

Table $\mathscr{C}$.1: Descriptive Statistics for Firms With vs. Firms Without Real Estate

Panel A: Means in Incept Year (typically 1992)

	Variable	All	Zero	Positive	$[\epsilon, 100\%]$	$[100\%, \infty]$
			Classified by *@Incept* Year Ratio redol/P			
(sort)	redol/P, @Incept	0.74	0.00	1.36	0.57	2.04
	lagPPE/lagAT	0.30	0.17	0.40	0.42	0.39
	lagPPE, in $	235	10	421	400	439
	lagAT, in $	624	35	1,111	949	1,251
	log(lagAT), in log-$	3.8	2.2	5.1	5.2	4.9
	capex/P	0.45	0.67	0.26	0.23	0.28
	Number of Firms	3,004	1,359	1,645	767	878

Panel B: Means in Incept and Final Years, Growth Rates

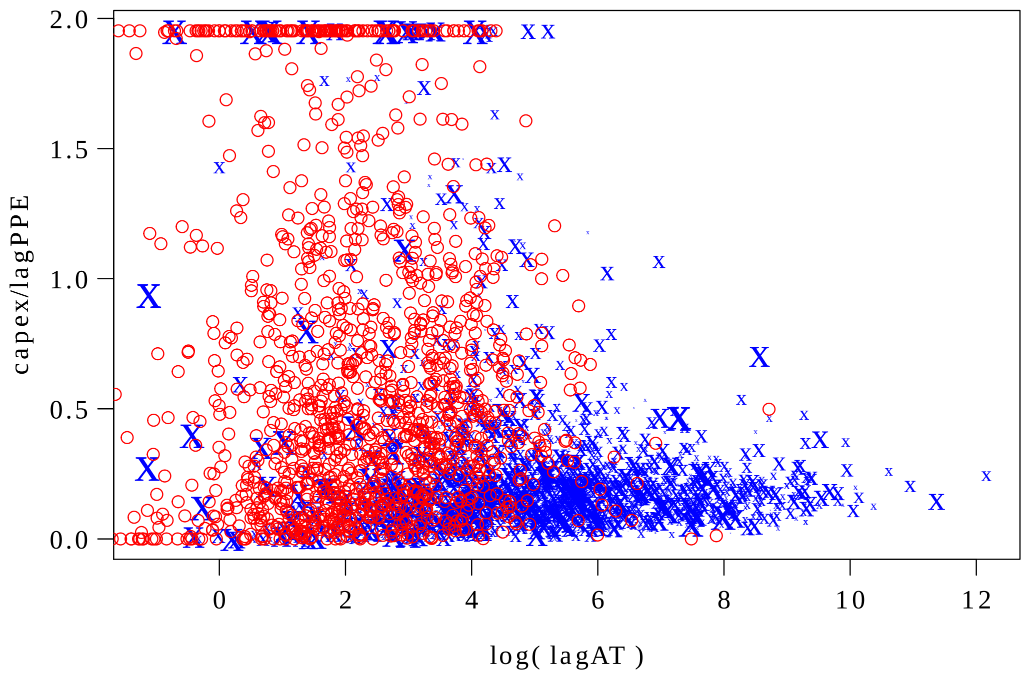
	Variable	All	Zero	Positive	$[\epsilon, 100\%]$	$[100\%, \infty]$
			Classified by *@Incept* Year Ratio redol/P			
(sort)	redol/P, @Incept	0.822	0.000	1.404	0.584	2.076
	capex/P, @Incept	0.44	0.68	0.26	0.24	0.28
	capex/P, @Final	0.30	0.41	0.22	0.22	0.21
	$(\Delta$capex$)$/P, @Incept	0.07	0.12	0.03	0.03	0.03
	$(\Delta$capex$)$/P, @Final	0.04	0.07	0.02	0.03	0.02
	Number of Firm-Years, All	23,589	9,772	13,817	6,228	7,589

Explanation: To be included, a firm had to have complete data for its panel. AT are book assets.

Interpretation: Firms without real estate started out with higher investment as a fraction of their lagged capital stock (0.68), but this declined over time (to 0.41). Firms with real estate started out with lower investment (0.26) and this remained about the same (0.22).

Source: 1descriptive.R. July 16, 2020

Figure 𝒞.1: Investment Ratios by Firm Size *in Incept Year*



Explanation: This plots firm size and capex investment ratios in the first year in which a firm appears in the sample. Red circles are firms without real estate. The size of the blue "X" is the log of 1+redol/P. The sample and winsorizations are as in CST.

Interpretation: Even at the start of the sample, firms without real estate were *observably* different from firms with real estate.

Source: 1descriptives.R: laglogassets-vs-CAPEX1-LAGPPE1.pdf.

Table 𝒞.2: Deterministic Trends For Firms With vs. Firms Without Real Estate

Panel A: To 2008

	y Dep	←	Indep (key) x + ···	on (key) REisPOS*repi Coef	T_{ols}	T_{clstr}
A0. As in CST12	capex/P	←	REisPOS*repi.offc + ···	0.213	3.39	2.27
A1. Add 1 Control control for REisPOS-specific trend	capex/P	←	REisPOS*repi.offc + ··· +REisPOS*yrtrend	–0.199	–2.37	–2.18
A2. Add 14 Controls control for REisPOS-specific year dummies	capex/P	←	REisPOS*repi.offc + ··· +REisPOS*{yrdummies}	–0.250	–2.75	–2.40

Panel B: Extended to 2018

	y Dep ← Indep (key) x + ···	on (key) REisPOS*repi Coef	T_{ols}	T_{clstr}
B1. linear trend	capex/P ← REisPOS*state + ··· +REisPOS*yrtrend	0.015	0.362	0.202
B2. year dummies	capex/P ← REisPOS*state + ··· +REisPOS*{yrdummies}	–0.105	–1.315	–1.055

Explanation: The data panel contains 18,031 observations for 2,258 firms over 14 years. Row A0 repeats the original CST-based results from Table 10. Rows A1 and B1 show regressions that add only one extra control variable: a dummy for whether the firm does or does not have real estate at incept time, multiplied by a deterministic year-trend (e.g., yrtrend $\equiv$ (year − 1992)/1000). Rows A3 and B3 replace the interacting linear year trend with a full set of interacted year dummies.

Interpretation: REisPOS*repi.offc (real-estate values) had picked up a non-stochastic linear differential investment trend between firms with real estate vs. firms without real estate. (Firms without real estate tended to reduce their capital expenditure ratios more over time than firms with real estate.)

Source: 2feregs-2007.Rout. July 14, 2020

𝒟 Placebos Source Code

Table 𝒟.1: Univariate Time-Series Placebo

```r
MC <- 100000
T <- 20
trend <- 1.02^(1:T)  ## this creates problems for the FE 1/d correction
posxy <- T

rw <- function(T) arima.sim(n=T-1, list(order=c(0,1,0)))
lag <- function(x) c(NA, x[1:(length(x)-1)])
msd <- function(x) c(mean(x), sd(x))
reg.CT <- function( formula ) round(coef(summary(lm( formula )))[2,c(1,3)], 3)
p2 <-  function( t1, t2 ) paste( t1, rep(t2,each=2), sep="." )
rbind.lapply <- function( ... ) do.call("rbind", mclapply( ... ))

run.ts <- function(i) {
    set.seed(i)

    x <- trend * ( posxy + rw(T) ); lx <- lag(x)
    y <- trend * ( posxy + rw(T) ); ly <- lag(y)
    d <- trend * exp(rnorm(T)/10);  ld <- lag(d)  ## jagged spikes, not rw

    v <- c(reg.CT( I( y/d )  ~  I(x/d) ),
           reg.CT( I( y/d )  ~  I(x/d) + I(1/d) ),
           reg.CT( I( y/d - ly/ld )  ~  I( x/d - lx/ld ) ),
           reg.CT( I( (y-ly)/ld )  ~  I( (x-lx)/ld ) ),
           reg.CT( I( (y-ly)/ld )  ~  I( x/d - lx/ld ) ),
           reg.CT( I( y/d - ly/ld )  ~  I( (x-lx)/ld ) ),
           msd(x), msd(y), msd(d), msd(1/d) )

    names(v) <- c( p2( c("C","T"), c("fe", "fe1", "rfd", "dafd", "ydafd", "xdafd") ),
                   p2( c("m", "sd"), c("x", "y", "d", "oned")) )
    v
}

many.ts <- rbind.lapply( 1:MC, run.ts )

print( iaw$kable( iaw$summary(  many.ts ), digits=3 ) )
                                        # or just print( summary( many.ts ) )
```

Explanation: This is the R source code executing the degenerate panel (univariate time-series) placebo generations and estimations that are summarized in Table 7.

Table 𝒟.2: A More Realistic Placebo

```r
## simul-cst.R: Ivo Welch, July 20 2020

N <- 100000      ## Firms
Y <- 15          ## Years
MC <- 100000     ## Monte-Carlo Simulations

library(fixest)  # fixed-effects regression, Laurent Berge's great package!

################ Placebo Generation
create.sandbox <- function(N, Y) {
                                        # a new firm-year matrix with starting vals
    newmat <- function( v ) { mt <- matrix(NA, ncol=Y, nrow=N); mt[,1] <- v; mt }
    lagmat <- function( m )  cbind( rep(NA,N), m[ 1:N, 1:(Y-1) ] )
    r.grow.lagppe <- function( lagppe.v, mn, sd ) rnorm(N,mn,sd)*lagppe.v  # frac, not 1+frac
    winsor <- function( v, r=c(0,Inf) ) ifelse( v<r[1], r[1], ifelse(v>r[2], r[2], v) )

    ################ repi is an index series
    repi <- newmat( 1.0 );
    for (y in 2:Y) repi[,y] <- winsor( repi[,y-1] * (1.0 + rnorm(N, 0.05, 0.05)) )

    ################ lagppe starts exponentially distributed for scale heterogeneity
    init93.lagppe <- exp( rnorm(N,3.0,4.0) )
    lagppe  <- newmat( init93.lagppe )
    for (y in 2:Y) lagppe[,y] <- winsor( lagppe[,y-1] + r.grow.lagppe(lagppe[,y-1],0.03,0.20) )

    ################ capex (can turn negative)
    capex <- newmat( init93.lagppe * rnorm(N,0.21,0.40) )  ## the initial value
    for (y in 2:Y) capex[,y] <- capex[,y-1] + r.grow.lagppe(lagppe[,y-1],0.01,0.16)

    ################ redol (must be pos);
    redol <- newmat( init93.lagppe * rbinom(N,1,prob=0.6) * rnorm(N,1.0,1.5)  )
    for (y in 2:Y) redol[,y] <- redol[,1]*repi[,y]

    ################ prepare properly formatted return

    m2v <- as.numeric
    aux <- data.frame(lagppe=m2v(lagppe), laglagppe=m2v(lagmat(lagppe)),
                      capex=m2v(capex), redol=m2v(redol),
                      lagcapex=m2v(lagmat(capex)), lagredol=m2v(lagmat(redol)))

    d <- data.frame(fm=rep(1:N, Y), yr=rep(1:Y, each=N),
                    cbind(repi=m2v(repi), d.repi=m2v( repi-lagmat(repi) )))

    ## create the ratio variables for the panel regregressions
    d <- with(aux, within(d, {
        capex.lagppe <- winsor( capex/lagppe, c( -1.5, 2.0) )
        one.lagppe   <- winsor( 1/lagppe, c(0, 6.0) )
        redol.lagppe <- winsor( redol/lagppe, c(0, 5.9) )

        d.bad.capex.lagppe <- winsor( capex/lagppe - lagcapex/laglagppe, c(-2,2) )
        d.bad.redol.lagppe <- winsor( redol/lagppe - lagredol/laglagppe, c(-2,2) )
        d.good.capex.lagppe <- winsor( (capex-lagcapex)/lagppe, c(-2,2) )
        d.good.redol.lagppe <- winsor( (redol-lagredol)/lagppe, c(-2,2) )
    }))

    d[complete.cases(d),]
}
```

```r
################ CST Panel Regressions

coefrow <- function( fixestresults, grepnames ) {
    ct <- coeftable(fixestresults); ct[  grepl(grepnames, rownames(ct)),  ] }
coefrow.T <- function( fixestresults, grepnames ) coefrow(fixestresults, grepnames)[1,3]

estimate.sandbox <- function( d ) {
                                        # the original CST12 regression
    fe.solo.re <- feols( capex.lagppe ~ redol.lagppe + repi | fm+yr, data=d )
                                        # the ad-absurdum (denominator) regression
    fe.solo.one <- feols( capex.lagppe ~ one.lagppe + repi | fm+yr, data=d )
                                        # the response CST20 regression includes 1/lagppe
    fe.both <- feols( capex.lagppe ~ redol.lagppe + one.lagppe + repi | fm+yr, data=d )

                                        # the incorrect differences
    fe.bad.delta <- feols( d.bad.capex.lagppe ~ d.bad.redol.lagppe + d.repi, data=d )
                                        # the correct differences
    fe.good.delta <- feols( d.good.capex.lagppe ~ d.good.redol.lagppe + d.repi, data=d )

    c(T.solo.redol= coefrow.T( fe.solo.re, ”redol.lagppe” ),
      T.solo.one= coefrow.T( fe.solo.one, ”one.lagppe” ),
      T.both.redol= coefrow.T( fe.both, ”redol.lagppe” ),
      T.both.one= coefrow.T( fe.both, ”one.lagppe” ),
      T.bad.redol= coefrow.T( fe.bad.delta, ”d.bad.redol” ),
      T.good.one= coefrow.T( fe.good.delta, ”d.good.redol”)
      )
}

################ The Main Program, MC (multicore monte-carlo) simulation runs
result <- do.call(”rbind”, mclapply( 1:MC, function(i) {
    set.seed(i)                         # deterministic to allow reproduction
    d <- create.sandbox(N, Y)
    c( mc=i, estimate.sandbox( d ) ) }))

options(digits=3)
print(summary(result))
```

Explanation: This is the R source code executing the panel placebo generations and estimations that are summarized in Table 8.

$\mathscr{E}$ Temporary Appendix

This appendix exists only to discuss aspects from earlier versions. It is here primarily in order not to obsolete Chaney, Sraer, and Thesmar (2020). Although I stand by my old version, my new version is simply much better. It no longer requires the more cumbersome previous approaches.

- The placebo has become simpler and is thus easier to understand and discuss.
- The control for the denominator (where I had used a static single-year PPE as the denominator) was replaced by the RoC approach.

The third subsection below takes issue with the common theoretical implication that a one-time shock to value necessarily has eternal effects on investment flow. Instead, collateral can facilitate investment that is not always immediately paid back.

I also note that the shock identification component of this current draft would probably move into an appendix were it not for the prominence that CST20 have given it. Removing it felt like pulling the rug from under their their (very useful) response. It may yet move into an appendix.

$\mathscr{E}$.1 Discussion of Placebos in Earlier Drafts

There are other placebos, in which including 1/lagPPE results in valid inference on redol.lagPPE. For example, this is the case in the unrealistic example in which all firms start with the same scale. This removes the initial scale-based correlation between capex and redol. I *do not* fully understand the *necessary* conditions which lead to biased coefficients and overconfident T-statistics. I *do* understand the *sufficient* condition. Even the naïve placebo processes in the paper generate the same mechanically-biased inference on redol/lagPPE in specifications with additive 1/lagPPE control, as in Chaney, Sraer, and Thesmar (2020).

My earlier draft entertained a different placebo in which I kept the initial values of each firm (usually 1993), and perturbed the subsequent real-estate values. Chaney, Sraer, and Thesmar (2020) objected that this induced severe correlation between their actual data and my placebo. This was not accidental, but intentional. It is thus important to be clarify what this placebo can and cannot claim to accomplish. My point was to show that *subsequent* real-estate price changes no longer matter, given the initial first-year configuration.

It was my understanding that CST did not intend to test that big firms had more real-estate and stable investment. It this were exclusive to the theory, they can indeed declare victory.

Chaney, Sraer, and Thesmar (2020) and I agree on the fact that if we take the initial configuration of redol, capex, and lagPPE as given, then we can estimate panel coefficients results like those empirical observed, and regardless of how subsequent real-estate prices are scrambled. Presumably, we then agree on the interpretation that the Chaney, Sraer, and Thesmar (2012) theory vis-a-vis this placebo rests on its claim to this initial configuration, and not on later (coincidental) real-estate price changes. Presumably, we agree then that to the extent that other theories can be responsible for the original configuration, they are as consistent with the panel regression coefficient evidence in Chaney, Sraer, and Thesmar (2012) as the CST theory.

My earlier placebo claimed "equal ownership" to the initial configuration, assuming that other non-CST forces could have been responsible for the initial correlations. This can be subject to disagreement. Chaney, Sraer, and Thesmar (2020) did not concede that my placebo was entitled to do so; and that, instead, their theory is *exclusively* entitled to use of the initial configuration—and especially the fact that capital expenditures and real-estate dollar holdings were correlated, if only due to scale, as now clarified by the new placebo in Appendix $\mathscr{D}$. The reader can judge.

> *I am not at all clear about the objection based on Figure 1 in Chaney, Sraer, and Thesmar (2020). The figure shows that there is a high correlation between the "actual \$ real-estate values 1993" and the "Welch Placebo \$ Value Real Estate 1993." If this figure indeed plots what it claims to plot, it plots values in dollars, not in percentages. Scale heterogeneity would induce a high correlation in dollar values among actual and placebo real-estate and capital expenditures values, just as they should (and do in my new placebo). The equation below their graph further adds to my confusion, because it states that they propose*
>
> $$\textit{RE Value Placebo}_{i,t} = \frac{\textit{RE Placebo}_{i,1993}}{\textit{PPE}_{i,t-1}} \cdot \frac{\textit{offprice}_{i,t}}{\textit{offprice}_{i,1993}} ,$$
>
> *but the LHS would be a ratio and not a value. I do not think my confusion is consequential, given the above and below, but I could have responded more clearly to their correlation if I had understood what Chaney, Sraer, and Thesmar (2020) did and meant.*

$\mathscr{E}$.2 Earlier Static Denominators

Table $\mathscr{E}$.1: (Obsolete) Static Denominators

	y Dep ← Indep (key) x + ···	Coef on (key) x: Coef	T-stat	Obs	Years	Firms
A1.	capex/LAGPPE92 ← redol/LAGPPE92 + ···	–0.032	–3.67	27,009		
A2.	capex/LAGPPE00 ← redol/LAGPPE00 + ···	–0.059	–8.27	27,009		
A3.	capex/LAGPPE07 ← redol/LAGPPE07 + ···	–0.007	–0.84	27,009		

Panel B: Static Normalizations, Column 7 like

	y Dep ← Indep (key) x + ···	Coef on (key) x: Coef	T-stat	Obs	Years	Firms
B1.	capex/LAGPPE92 ← REisPOS*repi.offc + ···	–0.067	–0.96	18,031	14	2,258
B2.	capex/LAGPPE00 ← REisPOS*repi.offc + ···	–0.207	–4.08	18,031	14	2,258
B3	capex/LAGPPE07 ← REisPOS*repi.offc + ···	0.012	0.21	18,031	14	2,258

Explanation: This table was in an earlier draft. It remains here only because Chaney, Sraer, and Thesmar (2020) mention it. The variables are normalized by static values in PPE. The table will be removed in the next draft, because the change-variant is much better. LAGPPE92 is the first PPE in the sample for each firm, LAGPPE00 is the PPE in 2000, and LAGPPE07 is the last PPE in the sample.

Source: 2feregs-2007.Rout: July 14, 2020.

$\mathscr{E}$.3 Further Challenges

Papers often raise as many questions as they answer. For a skeptic like myself, positive evidence of a collateral channel should further investigate the following questions:

1. Are there additional *unobservable* non-constant differences between firms that choose to hold real estate and firms that do not (Davidoff (2016))?
2. Is the behavior of the sample of publicly-traded firms also representative of that of privately-held firms? Does it aggregate to the economy or is it only in the cross-section?
3. Was there differential survivorship bias?[15]
4. What would be the effect of using broader measures of corporate investment, such as some that include R&D?
5. What would be the effect of mergers (of equals)? These would increase lagPPE and capex, but not the real-estate variable.
6. Is there a lag structure to corporate investment? (If so, how should we add up the effect?)
7. Is collateral a stock that is used up by a one-time investment, or does it replenish itself and lead to investment in every subsequent period? (See below.)
8. Are there standard error concerns because the independent variable is non-stationary while the dependent variable is stationary (Stambaugh (1999))?

$\mathscr{E}$.3.1 Theory Specification with Levels and Changes

For purposes of my analysis of the empirical CST results, I took the Chaney, Sraer, and Thesmar (2012) hypothesis about the relation between a nearly non-stationary stock variable (real-estate holdings) and a stationary flow variable (investment) as given. However, this is not necessarily the correct or only perspective.

The key question is this: Should a one-time *change* in real-estate value prompt a one-time *change* in investment, either sudden or perhaps tapered; or should it prompt eternal investment changes forever?

One-Time: Construct the simplest possible example: lagPPE‘ is constant. RE is about 40% of PPE at the outset, \$200. At some point, RE gains \$100 and becomes worth \$300. In this year, the company raises its investment by what I assume is an **8%** of the gain in RE. **Thus, with a RE gain of \$100, the investment goes up by \$8.** After the firm has increased its investment by \$8 due to the \$100 gain in (afterwards exhausted) collateral, investment returns to its prior level:[16]

[15] It is quite plausible that survivorship bias could have preferentially removed (smaller/riskier) firms with real-estate in volatile and increasing real-estate markets (who would have invested less; Their real-estate may have appreciated enough to make it worthwhile to liquidate the firm.)

[16] We can do this with two firms, one with and one without real-estate, adding both pc and an ind firm-fixed effect, and the inference comes out the same.

year:	1	2	3	4	5	6	7	8	9	10
lagPPE:	500	500	500	500	500	500	500	500	500	500
x=RE:	200	200	200	200	300	300	300	300	300	300
y=INV:	30	30	30	30	38	30	30	30	30	30

This naive one-time investment due to a collateral increase is not the model in CST's mind. With the CST specification, the estimated coefficient for x would be 0.013 and not 0.08. The 0.013 is the sum-total over all future non-increasing investment years, too.

Persistent: The CST specification presumes a persistent stock effect:

year:	1	2	3	4	5	6	7	8	9	10
lagPPE:	500	500	500	500	500	500	500	500	500	500
x=RE:	200	200	200	200	300	300	300	300	300	300
y=INV:	30	30	30	30	38	38	38	38	38	38

Here, the one-time collateral value delta is not exhausted, but can be used again and again. This seems appropriate, e.g., if the firm can reap a full investment return of $8 in the first year (due to the collateral allowing the extra $8 investment), pay back the $8 collateral, and can then invest another extra $8 again. Another possibility is that the firm has $100 extra in collateral, uses $8 in the first year, then a new $8 (for $16) in the second year, and so on.

A persistent effect is more difficult to estimate empirically, because it involves both stationary flow and non-stationary stock variables.

Both the one-time and the persistent specification seem too simplistic. This concern is not just an issue with the CST paper but also with other papers in the investments literature. The reader must be atuned to the differences in the meaning and interpretation for the CST estimate of $6-$7 per $100. It could map into significantly higher real-estate investment effects, if the true underlying model was the one-time model.