Appendix 1: Model Discussion

A. Potential Outcomes Framework

In this section, we derive the equations in Section II using a potential outcomes framework, similar to Imbens and Angrist (1994). A distinct feature of our setting is that we will consider a multivalued instrument. We have a discrete instrument, $Z_i \in \{0,1,2\}$, which represents an exogenous factor that may influence benefit enrollment. When $Z_i = 0$, the default benefit is the DB plan, but employees may switch into a DC plan. When $Z_i = 1$ all employees must enroll in the DB plan, and when $Z_i = 2$ all employees must enroll in the DC plan. We make the following assumptions regarding the exogeneity of the instrument Z_i:

Assumption A.1 (Instrument Exogeneity)

A.1.1. Independence: $\{L_i(B_i(0),0), L_i(B_i(1),1), L_i(B_i(2),2), B_i(0), B_i(1), B_i(2)\} \perp Z_i$;

A.1.2. Exclusion: $L_i(0,0) = L_i(0,1) \equiv L_i(0)$ *and* $L_i(1,0) = L_i(1,2) \equiv L_i(1)$.

Thus, the outcome of leaving is a function of benefit enrollment — $L_i(B_i) = B_i \cdot L_i(1) + (1- B_i) \cdot L_i(0)$ — and benefit enrollment is a function of the instrument — $B_i(Z_i) = \mathbf{1}\{Z_i = 0\} \cdot B_i(0) + \mathbf{1}\{Z_i = 2\}$.

When enrollment is determined by the employee — that is, $Z_i = 0$— the decision rule to enroll is determined by Equation 2:

(1.1)
$$B_i(0) = \begin{cases} 1 & \text{if } \phi_i \geq c_i \\ 0 & \text{if } \phi_i < c_i. \end{cases}$$

We now derive Equation 5:

$$\beta_{Endo} \equiv \mathbb{E}[L_i|B_i = 1, Z_i = 0] - \mathbb{E}[L_i|B_i = 0, Z_i = 0]$$

$$= \mathbb{E}[L_i(1)|B_i(0) = 1] - \mathbb{E}[L_i(0)|B_i(0) = 0]$$

(1.2) $$= \mathbb{E}[L_i(1)|B_i(0) = 1] - \mathbb{E}[L_i(0)|B_i(0) = 1]$$

$$+\{\mathbb{E}[L_i(0)|B_i(0) = 1] - \mathbb{E}[L_i(0)|B_i(0) = 0]\}$$

$$= \underbrace{\mathbb{E}[L_i(1) - L_i(0)|B_i(0) = 1]}_{\beta_1}$$

$$+ \underbrace{\{\mathbb{E}[L_i(0)|B_i(0) = 1] - \mathbb{E}[L_i(0)|B_i(0) = 0]\}}_{\beta_{Selection}},$$

In the second line, we have relied on Assumptions 1.1 and 1.2.

In the exogenous case the employer's nudge completely determines enrollment in the new benefit — $B_i(1) = 0$ and $B_i(2) = 1$. We can thus show, as in Equation 6, that:

$$\beta_{Exog} \equiv \mathbb{E}[L_i|B_i = 1, Z_i = 2] - \mathbb{E}[L_i|B_i = 0, Z_i = 1]$$

$$= \mathbb{E}[L_i(1) - L_i(0)]$$

(1.3) $$= \Pr(B_i(0) = 0) \cdot \underbrace{\mathbb{E}[L_i(1) - L_i(0)|B_i(0) = 0]}_{\beta_0}$$

$$+\Pr(B_i(0) = 1) \cdot \underbrace{\mathbb{E}[L_i(1) - L_i(0)|B_i(0) = 1]}_{\beta_1}$$

$$= \pi_0\beta_0 + \pi_1\beta_1,$$

where in the second line, we have again relied on Assumptions 1.1 and 1.2.

B. Derivation of a Lower Bound on the Selection Effect

As mentioned in Section II in the text, we can use the observed relationship between benefit enrollment and leaving under two distinct choice scenarios to establish a lower bound on the selection effect. We formally state this in the following proposition:

Proposition 1 *If the quasi-experimental estimate defined in Equation 6 is positive ($\beta_{Exog} \geq 0$) and the treatment on the treated is negative ($\beta_1 < 0$), OR if exogenous benefit enrollment increases leave propensity by more among those who would not have endogenously enrolled relative to those who would have enrolled ($\beta_0 \geq \beta_1$), then the difference between the endogenous (Equation 5) and exogenous (Equation 6) estimates is bounded from above by the selection effect defined in Equation 5. That is:*

$$\beta_{Endo} - \beta_{Exog} \leq \beta_{Selection}. \tag{1.4}$$

Before proving Proposition 1, we establish a useful lemma:

Lemma 1 (Selection and Observational Correlations) *If the treatment on the treated is negative ($\beta_1 < 0$), then the observed difference in leave probabilities by benefit type (B_i) defined in Equation 5 is bounded from above by the selection effect defined in Equation 5. That is:*

$$\beta_{Endo} \leq \beta_{Selection} \tag{1.5}$$

The implication of Lemma 1 is that if we observe a positive correlation between the probability of the leaving the firm and endogenous enrollment in the new benefit (that is, $\beta_{Endo} \geq 0$), then we can sign the selection effect as positive (that is, $\beta_{Selection} > 0$). This result is asymmetric, in that a negative correlation (that is, $\beta_{Endo} < 0$) is not informative about the sign of the selection effect.

Proof.

$$\beta_{Endo} = \beta_1 + \beta_{Selection}$$
$$\leq \beta_{Selection}$$

where the first line was shown in Equation 5 and, in the second third line, we have used the assumption $\beta_1 < 0$. The assumption that $\beta_1 < 0$ is guaranteed in this version of the model, due to the nonnegative enrollment cost. To see that, note:

$$\begin{aligned}\beta_1 &= \mathbb{E}[L_i(1)|B_i(0) = 1] - \mathbb{E}[L_i(0)|B_i(0) = 1] \\ &= \Pr(m_i > \phi_i|\phi_i \geq c_i) - \Pr(m_i > 0|\phi_i \geq c_i) \\ &< \Pr(m_i > 0|\phi_i \geq c_i) - \Pr(m_i > 0|\phi_i \geq c_i) \\ &= 0,\end{aligned}$$

where the second line follows from Equation 4, and in the third line, we used the fact that $\phi_i \geq c_i \rightarrow \phi_i \geq 0$, since c_i is nonnegative.[i] Thus, the endogenous effect is bounded above by the selection effect. It follows that a necessary condition for observing a positive β_{Endo} is a positive selection effect.

We now prove Proposition 1.:

Proof. Recall from Equation 6 that:

$$\beta_{Exog} = \pi_1\beta_1 + \pi_0\beta_0$$

Also, recall from above that $\beta_{Endo} = \beta_1 + \beta_{Selection}$. Next, the difference between β_{Endo} and β_{Exog} gives:

$$\begin{aligned}\beta_{Endo} - \beta_{Exog} &= \beta_1 + \beta_{Selection} - \pi_1\beta_1 - \pi_0\beta_0 \\ &= \beta_{Selection} + (1-\pi_1)\beta_1 - \pi_0\beta_0 \\ &= \beta_{Selection} + \pi_0[\beta_1 - \beta_0]\end{aligned}$$

If the second term in brackets, $[\beta_1 - \beta_0]$, is negative, then the results follows. We have focused on two sufficient conditions for this term to be negative. First, note that if $\beta_{Exog} \geq 0$, then we have:

$$\begin{aligned}0 &\leq \beta_{Exog} \\ &= \pi_1\beta_1 + \pi_0\beta_0 \\ &= \beta_1 - \pi_0[\beta_1 - \beta_0] \\ &\leq -\pi_0[\beta_1 - \beta_0] \\ &[\beta_1 - \beta_0] \leq 0\end{aligned}$$

where in the fourth line we have used the assumption that $\beta_1 < 0$. Alternatively, we can just assume that $[\beta_1 – \beta_0]$ is negative. In either case, the result follows.

The assumption that $[\beta_1 – \beta_0]$ is negative will in general be true if the new benefit is less likely to make those who would choose the benefit leave the firm than those *who would not* choose the benefit if given the choice. It makes sense that those for whom values of ϕ_i are high are less likely to have $m_i > \phi_i$, which is how this condition is represented in our model. However, this is not guaranteed to be negative and one could construct counter examples. When this assumption is true, we have the result and a necessary condition for $\beta_{Endo} - \beta_{Exog} \geq 0$ is that $\beta_{Selection} \geq 0$.

A couple of points are worth making about our stylized model. First, it may appear that the dynamics are completely suppressed in our model. In particular, we introduce a friction in decision-making by requiring the enrollment decision to be made before the leave decision, and furthermore do not model forward-looking behavior at the enrollment stage. However, the friction is meant to capture uncertainty about the future leave decision, or at least about the time span between enrollment and leaving. In addition, we can allow for the enrollment decision to be correlated with the leave decision directly through a correlation between ϕ and m, which we have thus far left unrestricted.[ii]

Second, we have to this point modeled a new benefit that only affects mobility, m, through its effect on $E[V_i(w_i,B_i)]$. However, the new benefit we examine in our context (the DC plan) has the potential to directly affect mobility, for example, by reducing or eliminating the vesting requirement for retirement benefits. This can be modeled by allowing η_i, the employment switching cost, to be a function of B_i. We have abstracted here from that interaction. However, we show next in Appendix 1.C. that Proposition 1 still holds in this case, so long as we still assume that $\beta_0 \geq \beta_1$.

C. Allowing for a Direct Effect of Benefit Enrollment on Mobility

In the previous section, we restricted the effect of the new benefit on m to an effect on $E[V_i(w_i,B_i)]$. We now show that an amended version of Proposition 1 still holds once this restriction is relaxed. We now define a new “mobility" parameter, $\tilde{m}$, as the value of mobility, net the switching cost

$$\tilde{m}_i \equiv \mathbb{E}[V_i^o(w_i^o, B_i^o)] - \mathbb{E}[V_i(w_i, 0)].$$

Furthermore, we now allow the employment switching cost to be a function of benefit enrollment, B_i. Without loss of generality, we normalize the switching cost to zero in the absence of the new benefit and define this new function $\tilde{\eta}(B_i)$ as follows:

$$\tilde{\eta}_i(B_i) \equiv B_i \cdot \eta_i$$

It follows that the net benefit of mobility is now:

$$m_i \equiv \tilde{m}_i - \tilde{\eta}_i,$$

and the decision to leave is now made according to the following rule:

$$L_i(B_i) = \begin{cases} 1 \; if \; (\phi_i + \eta_i) \cdot B_i < \tilde{m}_i \\ 0 \; if \; (\phi_i + \eta_i) \cdot B_i \geq \tilde{m}_i. \end{cases}$$

Heterogeneity is now captured by the quadruplet (ϕ, c, $\tilde{m}$, η). The incentive effect is now $\phi + \eta$, and without any further restrictions on η, Lemma 1 no longer holds. In particular, notice that the when $\eta_i < 0$, the benefit enrollment may increase the likelihood of leaving the firm. That is, we may have $\beta_1 \geq 0$. This is the case, for example, when the new benefit does not have as demanding a vesting requirement. Nonetheless, the following, amended version of Proposition 1 is obtained:

Proposition 1a (Selection, Observational Correlations and Quasi-Experimental Estimates with Direct Mobility Effects). *If exogenous benefit enrollment increases leave propensity by more among those who would not have endogenously enroll relative to those who would have enrolled (that is, $\beta_0 \geq \beta_1$), then the difference between the endogenous (Equation 5) and exogenous (Equation 6) estimates is bounded from above by the selection effect defined in Equation 5. That is:*

$$\beta_{Endo} - \beta_{Exog} \leq \beta_{Selection}$$

Proof. To prove this, we use the same steps as above to show:

$$\beta_{Endo} - \beta_{Exog} = \beta_{Selection} + \pi_0[\beta_1 - \beta_0],$$

and the result follows.

D. Derivation of Lower Bound when a LATE is Estimated

As mentioned in Section 2, our results require an estimate of the average treatment effect, β_{Exog}, and our method of 2SRI technically recovers an average treatment effect. However, one may alternatively interpret our estimates as a local average treatment effect (LATE), which is a common interpretation of IV estimates (see for example, Imbens and Angrist 1994). In that case, we must use additional assumptions to establish a lower bound on the selection effect. To see this, redefine the instrument Z_i as a binary variable that takes a value of zero when the default is the DB plan and one when the default is the DC plan. We define the subpopulation of compliers as those who would enroll in the DB in the absence of this default, but who enroll in the DC plan in the presence of it — that is, those for whom $B_i(1) > B_i(0)$. The LATE, then, is defined as follows:

(1.6) $$\beta_{LATE} \equiv \mathbb{E}[L_i(1) - L_i(0) | B_i(1) > B_i(0)].$$

Note that our previously define treatment on the untreated, β_0 is related to the LATE as follows:

(1.7)
$$\begin{aligned}
\beta_0 &\equiv \mathbb{E}[L_i(1) - L_i(0) | B_i(0) = 0] \\
&= \Pr(B_i(1) > B_i(0) | B_i(0) = 0) \cdot \mathbb{E}[L_i(1) - L_i(0) | B_i(1) > B_i(0)] \\
&\quad + \Pr(B_i(1) = B_i(0) = 0 | B_i(0) = 0) \cdot \mathbb{E}[L_i(1) - L_i(0) | B_i(1) = B_i(0) = 0] \\
&= \frac{\pi_C}{\pi_C + \pi_{NT}} \beta_C + \frac{\pi_{NT}}{\pi_C + \pi_{NT}} \beta_{NT} \\
&= \frac{\pi_C}{\pi_C + \pi_{NT}} \beta_{LATE} + \frac{\pi_{NT}}{\pi_C + \pi_{NT}} \beta_{NT},
\end{aligned}$$

where the "C" subscript denotes the complier subpopulation and the "NT" subscript refers to the "never-taker" subpopulation — those for whom $B_i(0) = B_i(1) = 0$. Rearranging terms from Equation 1.7, we have:

(1.8)
$$\begin{aligned}
\beta_{LATE} &= \frac{\pi_C + \pi_{NT}}{\pi_C} \beta_0 - \frac{\pi_{NT}}{\pi_C} \beta_{NT} \\
&= \beta_0 - \frac{\pi_{NT}}{\pi_C} (\beta_{NT} - \beta_0).
\end{aligned}$$

Suppose that $\beta_{Exog} = \beta_{LATE}$, then our key derivation is altered:

(1.9) $$\beta_{Endo} - \beta_{Exog} = \beta_1 + \beta_{Selection} - \beta_0 + \frac{\pi_{NT}}{\pi_C}(\beta_{NT} - \beta_0)$$

$$= \beta_{Selection} + (\beta_1 - \beta_0) + \frac{\pi_{NT}}{\pi_C}(\beta_{NT} - \beta_0)$$

The assumption that $\beta_0 > \beta_1$ is now no longer sufficient to establish a lower bound, but rather we require that the sum of the second and third terms in Equation 1.9 be negative.

In the main text, we maintain the assumption that our method recovers an average treatment effect. In a literal sense, the assumptions required to implement our 2SRI method imply that our estimates recover an average treatment. In addition, the standard results that equate IV estimates to a local average treatment effect, for example Imbens and Angrist (1994), do not technically apply in the case of a nonlinear specification, such as ours.

However, as we show below in Appendix 4, there is an analogous method, the local average response function (LARF) method, that recovers an average treatment effect among the compliers, even in the case of a nonlinear specification. In that case, our empirical estimates are nearly identical to those using the 2SRI method. This either suggests that the effect among the compliers is comparable to the average treatment effect, or that the variation used to identify the 2SRI essentially recovers a local effect. In the former case, we are justified in interpreting our effect as an average treatment effect, while in the latter case, we are not.

Even if our method only identifies a local treatment effect, we have two additional arguments as to why our lower bound is likely to still hold. First, should the third additional term in Equation 1.9 be positive, it is attenuated by a factor or π_{NT}/π_C which is roughly 1/3 in our sample, given our first stage results (available upon request). Second, we estimate the average characteristics of the complier subpopulation in Appendix 1.E. below. In Table 1.1 we compare the complier population to the general sample. We find evidence that compliers are lower tenured, and more likely to be Hispanic than the general sample; however, there is no evidence that they differ in their weekly hours, annual salary, or gender.

E. Complier Analysis

We provide some characteristics of the complier population by using the method described in Autor and Houseman (2005) to estimate the characteristics of the marginal DC enrollee, and report the results in Table 1.1. Column (1) reports the means of various observable characteristics in our sample. Column (2) reports the estimated average characteristic of the "compliers," or those individuals who would not have enrolled in the DC plan were it not for the fact that they were defaulted into the DC plan. Column (3) reports the difference along with standard errors. In all cases, the estimates are regression-adjusted for age.

Table 1.1

Estimated Complier Characteristics

	(1) Sample Average	(2) Complier Mean	(3) Difference
Female = 1	0.162***	0.172***	-0.010
	(0.006)	(0.032)	(0.031)
Black	0.107***	0.116***	-0.009
	(0.005)	(0.025)	(0.024)
Hispanic	0.278***	0.431***	-0.154***
	(0.007)	(0.040)	(0.039)
Asian/Am. Indian/Other	0.158***	0.117***	0.041
	(0.006)	(0.030)	(0.030)
Tenure	11.079***	6.763***	4.316***
	(0.143)	(0.469)	(0.476)
Weekly Hours	39.573***	39.458***	0.115
	(0.040)	(0.211)	(0.206)
Salary (in $1,000s)	46.584***	45.776***	0.808
	(0.201)	(1.064)	(1.047)
N	4,153	4,153	4,153

Appendix 2: Two-Stage Residual Inclusion (2SRI)

Here we demonstrate the control function approach, Two-Stage Residual Inclusion (2SRI) (Terza, Basu, and Rathouz 2008, for example see). Suppose we have a binary outcome, Y_i, a key regressor of interest D_i, a set of predetermined covariates X_i and an instrument Z_i The binary outcome is modeled using a standard probit model:

$$Y_i = \mathbf{1}\{\tilde{\lambda} D_i + \tilde{\Gamma}_1 X_i + u_i > 0\} \tag{2.10}$$

Under the assumption that $u_i \sim \mathcal{N}(0,1)$ is independent of (D_i, X_i), we have the following:

$$\mathbb{E}[Y_i | D_i, X_i] = \Phi(\tilde{\lambda} D_i + \tilde{\Gamma}_1 X_i) \tag{2.11}$$

However, we are interested in the case where D_i may be and endogenous regressor, that is D_i D_t and u_i may be correlated. In this case, the naïve probit regression of Y_i on D_i and X_i will be inconsistent, and in particular, the coefficient on D_i will be biased. Let (λ^*, Γ_1^*) be the parameters estimated from the naïve probit regression. Define the average partial effect of D_i on Y_i using the parameters from this naïve regression as:

$$\beta_{Endo} = \mathbb{E}[\Phi(\lambda^* + \Gamma_1^* X_i)] - \mathbb{E}[\Phi(\Gamma_1^* X_i)]. \tag{2.12}$$

The average partial effect will, by extension, also be biased. Consider the following ancillary regressions:

(2.13) $$D_i = \gamma Z_i + \Gamma_2 X_i + v_i$$

(2.14) $$u_i = \tilde{\alpha} v_i + e_i,$$

where $\tilde{\alpha} \neq 0$ captures the endogeneity of D_i. We make the following identifying assumptions:

Assumption 2.1 (2SRI)

1. *First Stage:* $\gamma \neq 0$.
2. *Independence 1: Conditional on* X_i, Z_i *is independent of* (u_i, v_i, e_i).
3. *Independence 2:* v_i *is independent of* e_i.

We now demonstrate identification of the average partial effect of D_i on Y_i. First, substitute for u_i in Equation 2.10 using Equation 2.14 and we have:

(2.15) $$Y_i = \mathbf{1}\{\tilde{\lambda} D_i + \tilde{\Gamma}_1 X_i + \tilde{\alpha} v_i + e_i > 0\}$$

Note that the error term e_i is independent of the regressors. By normality of u_i, we have $e_i \sim \mathcal{N}(0, \sigma_e)$. Applying standard probit regression results, we have:

(2.16) $$\begin{aligned} \mathbb{E}[Y_i | D_{i,} X_i, v_i] &= \Pr(\tilde{\lambda} D_i + \tilde{\Gamma}_1 X_i + \tilde{\alpha} v_i + e_i > 0) \\ &= \Phi\left(\frac{\tilde{\lambda}}{\sigma_e} D_i + \frac{1}{\sigma_e} \tilde{\Gamma}_1 X_i + \frac{\tilde{\alpha}}{\sigma_e} v_i\right) \\ &= \Phi(\lambda D_i + \Gamma_1 X_i + \alpha v_i) \end{aligned}$$

We do not directly observe v_i, but we can obtain a consistent estimate using the residuals from a linear regression of D_i on Z_i and X_i, as per Equation 2.13. We then estimate a probit regression of

Y_i on D_i, X_i and the estimated v_i. The parameters from the probit estimation of (2.16) are then used to calculate the average partial effect of D_i on Y_i:

(II.17) $$\beta_{Exog} = \mathbb{E}[\Phi(\lambda + \Gamma_1 X_i + \alpha v_i) - \Phi(\Gamma_1 X_i + \alpha v_i)].$$

The variance covariance matrix for the estimate parameters are adjusted for the two-step procedure, using standard results (Newey and McFadden, 1994). Standard errors for the average partial effect are obtained via the delta method.

Appendix 3: Supplemental Results

A. Robustness to Age Definition

Table 3.2
Effect of DC Plan on One-Year Leave Probability and Test for Selection, Alternative Age Comparison

	(1)	(2)	(3)
β_{Endo}	-0.016	-0.013	-0.015
	(0.016)	(0.016)	(0.016)
β_{Exog}	-0.054***	-0.049***	-0.045**
	(0.017)	(0.017)	(0.018)
H_0: $\beta_{Endo} \leq \beta_{Exog}$	0.005	0.012	0.031
$\mathrm{E}[L_i]$	0.077	0.077	0.077
Controls	No	Yes	Yes
Age FEs	No	No	Yes
Year FEs	No	No	Yes
N	4,164	4,145	4,145
First Stage F-stat	622	621	628

Note: Sample includes employees in the years 1999 - 2002. β_{Endo} estimates are from a simple probit regression. β_{Exog} estimates are calculated using Two-Stage Residual Inclusion (2SRI) with a linear first-stage where DC is instrumented for using the default pension plan type based on age in current year and probit second stage. Average partial effects are reported. Robust standard errors are adjusted for first-stage estimation. P-value for H_0 reported for evaluating implication of Equation 7. Demographic controls include gender, race, a cubic in tenure dummies, hours worked per year, and base pay rate. * Significantly different at the 10% level; ** at the 5% level; *** at the 1% level.

Table 3.3

Effect of DC Plan on Two-Year Leave Probability and Test for Selection, Alternative Age Comparison

	(1)	(2)	(3)
β_{Endo}	-0.012	0.001	-0.004
	(0.024)	(0.025)	(0.024)
β_{Exog}	-0.110***	-0.097***	-0.093***
	(0.022)	(0.024)	(0.024)
H_0: $\beta_{Endo} \leq \beta_{Exog}$	0.000	0.000	0.000
$\mathrm{E}[L_i]$	0.146	0.146	0.146
Controls	No	Yes	Yes
Age FEs	No	No	Yes
Year FEs	No	No	Yes
N	3,146	3,146	3,146
First Stage F-stat	622	618	624

Note: Sample includes employees in the years 1999, 2000 and 2002. β_{Endo} estimates are from a simple probit regression. β_{Exog} estimates are calculated using Two-Stage Residual Inclusion (2SRI) with a linear first-stage where DC is instrumented for using the default pension plan type based on age in current year and probit second stage. Average partial effects are reported. Robust standard errors are adjusted for first-stage estimation. P-value for H_0 reported for evaluating implication of Equation 7. Demographic controls include gender, race, a cubic in tenure dummies, hours worked per year, and base pay rate. * Significantly different at the 10% level; ** at the 5% level; *** at the 1% level.

Table 3.4

Effect of DC Plan on Three-Year Leave Probability and Test for Selection, Alternative Age Comparison

	(1)	(2)	(3)
β_{Endo}	-0.014	0.004	0.000
	(0.026)	(0.027)	(0.027)
β_{Exog}	-0.140***	-0.114***	-0.110***
	(0.032)	(0.035)	(0.036)
H_0: $\beta_{Endo} \leq \beta_{Exog}$	0.000	0.000	0.000
$E[L_i]$	0.188	0.188	0.188
Controls	No	Yes	Yes
Age FEs	No	No	Yes
Year FEs	No	No	Yes
N	2,049	2,038	2,038
First Stage F-stat	621	590	592

Note: Sample includes employees in the years 1999 and 2002. β_{Endo} estimates are from a simple probit regression. β_{Exog} estimates are calculated using Two-Stage Residual Inclusion (2SRI) with a linear first-stage where DC is instrumented for using the default pension plan type based on age in current year and probit second stage. Average partial effects are reported. Robust standard errors are adjusted for first-stage estimation. P-value for H_0 reported for evaluating implication of Equation 7. Demographic controls include gender, race, a cubic in tenure dummies, hours worked per year, and base pay rate. * Significantly different at the 10% level; ** at the 5% level; *** at the 1% level.

B. Robustness to Vesting Status

Table 3.5
Effect of DC Plan on One-Year Leave Probability and Test for Selection, Vested Sample

	(1)	(2)	(3)
β_{Endo}	-0.026*	-0.024	-0.024
	(0.015)	(0.016)	(0.016)
β_{Exog}	-0.049***	-0.049***	-0.046***
	(0.015)	(0.015)	(0.016)
H_0: $\beta_{Endo} \leq \beta_{Exog}$	0.078	0.084	0.096
$E[L_i]$	0.060	0.060	0.060
Controls	No	Yes	Yes
Age FEs	No	No	Yes
Year FEs	No	No	Yes
N	2,688	2,630	2,630
First Stage F-stat	380	372	373

Note: Sample includes employees in the years 1999 - 2002 who have at least 5 years of service. β_{Endo} estimates are from a simple probit regression. β_{Exog} estimates are calculated using Two-Stage Residual Inclusion (2SRI) with a linear first-stage where DC is instrumented for using the default pension plan type based on age in 2002 and probit second stage. Average partial effects are reported. Robust standard errors are adjusted for first-stage estimation. P-value for H_0 reported for evaluating implication of Equation 7. Demographic controls include gender, race, a cubic in tenure dummies, hours worked per year, and base pay rate. * Significantly different at the 10% level; ** at the 5% level; *** at the 1% level.

Table 3.6
Effect of DC Plan on Two-Year Leave Probability and Test for Selection, Vested Sample

	(1)	(2)	(3)
β_{Endo}	-0.034	-0.033	-0.034
	(0.025)	(0.025)	(0.024)
β_{Exog}	-0.102***	-0.106***	-0.103***
	(0.024)	(0.022)	(0.023)
H_0: $\beta_{Endo} \leq \beta_{Exog}$	0.003	0.003	0.004
$E[L_i]$	0.126	0.126	0.126
Controls	No	Yes	Yes
Age FEs	No	No	Yes
Year FEs	No	No	Yes
N	2,023	2,004	2,004
First Stage F-stat	380	376	372

Note: Sample includes employees in the years 1999, 2000 and 2002 who have at least 5 years of service. β_{Endo} estimates are from a simple probit regression. β_{Exog} estimates are calculated using Two-Stage Residual Inclusion (2SRI) with a linear first-stage where DC is instrumented for using the default pension plan type based on age in 2002 and probit second stage. Average partial effects are reported. Robust standard errors are adjusted for first-stage estimation. P-value for H_0 reported for evaluating implication of Equation 7. Demographic controls include gender, race, a cubic in tenure dummies, hours worked per year, and base pay rate. * Significantly different at the 10% level; ** at the 5% level; *** at the 1% level.

Table 3.7
Effect of DC Plan on Three-Year Leave Probability and Test for Selection, Vested Sample

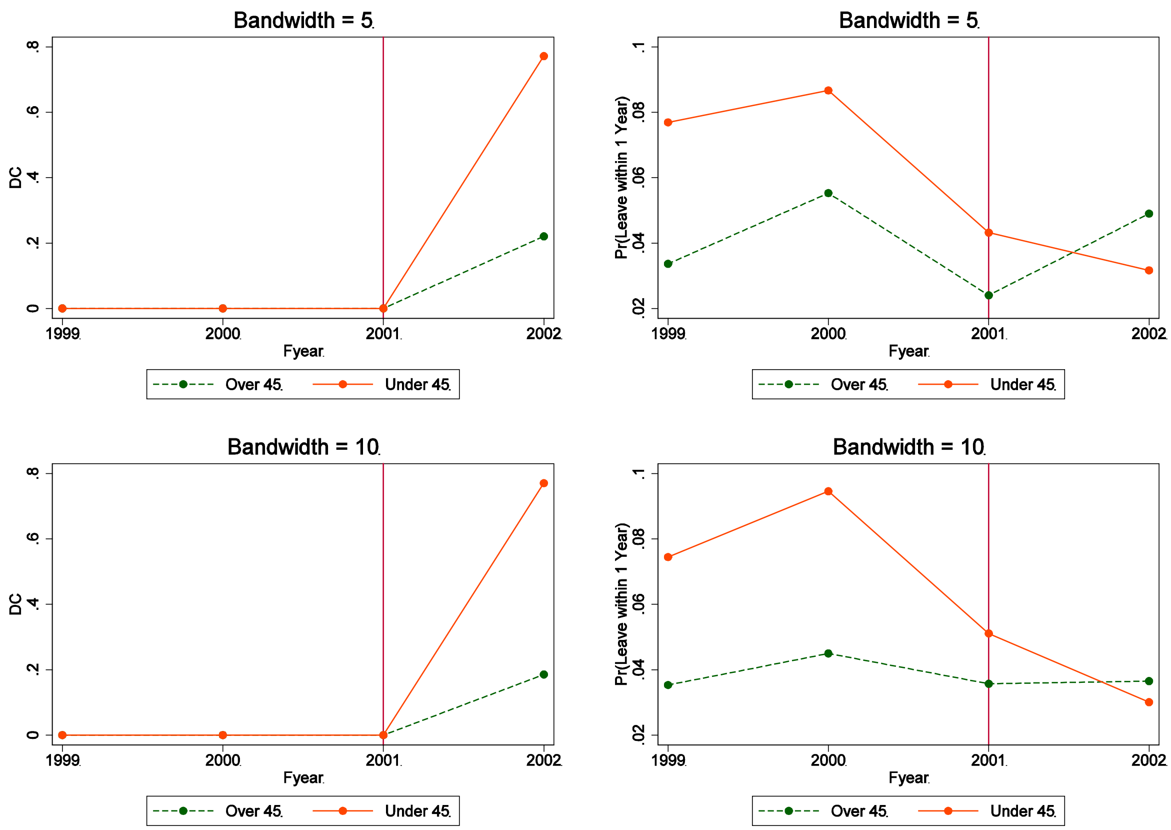
	(1)	(2)	(3)
β_{Endo}	-0.056**	-0.051*	-0.052**
	(0.028)	(0.028)	(0.028)
β_{Exog}	-0.149***	-0.152***	-0.148***
	(0.034)	(0.031)	(0.033)
H_0: $\beta_{Endo} \leq \beta_{Exog}$	0.001	0.000	0.001
$E[L_i]$	0.171	0.169	0.169
Controls	No	Yes	Yes
Age FEs	No	No	Yes
Year FEs	No	No	Yes
N	1,361	1,354	1,354
First Stage F-stat	380	372	361

Note: Sample includes employees in the years 1999 and 2002 who have at least 5 years of service. β_{Endo} estimates are from a simple probit regression. β_{Exog} estimates are calculated using Two-Stage Residual Inclusion (2SRI) with a linear first-stage where DC is instrumented for using the default pension plan type based on age in 2002 and probit second stage. Average partial effects are reported. Robust standard errors are adjusted for first-stage estimation. P-value for H_0 reported for evaluating implication of Equation 7. Demographic controls include gender, race, a cubic in tenure dummies, hours worked per year, and base pay rate. * Significantly different at the 10% level; ** at the 5% level; *** at the 1% level.

C. Robustness to Bandwith

Figure 3.1

DC Plan Enrollment and Probability of Leaving within One Year by Default Assignment: 1999-2002, Alternative Bandwidths



(a) DC Plan Enrollment

(b) Probability of Leaving in One

Notes: Over 45 represents employees age 45 or older on September 1, 2002. Under 45 represents employees younger than age 45 on September 1, 2002. Employees over 45 were defaulted to remain in the DB plan for 2002 and later, while employees under 45 were defaulted to switch to the DC plan.

Table 3.8

Effect of DC Plan on One-Year Leave Probability and Test for Selection, Alternative Bandwidths

	(1)	(2)	(3)	(4)	(5)	(6)
β_{Endo}	-0.009	-0.009	-0.008	-0.014	-0.014	-0.014
	(0.023)	(0.022)	(0.023)	(0.017)	(0.017)	(0.016)
β_{Exog}	-0.057**	-0.058***	-0.052**	-0.046***	-0.046***	-0.042**
	(0.022)	(0.022)	(0.022)	(0.017)	(0.016)	(0.017)
H_0: $\beta_{Endo} \leq \beta_{Exog}$	0.048	0.043	0.059	0.056	0.049	0.072
$\mathrm{E}[L_i]$	0.052	0.052	0.052	0.054	0.054	0.054
Controls	No	Yes	Yes	No	Yes	Yes
Age FEs	No	No	Yes	No	No	Yes
Year FEs	No	No	Yes	No	No	Yes
Bandwidth	5	5	5	10	10	10
N	1,499	1,499	1,499	2,584	2,584	2,584
First Stage F-stat	155	155	156	320	321	323

Note: Sample includes employees in the years 1999 - 2002. β_{Endo} estimates are from a simple probit regression. β_{Exog} estimates are calculated using Two-Stage Residual Inclusion (2SRI) with a linear first-stage where DC is instrumented for using the default pension plan type based on age in 2002 and probit second stage. Average partial effects are reported. Robust standard errors are adjusted for first-stage estimation. P-value for H_0 reported for evaluating implication of Equation 7. Demographic controls include gender, race, a cubic in tenure dummies, hours worked per year, and base pay rate. * Significantly different at the 10% level; ** at the 5% level; *** at the 1% level.

Appendix 4: Local Average Response Function (LARF) Results

A. Overview of Method

We consider an alternative approach to addressing endogenous regressors in the context of a nonlinear, probit specification. In particular, we apply the Local Average Response Function (LARF) approach to our context (Abadie, 2003). We recast our econometric model within a potential outcomes framework. Let $Y(D(Z), Z)$ be a binary outcome of interest, which is a function of a binary treatment variable, $D(Z)$ and a binary instrument, Z. Define Y_1 and Y_0 as the potential outcomes, as a function of the treatment variable D. For a given individual, the observed outcome is $Y = D \cdot Y_1 + (1 - D) \cdot Y_0$. Likewise, define the potential treatments as D_1 and D_0, which are functions of the instrument Z. For a given individual, the observed treatment status is $D = Z \cdot Y_1 + (1 - Z) \cdot Y_0$. Let X be a vector predetermined covariates. Using the shorthand $Y_{dz} = Y(d, z)$ for the potential outcomes, we make the following assumptions regarding the instrument, Z:

Assumption 4.1 (LARF)

4.1.1. *Independence: Conditional on X, the random vector* $(Y_{00}, Y_{01}, Y_{10}, Y_{11}, D_0, D_1)$ *is independent of* Z

4.1.2. *Exclusion:* $\Pr(Y_{d1} = Y_{d0} \mid X) = 1$ *for* $d \in \{0,1\}$

4.1.3. *First Stage:* $0 < \Pr(Z = 1 \mid X) < 1$ *and* $\Pr(D_1 = 1 \mid X) > \Pr(D_0 = 1 \mid X)$

4.1.4. *Montonicity:* $\Pr(D_1 \geq D_0 \mid X) = 1$

Abadie (2003) shows that the instrument Z can be used to estimate a Local Average Response Function (LARF). We briefly sketch the results from Abadie (2003) and apply them to our specific context. Let the average response function be $\mathbb{E}[Y(D)|X]$ — that is, the average relationship between the expected outcome and treatment variable. We define the LARF as $E[Y(D) \mid X, D_1 > D_0]$ — that is, the average response among the complier subpopulation, or the group for whom $D_1 > D_0$. Consider the following weight κ:

(4.18) $$\kappa = 1 - \frac{D(1-Z)}{\Pr(Z=0|X)} - \frac{(1-D)Z}{\Pr(Z=1|X)}$$

Let $g(Y, D, \underline{X})$ be a general function with bounded expectation. Abadie (2003) proves that under the assumptions above, we have the following:

(4.19) $$\mathbb{E}[g(Y,D,X)|D_1 > D_0] = \mathbb{E}\left[\frac{\kappa}{\mathbb{E}[\kappa]} \cdot g(Y,D,X)\right]$$

In words, we can estimate any statistical moment among the subpopulation of compliers by using a weighted expectation over the entire population. Intuitively, the LARF generalizes the classic Local Average Treatment Effect (LATE) (Imbens and Angrist 1994) to a broad class of nonlinear models. Indeed, if we were to model our outcome using a linear probability model the LARF and 2SLS LATE estimates are identical.

In our context, we assume the local average response function takes on a probit form:

(4.20) $$\mathbb{E}[Y(D)|X, D_1 > D_0] = \Phi(\lambda_{LARF} D + \Gamma_{LARF} X)$$

The parameters of interest maximize a probit likelihood function among compliers:

(4.21) $$(\lambda_{LARF}, \Gamma_{LARF}) = argmax_{\lambda,\Gamma}\mathbb{E}[Y \ln\Phi(\lambda D + \Gamma X) - (1 - Y)\ln\big(1 - \Phi(\lambda D + \Gamma X)\big)|D_1 > D_0.]$$

We cannot estimate the sample analog of Equation 4.21 because we do not simultaneously observe D_1 and D_0. However, using the result above in Equation 4.19, we can yet recover the parameters as follows:

(4.22) $$(\lambda_{LARF}, \Gamma_{LARF}) = argmax_{\lambda,\Gamma}\mathbb{E}\left[\frac{\kappa}{\mathbb{E}[\kappa]}(Y \ln\Phi(\lambda D + \Gamma X) + (1 - Y)\ln\big(1 - \Phi(\lambda D + \Gamma X)\big))\right]$$

The resulting parameters can thus be used to calculate an average partial effect among the compliers:

(4.23) $$\beta_{LARF} = \mathbb{E}[\Phi(\lambda_{LARF} + \Gamma_{LARF}X) -. \Phi(\Gamma_{LARF}X)|D_1 > D_0] = \mathbb{E}\left[\frac{\kappa}{\mathbb{E}[\kappa]}(\Phi(\lambda_{LARF} + \Gamma_{LARF}X) - \Phi(\Gamma_{LARF}X))\right]$$

In practice the weight κ, and in particular $Pr(Z = 1|X)$ must be estimated in a first stage. We specify a linear model as follows:

(4.24) $$Z = \pi X + \mu.$$

We then estimate the LARF parameters using the sample analog of Equation 4.22. Inference is performed accounting for the fact that the weight κ is estimated in a first stage (Newey and

McFadden 1994). The average partial effect in Equation 4.23 now holds a causal interpretation — the LARF function returns the *ceteris paribus* effect of variation in D on Y among a consistent group, the compliers, and therefore does not suffer from the selection bias that confounds the naïve, endogenous regression in Equation 10 in the text.

In our context, the outcome of interest is L_i, the endogenous regressor is DC_i, the set of controls are ($Post2002$i, $Under45_i$, X_i) and the instrument is $Post2002_i \times Under45_i$. We approximate the endogenous average partial effect as before using the naïve probit in Equation 10 and the exogenous average partial effect with β_{LARF}. We then test the key inequality in Equation 7. Inference is adjusted to account for the sample correlation between these two parameters.

B. LARF Results

Table 4.9

Effect of DC Plan on One-Year Leave Probability and Test for Selection, LARF Estimates

	(1)	(2)	(3)
β_{Endo}	-0.022	-0.020	-0.020
	(0.015)	(0.015)	(0.015)
β_{Exog}	-0.063***	-0.060***	-0.056***
	(0.013)	(0.014)	(0.015)
H_0: $\beta_{Endo} \leq \beta_{Exog}$	0.001	0.002	0.005
$\mathbb{E}[L_i]$	0.077	0.077	0.077
Controls	No	Yes	Yes
Age FEs	No	No	Yes
Year FEs	No	No	Yes
N	4,153	4,134	4,134
First Stage F-stat	622	622	626

Note: Sample includes employees in the years 1999 - 2002. β_{Endo} estimates are from a simple probit regression. β_{Exog} estimates are calculated using Local Average Response Function (LARF) with a linear first-stage and probit second stage. Average partial effects are reported. Robust standard errors are adjusted for first-stage estimation. P-value for H_0 reported for evaluating implication of Equation 7. Demographic controls include gender, race, a cubic in tenure dummies, hours worked per year, and base pay rate. * Significantly different at the 10% level; ** at the 5% level; *** at the 1% level.

Table 4.10
Effect of DC Plan on Two-Year Leave Probability and Test for Selection, LARF Estimates

	(1)	(2)	(3)
β_{Endo}	-0.024	-0.020	-0.021
	(0.023)	(0.023)	(0.023)
β_{Exog}	-0.121***	-0.120***	-0.116***
	(0.018)	(0.018)	(0.019)
H_0: $\beta_{Endo} \leq \beta_{Exog}$	0.000	0.000	0.000
$\mathrm{E}[L_i]$	0.146	0.146	0.146
Controls	No	Yes	Yes
Age FEs	No	No	Yes
Year FEs	No	No	Yes
N	3,137	3,123	3,123
First Stage F-stat	622	623	623

Note: Sample includes employees in the years 1999, 2000 and 2002. β_{Endo} estimates are from a simple probit regression. β_{Exog} estimates are calculated using Local Average Response Function (LARF) with a linear first-stage and probit second stage. Average partial effects are reported. Robust standard errors are adjusted for first-stage estimation. P-value for H_0 reported for evaluating implication of Equation 7. Demographic controls include gender, race, a cubic in tenure dummies, hours worked per year, and base pay rate. * Significantly different at the 10% level; ** at the 5% level; *** at the 1% level.

Table 4.11

Effect of DC Plan on Three-Year Leave Probability and Test for Selection, LARF Estimates

	(1)	(2)	(3)
β_{Endo}	-0.033	-0.029	-0.029
	(0.025)	(0.025)	(0.025)
β_{Exog}	-0.168***	-0.170***	-0.164***
	(0.027)	(0.026)	(0.027)
H_0: $\beta_{Endo} \leq \beta_{Exog}$	0.000	0.000	0.000
$\mathrm{E}[L_i]$	0.187	0.187	0.187
Controls	No	Yes	Yes
Age FEs	No	No	Yes
Year FEs	No	No	Yes
N	2,040	2,039	2,039
First Stage F-stat	621	624	616

Note: Sample includes employees in the years 1999 and 2002. β_{Endo} estimates are from a simple probit regression. β_{Exog} estimates are calculated using Local Average Response Function (LARF) with a linear first-stage and probit second stage. Average partial effects are reported. Robust standard errors are adjusted for first-stage estimation. P-value for H_0 reported for evaluating implication of Equation 7. Demographic controls include gender, race, a cubic in tenure dummies, hours worked per year, and base pay rate. * Significantly different at the 10% level; ** at the 5% level; *** at the 1% level.

[i] In Appendix 1.C we relax the assumptions that ensure $\beta_1 < 0$ and show that our main result still holds.

[ii] In fact, if we had not allowed any friction, then our model would generate the unrealistic prediction that no one who enrolls then leaves the firm, as it would not be optimal to pay the cost of enrolling knowing that one would be leaving the firm.