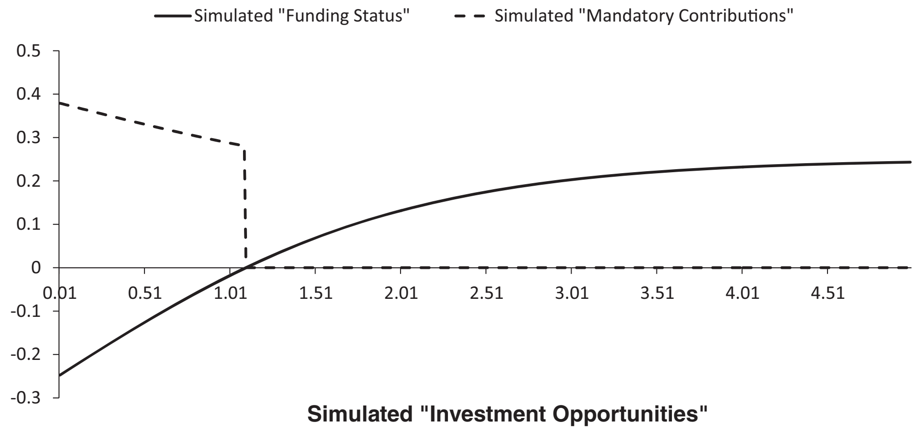

Figure 1. Simulated mandatory contributions. On the horizontal axis is simulated investment opportunities. Simulated funding status is a logistic function of investment opportunities, and simulated mandatory contributions are a kinked function of funding status.

However, although these kinks and jumps clearly diminish any correlation that might be present, they do not necessarily set it to zero. This point is easiest to see in a more general setting. Suppose that two random variables are correlated, and a third is a nonlinear function of the second. Then the claim is that the third is necessarily uncorrelated with the first, conditional on the second. This statement cannot be true if the conditional correlation is calculated via a linear regression because a linear regression cannot provide perfect conditioning in the presence of nonlinear relationships.

We also provide a simple counter example to illustrate this point. Suppose a random variable, y, has a standard lognormal distribution; a random variable, x, is a logistic function of y; and a third variable, z, is zero if $x > 0$ but equals $0.2 + 0.4\,(0.4 - x)$ if $x < 0$. This example captures the main features of the general relation between investment opportunities, funding status, and mandatory contributions; the relations between these variables are depicted in Figure 1. When we simulate a sample of 5,000 observations using these assumptions, we find that z (mandatory contributions) is correlated with y (investment opportunities), conditionally on x (funding status). The t-statistic is around 20. To summarize, it is entirely possible that mandatory contributions are conditionally correlated with investment opportunities and that the regressions in Rauh (2006) do not solve the endogeneity problem. This counterexample shows that identification strategies based on loose analogies to RDD need not be logically correct. However, they may be empirically relevant, so we now turn to examining this latter question.

III. Data

For our empirical analysis, we attempt to reconstruct the data set in Rauh (2006). We start with an unbalanced panel of Compustat firms from the 2007

Standard and Poor's Compustat industrial files. Definitions of the Compustat variables we use are in the Appendix. As in Rauh (2006), we restrict our attention to the subsample of Compustat firms that file an IRS 5500 form with the Department of Labor (DOL) and that sponsor DB pension plans. We need to impose the first restriction to calculate mandatory contributions accurately. All pension plans over a certain size must file an IRS form 5500 yearly. We impose the second restriction because only firms with DB pension plans must make mandatory contributions.

Compustat provides pension data based on firm Securities and Exchange Commission (SEC) filings; however, we do not use these data. As explained in more detail in Rauh (2006), Compustat pension data are aggregated to the firm level. They are therefore inadequate for our purposes because mandatory contributions are determined at the pension plan level—not the firm level. Approximately one-third of the firms in our sample have more than one pension plan, and relying on Compustat data would therefore lead to inaccurate computation of firms' mandatory contributions. Furthermore, firms have significantly more accounting discretion when submitting SEC filings than when filing form IRS 5500. SEC pension data also include both international and domestic pension plans, whereas only domestic plans are required to pay mandatory contributions. Finally, the methods for computing pension liabilities and costs for SEC filings are different from those that are required for computing mandatory contributions.

The sample period runs from 1990 to 1998. The sample starts in 1990 because IRS 5500 forms are first available in standardized form from the DOL in this year. The sample stops in 1998 because we match plans to firms primarily by using the CUSIP of the plan's sponsor, and in 1998 reporting requirements no longer forced pension plans to list the CUSIP pertaining to the plan. Although we match primarily by CUSIPs, missing data occasionally force us to match by employee identification numbers (EINs) or by hand using exact firm names. EINs and firm names only match plans that directly pertain to the Compustat firm. For instance, if a firm's subsidiary sponsors a plan, this plan cannot be matched without access to the CUSIP of the parent firm. Although we attempt to mitigate this problem by using subsidiary names from the Compustat Business Information File, without the CUSIP of the parent of the subsidiary sponsoring the plan, we face potential sample selection issues. After this matching process, we end up with 7,905 firm-year observations, a number somewhat smaller than that in Rauh (2006).

For each plan-year, we extract the following variables from the plan's IRS 5500 filing: pension assets, pension liabilities, funding credits, total contributions, and normal cost, which is the present value of the pension benefits that accrue during the current year. Pension assets are the present value of assets in the plan at the beginning of the year. Pension liabilities are the accumulated benefit obligation from Schedule B of the IRS 5500. From 1991 to 1994, the calculation of pension liabilities follows the regulations in the Omnibus Budget Reconciliation Act of 1987. After 1994, these calculations follow the regulations in the Retirement Protection Act of 1994.

How are mandatory contributions determined? Over the sample period from 1990 to 1998, mandatory contributions consist of two components: the minimum funding contribution (MFC) and the deficit reduction contribution (DRC). Firms must contribute the larger of the MFC or the DRC. The MFC is defined as the sum of the normal cost and an installment payment on unfunded liabilities. It is the normal cost component that provides the necessary discontinuity at the point of underfunding. Given that some discretion is allowed in calculating this installment, we follow Munnell and Soto (2004) and Rauh (2006) and estimate it as 10% of the funding gap. Therefore, the MFC is given by

$$MFC = (\text{Normal Cost}) + 0.1(\text{Funding Gap}).$$

The MFC can be offset by accumulated funding credits, which can be estimated from the 5500 filings. The DRC is given as a fraction of funding status:

$$\begin{aligned}\frac{DRC}{\text{Funding Gap}} &= \min\left[0.30, \left(0.30-0.4\left(\frac{\text{Pension Assets}}{\text{Pension Liabilities}}-0.6\right)\right)\right], \qquad \text{before 1995}\\ &= \min\left[0.30, \left(0.25-0.4\left(\frac{\text{Pension Assets}}{\text{Pension Liabilities}}-0.35\right)\right)\right], \qquad \text{1995 and after.}\end{aligned}$$

The mandatory contribution (MPC) is then given by the following function

$$MPC = \max\,[MFC, DRC],$$

if the funding gap is negative. The minimum and maximum in the above definitions create kinks in the mandatory contribution function. At these kinks, the proportion of the funding gap that must be filled changes. As long as one defines the threshold event as a change in this proportion, these kinks can also be used to define threshold events.

The change to the DRC in 1995 was also accompanied by two provisions that created further discontinuities in the mandatory contribution function. First, plans that were at least 90% funded were exempt from the DRC, and second, plans that were at least 80% funded and that had a recent history of being overfunded were also exempt. These thresholds imply actual discontinuities because at these thresholds the DRC can be much larger than the MFC.

In this setting, two natural definitions of the funding gap arise. In either case, for a firm with one plan, defining the funding gap is straightforward. However, defining a funding gap for a firm with multiple plans is more complicated. One alternative, from Rauh (2006), is to sum the gaps across all pension plans within a firm. This definition, which we refer to as the average funding status, correctly identifies the zero underfunding point only if firms can easily shift funds across plans. Another alternative is to define the funding gap as the smallest gap if all plans have positive funding status, and as the sum of the negative gaps if any plans have negative funding status. This definition captures the notion of the distance to the point of violation if firms cannot shift funds across plans. As we show below, the use of one definition or another is immaterial for our full-sample regressions.

IV. Full-Sample Analysis

We first examine the entire spectrum of funding statuses to pinpoint exactly which kinks and jumps in the mandatory contribution function are likely to provide a causal effect on investment. Here, we use both informal visual and formal statistical analysis and pay close attention to the characteristics of the firms at different points along the function.

A. Visual Analysis

Panel A of Figure 2 plots a kernel regression of the ratio of investment to assets on average funding status, where the latter is scaled by pension liabilities for ease of interpretation. The relevant vertical axis is on the left. The solid line is the fitted regression line, and the dotted lines are 5% confidence bands. This regression closely replicates the result in Rauh (2006) that investment appears to drop off sharply after the point at which pension assets equal pension liabilities. However, the kernel regression does not tell the whole story. The solid line plots a histogram of the funding status, superimposed on the kernel regression, with the vertical axis on the right indicating the number of observations. The drop in investment does not occur until underfunding is between 10% and 15% of pension liabilities, but the histogram indicates that few observations in the sample can be found at or below this point. Indeed, the tenth percentile is at the 12% level of underfunding, and the fifth percentile is at the 22% level of underfunding.

This finding highlights one of the most important concerns facing studies that emphasize identification of causal effects: internal versus external validity. The figure suggests informally that although pension funding problems may indeed cause drops in investment, if this causal effect exists it is likely limited to a few firms. Providing more definitive evidence on this conjecture is our next task.

As a first step, we continue with our informal visual analysis to determine whether there are any causal effects around the point of zero underfunding, where a reasonably large part of the sample lies. Panel B of Figure 2 plots two kernel regressions of investment on the funding gap, one using only observations with a positive gap, and the other using only observations with a negative gap. This type of informal analysis is recommended in Imbens and Lemieux (2008), who comment that formal statistical analysis is essentially just a sophisticated version of this sort of basic plot. The intent is to find a discontinuity in the outcome of interest, which in our case is investment. If no discontinuity can be observed, there is little chance that formal analysis will lead to estimates with economically and statistically significant magnitudes. We find that the two regression lines almost meet at the discontinuity point