# BLS WORKING PAPERS

U.S. Department of Labor
U.S. Bureau of Labor Statistics
Office of Compensation and Working Conditions

## Matching Matters in 401(k) Plan Participation

**Keenan Dworak-Fisher**, U.S. Bureau of Labor Statistics

Working Paper 434
February 2010

# Matching Matters in 401(k) Plan Participation

**February 2010**

**Keenan Dworak-Fisher**
**Research Economist**
**Compensation Research and Program Development Group**
**Office of Compensation and Working Conditions**
**U.S. Bureau of Labor Statistics**
**Dworak-fisher.keenan@bls.gov**

## Abstract

This study offers new evidence on the effects of the matching contributions made by employers to 401(k) plan accounts on plan participation rates, exploiting microdata from the National Compensation Survey, a large, nationally representative, establishment dataset. It addresses the potential endogeneity of the matching contributions by employing coworker and labor market characteristics as instruments. The results indicate that employer matches have substantial effects. They also indicate that higher match rates tend to be correlated with workers having lower propensities to save; correcting for this endogeneity produces estimates that are bigger than those seen through direct cross-sectional comparisons.

**Acknowledgments:** I would like to thank John Bishow, Brandy Dickerson, Mike Lettau, and Erin McNulty for help in understanding and assembling the data; and Al Blostin, Julia Coronado, Maury Gittleman, Mark Loewenstein, Brooks Pierce, and Bill Wiatrowski, and seminar participants at the University of North Carolina – Greensboro and the 2005 AEA-ASSA meetings for helpful comments on earlier versions of this paper.

## Introduction

How do the provisions of a 401(k) plan affect the participation rates of employees? As traditional pensions continue their long decline and various changes to Social Security are contemplated, this question is increasingly crucial to those concerned about the sufficiency of retirement savings among US workers. In 2003, 39.9 percent[1] of U.S. private industry workers had access[2] to a 401(k) plan in which employees must contribute to participate and employers matched some of those contributions. Yet, only 67.9 percent of those with access to these plans participated; among lower-paid workers, the take-up rate was even lower (59.6 percent). These facts feed the concern that many workers may be saving too little for retirement and strengthen the imperative for plan provisions that promote participation effectively.

The literature on the effects of 401(k) plan design has produced conflicting accounts of how 401(k) plan provisions affect participation. One picture of these effects has been portrayed in a number of papers by Choi, Laibson, and Madrian (2004). Primarily exploiting an extensive administrative database collected by Hewitt Associates, a large human resources consulting company, these authors have found that a significant fraction of workers act passively with regard to their 401(k)-related saving decisions.

---

[1] Author's calculations using the National Compensation Survey (NCS) microdata collected from newly-initiated NCS sample members in 2003.

[2] Access to a benefit plan is defined in the National Compensation Survey according to the presence of a plan in the job/establishment pair; some workers are defined as having access even if they do not meet the applicable eligibility requirements.

Consistent with this behavior, Choi, *et al* find that the rate at which employers match employee contributions has a small, if any, effect on participation; the impact of automatic enrollment is much greater. Beshears, Choi, Laibson, and Madrian (2007) have additionally argued that the presence of an automatic enrollment provision diminishes the need for employers to provide generous matches. Yet, a different set of results has emerged when researchers have exploited an extensive, administrative data set from The Vanguard Group, a large investment management company. Huberman, Iyengar and Jiang (2007) and Mitchell, Utkus, and Yang (2005) both find that employer match rates significantly increase 401(k) participation.

The variety of results emerging from these non-representative, administrative datasets underscores the need for evidence from representative samples. But such evidence has been limited by deficiencies in the available datasets. In particular, the challenge of obtaining accurate and complete information about the relevant plan parameters is a high obstacle for representative surveys to clear. Household surveys like the Current Population Survey generally rely on workers to report the details of their retirement plans, resulting in a large error rate.[3] Establishment-based sources such as the Form 5500 tax data filed by employers with the IRS do not generally seek information about plan provisions in a very detailed manner. And special interest surveys such as the Health and Retirement Survey that obtain both establishment and household data generally limit their scope to a subset of the relevant population, making generalization to all relevant workers an uneasy proposition.

---

[3] See, e.g., Herz, Meisenheimer and Weinstein (2000) and Chan and Stevens (2008).

In this study, a large, nationally representative dataset from the National Compensation Survey is exploited to provide measures of the effects of 401(k) plan provisions on the participation rates of employees. The dataset contains accurate measures of plan details, including those governing employer matches, and its coverage is very broad, having been sampled to represent virtually all US private industry workers in 2002-2003. The completeness of its description of plans' match provisions allows the study to explore functional forms that offer separate measures of causal and sorting effects. The dataset also contains information about coworkers working at the same establishment, as well as the generosity of retirement benefits among other employers in the same labor market. These additional pieces of information are used to construct instrumental variables estimates of the causal effect of employer matches on plan participation. Together, the results of these various inquiries provide evidence in support of studies using data from The Vanguard Group: the rate at which employers match employee contributions has a significant, positive effect on plan participation. Further, the results support the findings of Even and MacPherson (2005) and others in the literature who have concluded that direct cross-sectional comparisons may *under*-estimate these causal effects due to negative selection arising from employers' efforts to remediate the behavior of low-saving workers.

## Data

The data come from the National Compensation Survey (NCS), a large, nationally representative survey conducted by the U.S. Bureau of Labor Statistics. Data from the NCS is used to calculate the Employment Cost Index, which estimates the growth in compensation costs, including those arising from employer-provided benefits, for a fixed bundle of workers. The NCS is collected with a rotating panel design, with a new panel initiated approximately once per year. When a panel is initiated, brochures for employers' benefit plans are collected along with the employer cost and benefit participation information. The details of these plan brochures are coded into the NCS database, and the incidence of various detailed plan provisions are reported in official bulletins. This study uses NCS microdata from the respondents initiated in 2002 and 2003, focusing on the detailed provisions data collected from 401(k) plan brochures and the contemporaneous participation data collected from the corresponding establishments.

The NCS microdata are collected at the job level: within each sampled establishment, a small number of narrowly defined jobs are selected.[4] The resulting wage, benefit costs, and participation data consist of averages among the employees at the establishment having that job description. Jobs are defined at as detailed a level as possible, identifying a specific set of job duties, required skills, and responsibilities. Each job corresponds roughly to a 6-digit occupational unit within the sampled establishment, further narrowed down so that all workers in the job have the same union and full-time/part-time statuses, the same pay basis (time vs. incentive), and the same benefit offerings. This taxonomy does not guarantee that workers in a given job work together, but it means that all workers in a job perform very similar functions. For example, an establishment might employ various classes of Accountants in a number of different divisions (accounts receivable, accounts payable, etc.), where each class indicates a different level of responsibility. The NCS would classify each class of

[4] Depending on the size of the establishment, between 4 and 8 jobs are sampled.

accountant as a separate job, but would not necessarily specify different jobs for different divisions.

Most job-level observations in the NCS correspond to a small number of individuals, but there are some exceptions, mostly accruing to necessary compromises made in collection. For example, some employers may make their compensation available only at more aggregate levels, necessitating job quotes that have large numbers of workers. An investigation of 2007 data showed that the median number of workers in a job was 5, and the third quartile worker count was 22. But a few large quotes drove the mean up to 90.2. Within jobs, quotes are relatively homogenous in compensation offerings; in 2007, 97.3 percent of the wage variance among individual workers in the NCS sample occurred between job quotes.

The focus of this study is on one variant of 401(k) plans: the savings and thrift. Such plans entail voluntary (tax deductible) contributions by the employee that are matched to some extent by the employer. This is easily the most prevalent form of 401(k) plan, making up more than 80 percent of 401(k) plans in which the employer made some contributions in 2002-2003.[5] Not included in the study are plans to which employers make no contributions, which are also fairly prevalent.[6]

---

[5] Author's calculations using the National Compensation Survey (NCS) microdata collected from newly-initiated NCS sample members in 2002 and 2003.

[6] In 2005, an estimated 16 percent of private industry workers had access to cash deferred arrangements with no employer contributions. These are not considered to be retirement benefit plans by the BLS, (BLS Summary 05-01) so their participation rates are not

Table 1 provides some summary statistics about the plans in the sample. The average participation rate in the sample, defined as the fraction of workers in the job that participate in the plan is .72. This variable can generally be considered a take-up rate: almost all employees in a job with access to the plan are eligible, although some plans have eligibility requirements based on months of service. 82 percent of the sample is made up of plans with flat match rate profiles – one percentage is applied to each employee's contributions, up to a specified percentage of the employee's salary. Another 13 percent apply a variable match rate, where employees receive one match rate to a first amount of their contributions and another (usually lower) rate on additional contributions, up to some limit. The remainder of the sample has different match profiles for different employees within a job, depending on the employees' tenure.[7] The average match rate on the first dollar contributed by employees is 75.37 percent; the log of this first-dollar match averages 4.21 with a standard deviation of 0.49.[8] Plans in the sample provided matches on employee contributions up to 5.16 of the corresponding salary, on average.

---

collected. For more details about these "zero-match" plans, see section 9.5 of Holmer, Janney and Cohen (2008).

[7] A small number of plans have matches that vary from year to year, depending on employer profits or the employer's discretion; these were dropped from the sample.

[8] To calculate measures among the plans whose match profiles vary by tenure, the tenure distribution of each corresponding record was imputed based on the available direct information and detailed occupational averages. The match provisions were then averaged across these imputed distributions.

To capture the overall generosity of the plans, we calculated the "total potential match" – the amount that employers contribute, as a percentage of wages, when employees contribute enough to exhaust the employer's match offer. For example, if a plan offers a 50% match on the 6 percent of wages the employee contributes, then the total potential match is 3 percent. The average total potential match in the sample is 3.57 percent of salary, while the log of the total potential match has a mean of 1.12 and a standard deviation of 0.58.

The data give a fairly complete picture of the benefit offerings of the employers in the sample, including other salient features of the 401(k) plans themselves. A majority of the sample allows employees to choose how the funds in the 401(k) account are invested, with control over funds contributed by the employee slightly more prevalent than control over funds contributed by the employer. A small percentage of the observations (6 percent) in the sample are governed by the automatic enrollment provisions advocated by Choi, *et al*. Meanwhile, 40 percent of the jobs indicate that they also provide a defined benefit plan, while 21 percent provide an additional defined contribution plan.[9] The average observation has a wage of $22.66 per hour, a health benefit costing the employer $2.21 per hour worked, and a defined benefit cost of $0.52 per hour worked. The data also contain detailed (6-digit) occupation and industry identifiers, as well as the location and employment of the establishments and whether workers in the job are unionized.

[9] A very small fraction of sample members have more than one savings and thrift plan. In such cases, we focus only on the plan that had the highest participation rate.

The sample consists of 2,708 jobs in 587 establishments[10], with 67 percent of jobs observed in 2003 and the rest observed in 2002.

## Empirical Analysis

We begin by laying out a simple model to describe some key aspects of the relationship between plan participation and the employer match on 401(k) contributions. We then use the model to describe the identifying assumptions inherent in the various empirical strategies we employ.

Consider the participation decisions of workers in a given establishment offering a plan with given provisions. Let the matching provisions at employer k be defined by one generosity parameter, $M_k$. Let other observed characteristics of the employer, such as other provisions of its 401(k) plan, be denoted as $E_k$, and let relevant, unobserved characteristics of the employer, such as its culture as regards retirement saving, be denoted as $c_k$. Workers in job j at employer k determine whether or not to participate in the plan according to $M_k$, $E_k$, $c_k$, and their own attributes – both observed attributes such as their income levels, denoted as $X_{jk}$, and unobserved attributes such as their innate attitudes toward retirement saving, denoted as $a_{jk}$. Letting $P_{jk}$ be the participation rate of workers in job j of establishment k, allow expected participation to have the form:

[10] This sample reflects all NCS sample members initiated in 2002 or 2003 for which valid data on match rates and participation were collected, with 1 establishment dropped due to outlying benefit cost values. Among establishments appearing in the sample in both years, only 2003 data were used.

$$E\left[P_{jk}\right] = \Phi(\beta_0 + \beta_1 \cdot a_{jk} + \beta_2 \cdot X_{jk} + \beta_3 \cdot c_k + \beta_4 \cdot E_k + \beta_5 \cdot M_k) \tag{1}$$

where Φ denotes the cumulative Normal distribution function.

Now consider the determination of $M_k$, which we divide into two steps: first the overall generosity of the plan's match provision is set; then the structure of the provision is determined. As discussed above, a good measure of a plan's generosity is the total potential match that a worker can receive by contributing up to the match threshold. We denote the total potential match offered by employer k as $T_k$.

Several factors affect the relationship between $T_k$ and the worker and establishment characteristics $a_{jk}$, $X_{jk}$, $E_k$, and $c_k$. Most saliently, workers with high propensities to save may disproportionately choose to work at establishments with high levels of plan generosity. Accordingly, some employers may purposely set their plan generosity levels to attract and maintain these workers (Ippolito, 2002). This sorting dynamic is worked out within the context of the alternatives available to workers; the same sorting outcome for establishment k would be expected to require a higher plan generosity, the higher the generosity of other employers in the same labor market. Letting $\overline{a}_k$ and $\overline{X}_k$ represent averages across the establishment of $a_{jk}$ and $X_{jk}$, we represent this relationship as:

$$T_k = \gamma_0 + \gamma_1 \cdot \overline{a}_k + \gamma_2 \cdot \overline{X}_k + \gamma_3 \cdot c_k + \gamma_4 \cdot E_k + \gamma_5 \cdot O_k + \gamma_e\,, \tag{2}$$

where $O_k$ represents the generosity of other employers in the same labor market as employer k.

Next, the profile of the plan's employer match is derived. Will a given level of plan generosity be achieved by a relatively high match rate applied to a relatively small range of employee contributions, or by a relatively low match rate applied to a relatively large range of employee contributions? The factors driving this determination are somewhat less clear than the sorting dynamic governing $T_k$. We explore two different possibilities. The first is that the first dollar match rate $M_k$ is random with respect to the other variables in the model:

$$M_k = \phi_0 + \phi_1 \cdot T_k + \phi_e \tag{3a}$$

Here, we assume that, once the employer chooses an overall generosity level, the marginal match rate(s) and corresponding range of employee contributions to which it (they) apply is chosen according to the idiosyncratic tastes of decision-makers, historical accident, or some other extraneous considerations.

The second possibility we consider is that $M_k|T_k$ is set to elicit a particular savings response from workers at the establishment; e.g., to raise savings among some or all workers. Perhaps the employer has a target participation rate, reflecting either its long-term budgeting goals or its need to meet the non-discrimination requirements governing 401(k) plans under ERISA. In this case, $M_k$ may be better captured by:

$$M_k = \phi_0 + \phi_1 \cdot \bar{a}_k + \phi_2 \cdot \bar{X}_k + \phi_3 \cdot c_k + \phi_4 \cdot E_k + \phi_5 \cdot T_k + \phi_e . \tag{3b}$$

Unlike in equation (2), we expect that the coefficients on $\bar{a}_k$ and $\bar{X}_k$ may be negative: if the first-dollar match rate has a significant effect on participation rates, the employer might choose a steeper profile the lower the savings propensity its employees to have. Such a dynamic is consistent with other components of retirement plan management documented elsewhere. For example, Bernheim and Garrett (2003) and Bayer, Bernheim and Scholz (1996) provide evidence that a remedial impetus is prevalent for employer-provided financial education programs. Madrian and Shea (2001) note that the employer they study adopted its automatic enrollment provision because it was having trouble meeting non-discrimination standards.

The model may be consulted in a number of ways to measure the causal effect of the first-dollar match $M_k$ on plan participation $P_{jk}$. One approach is to estimate equation (1) directly, including as many controls for the important worker and establishment characteristics as possible. We pursue this strategy by implementing the Bernoulli Quasi-Maximum Likelihood Estimator (BQMLE) developed by Papke and Wooldridge (1996), using the log of the first dollar match rate as the key explanatory variable. The BQMLE assumes that the expected value of $P_{jk}$ is captured by the standard normal cumulative density function conditional on the specified explanatory variables ($Z_{jk}$):

$$E(P_{jk} \mid x_{jk}) = \Phi(Z_{jk}\beta) \tag{4}$$

and is computed by maximizing

$$\ell_{jk}(b) = P_{jk}\log[\Phi(Z_{jk}b)] + (1 - P_{jk})\log[1 - \Phi(Z_{jk}b)]. \tag{5}$$

Estimating plan participation using the BQMLE has the advantage of dealing appropriately with fractional dependent variables having masses in the distribution at 0 and 1. It results in coefficient estimates similar to those resulting from probit analysis; we transform these coefficients into average partial effects estimates (APEs). As discussed by Wooldridge (2005), APEs are the logical item of interest for standard analytical frameworks, corresponding intuitively to OLS estimates. In addition, focusing on APEs allows us to avoid concerns over attenuation biases accruing from unobserved heterogeneity.

Table 2 gives the APEs when the BQMLE is applied directly to the cross-section. In the first two columns, the estimates reflect a direct application of equation (1), with observable job characteristics, meant to stand in for worker attributes $X_{jk}$, and observable employer characteristics $E_k$ included as controls. Since this type of specification is prevalent in previous work, we refer to it as the "base model." The controls include a dummy for whether the job is unionized, dummies for 9 occupational groups, the average compensation paid workers in the job, and the average compensation squared, as well as establishment size, and other provisions of the 401(k) plan. Also included are imputed values for four demographic variables: the average age of workers in the job, the percentage of workers who are male, the percentage having graduated from college, and the percentage who are white.[11] In addition, controls accounting for the composition of workers' compensations have been included: the wage component of compensation, the health care component, the component associated with any Defined Benefit plan present for the job, and a dummy indicating whether workers in the job have access to another Defined Contribution plan. These controls are entered as proxies for workers' unobserved savings propensity $a_j$. The logic of these proxies is that high savers are more oriented towards minimizing future risk, so for a given compensation level they are likely to prefer other benefits such as health insurance instead of wage. Finally, broad controls for 4 geographic regions and 4 industrial groupings are

[11] These imputations were generated by matching the detailed (3-digit) industry and occupation information, along with the observed wage rate in the job, to 2002 Current Population Statistics data and using regression analysis to predict values for each job.

contemplated as controls for unobserved employer characteristics $c_j$. We include two specifications: one with and one without region and industry dummies.[12]

The results of the base model indicate a small but significant effect of employer match rates on plan participation, with an APE of .0567 or .0580. This means that, on average, a one standard deviation change in the log match rate corresponds to a 2.8 or 2.9 percentage point increase in participation. These results might be questioned, however, to the extent that we have not been able to control for all the relevant correlates that appear in both equations (1) and (2) – i.e., if significant variation in the unobserved factors $a_j$ and $c_k$ remain. Most saliently, if workers with high (unobserved) savings propensities sort into employers with higher match rates as envisioned by a positive value of $\gamma_1$ in equation (2), then this "sorting effect" will positively bias the base model's estimates of the causal effect of $M_k$ on $P_{jk}$.

Before exploring alternative strategies for estimating match effects, we note the measured effects of several other variables in the base model. Automatic enrollment provisions are associated with an increase in participation by 7.75 percentage points. This is likely a downwardly biased estimate of the causal effect of automatic enrollment, because in some sample members the provision may only apply to a fraction of the workers, such as those who have recently been hired.[13] Still, it is within the margin of

---

[12] Dummies for year of observation and eligibility requirements of at least 1 year of service are also included in both specifications.

[13] Note, however, that the sample also likely includes many plans in which the automatic enrollment provision has only recently been added. In such cases, the APEs measured do not reflect the long run effects of automatic enrollment, tending instead to be higher.

error of the 11 point increase that Madrian and Shea (2001) find studying one large employer. Providing workers with a choice of how to invest their own contributions has a small but significant, negative association with participation. This is consistent with the results of Iyengar, Jiang and Huberman (2003), who argue that too much choice can impart complexity costs that reduce plan enrollment. But having choice over the employer's contributions does not appear to have any appreciable effect on participation. Both of these APEs contradict Papke (2003), who finds dramatic positive effects. The ability to draw loans from one's account appears to have an insignificant effect on participation.

The included controls for compensation level and imputed demographic traits show results that are broadly consistent with the previous literature. The two characteristics that have most consistently been found to have positive, significant effects on participation are income and age; we replicate that here. Some evidence (Even and MacPherson, 2005; Englehardt and Kumar, 2007) also suggests that white workers are more likely to participate than are blacks; the base model results agree, albeit with large standard errors. Education and gender are sometimes found to be significant coorelates of 401(k) participation, but they have often been estimated to have insignificant effects in multivariate analyses (Munnell, Sunden and Taylor, 2001/2; Bernheim and Garrett, 2003; Mitchell, Utkus and Yang, 2005). The base model finds no significant effects of percent of workers having a college graduation or percent male.

Finally, consider the measured effects of the composition of workers' compensation, which have not been included in previous studies. The results show that having a higher health plan component of compensation is significantly associated with higher participation in one's 401(k). The presence of other Defined Contribution plans is also associated with higher participation. These results suggest some savings propensity-related job sorting on these two benefit categories. But even with these controls, the estimated effect of the employer match remains positive and significant.

A second approach to estimating the effects of the employer match rate is to add a control for the overall generosity of the plan. To see the basis for this approach, substitute equation (3a) directly into equation (1). The effect of $M_k$ on $P_{jk}$ is then estimated based on the residual variation in $M_k$ denoted as $\varphi_e$ in equation (3a). The results of pursuing this approach are shown in the third and fourth columns of Table 2. The main thing to notice about these results is that, controlling for overall plan generosity, the effect of the match on participation remains significant, although it is diminished by about a quarter: a one standard deviation change in the log match rate causes the participation rate to rise by 2.0 to 2.3 percentage points. The APEs of the other correlates are qualitatively similar to those seen in the base model. These results suggest that the positive association between participation and match rates seen in the base model embody, to a large extent, a causal relationship. But such a conclusion rests on the validity of the restriction embodied in equation (3a): that the determination of $M_k|T_k$ is random with respect to the unobserved variables $a_j$ and $c_k$.

We can relax this assumption by turning to equation (3b), which models $M_k|T_k$ as a function of observed and unobserved employer and employee characteristics. Consider the equation obtained by substituting equation (2) into equation (3b):

$$M_k = \xi_0 + \xi_1 \cdot \bar{a}_k + \xi_2 \cdot X_{jk} + \xi_3 \cdot \bar{X}_{\sim jk} + \xi_4 \cdot c_k + \xi_5 \cdot E_k + \xi_6 \cdot O_k + \gamma_e . \quad (6)$$

This equation contains two elements that do not appear in the participation equation (1). First, $P_{jk}$ depends on only the characteristics of job j workers $X_{jk}$ and not on the characteristics of other workers in the establishment, while $T_k$ and $M_k|T_k$ are both determined in relation to all of establish k workers. This distinction is made explicit in equation (6) by splitting $\overline{X}_k$ into own-job and coworker components, $X_{jk}$ and $\overline{X}_{\sim jk}$. Second, the labor market variables $O_k$ also do not affect the worker's participation decision despite playing a key role in the sorting equilibrium described by equation (2). These two elements can be used to implement an instrumental variables strategy of estimating the effects of $M_k$ on participation, although some fairly strong assumptions are required: the instruments cannot directly affect the plan participation of worker j, and they cannot be correlated with the unobserved variables $a_j$ and $c_k$.

Several measures of $\overline{X}_{\sim jk}$ were calculated from the data, reflecting the average compensation, age, proportion with a college degree, proportion male, and proportion white. To capture $O_k$, two variables were generated from NCS data measuring the average proportion of compensation paid to defined contribution plans among other employers in the corresponding labor market.[14] The first of these measures used geography to define the relevant labor market, taking advantage of the cluster sample design of the NCS, in which a small set of (predominantly metropolitan) areas is selected as primary sampling units. Within each of these areas, the average fraction of

[14] Measures of $O_k$ were calculated using the larger NCS dataset measuring employer costs for all units in the NCS panel, pooling NCS observations over the 2001-2005 period. The resulting sample for creating the instruments contained 6768 employers.

compensation spent by employers on Defined Contribution plans was calculated. The second measure of $O_k$ was calculated similarly, but using 2-digit industry definitions as the relevant labor market concept.

These instruments were applied to the data using the instrumental variables methodology described Wooldridge (2005). First, equation (6) was estimated using OLS, and the residuals, $\hat{\eta}_e$, were added to the base BQMLE model. Then, the corrected standard errors were estimated using the methodology described in Papke and Wooldridge (2008). This methodology also readily allows testing of the validity of the instrumental approach: standard t-tests (using the corrected standard errors) can be applied to the estimated coefficient on $\hat{\eta}_e$. The resulting APEs are listed in the top panel of Table 3. The top row contains the APEs of the first-dollar match rate, and the second row contains the APEs of the first-stage residuals, which measure the effects of endogenous variation in $M_k$. The results for different sets of instruments are shown: the co-worker instruments $\overline{X}_{\sim jk}$, the labor market instruments $O_k$, and the combination of all instruments. As in Table 2, we consider specifications with and without broad controls for area and industry.[15]

All of the APEs of the first-dollar match rate in the table are greater than the base model estimates, and all are statistically significant. Using the co-worker instruments alone, the estimated APE is .091 when broad area and industry controls are not included and .199 when they are included, indicating that a one standard deviation increase in the

[15] These estimates were generated using a slightly reduced sample of 2,372 observations in 464 establishments; establishments with data for only one collected job are excluded.

match rate increases participation by 4.5 to 9.8 percentage points. Using only the labor market instruments, the estimated APEs are substantially greater: the APE of .426 when broad area and industry controls are included implies a 21.1 percentage point effect from a one standard deviation increase in $M_k$. Using all co-worker and labor market instruments, the estimated APEs are similar to those using only the coworker instruments, albeit with a slightly larger range.

A salient feature of the results in Table 3 is that the estimated APEs are markedly greater when broad controls for area and industry are included in the specification. This indicates some negative covariation between match rates and participation at these broad levels. Even and MacPherson (2005) provide a potential explanation: they note that a negative bias in the cross-section can occur if workforce sectors having low average savings propensities prompt employers in those sectors to employ higher matches to attract workers at the top of the savings propensity distribution. If this interpretation is accurate, then it is appropriate to include the controls in our estimation. Note, however, that we did not see this pattern in the base model results of Table 2; the sensitivity of the results to the controls may also reflect the reduced explanatory power of our instrumental variables approach. So while we prefer the specifications that include the broad controls, we continue to show the results in the remaining tables both with and without the broad controls.

The bottom panel of Table 3 provides information from the first stage of the estimation. The coefficients on the coworker instruments depict a mix of influences on the match rate. Workers with relatively well-compensated co-workers tend to have higher match rates, all else equal; this suggests that positive sorting as we might expect from equation (2) plays a key role. On the other hand, match rates also tend to be higher when co-workers are younger, and when they are non-white; this suggests an important role for the remedial impetus discussed earlier in regards to equation (3b). The coefficients on the labor market instruments suggest that employers consider the offerings of other employers in the same industry when determining the generosity of their retirement plans, but they indicate no significant role for geographic competitors. At the bottom of the table, the partial R-Squared and F-Test on the excluded instruments as discussed by Bound, Jaeger and Baker (1995) and Shea (1997) are listed. These indicate that the instruments are relatively weak, together explaining less than 2 percent of the residual variation in the match rate, but that they are strong enough to assuage concerns about finite-sample bias.

The increased magnitudes of the estimated match effects in Table 3 are notable. Sorting effects have been the main concern in the literature when match rate estimates have been considered, and the usual intuition is that such effects bias cross-sectional estimates upward.[16] But the results in Table 3 suggest that the cross-sectional estimates of our base model are negatively biased, if anything. In fact, in specifications including broad area and industry controls, the coefficients on the first-stage residuals indicate negative correlations between $M_k$ and $P_{jk}$ that are both substantively and statistically significant.

---

[16] Note, however, that some other studies (e.g., Even and MacPherson, 2005; Englehardt and Kumar, 2007) also find that addressing endogeneity in the relationship between $M_k$ and $P_{jk}$ results in higher estimated match effects.

One possible explanation is that the instruments are not valid – that they have a direct relationship with the dependent variable $P_{jk}$ as well as affecting $M_k$. Such a direct relationship could occur through peer effects – high savings propensities may be contagious. Duflo and Saez (2002) show that such networking effects are prevalent among co-workers who have frequent contact with each other. But Duflo and Saez also find that, when co-worker measures match dissimilar workers, networking effects become insignificant. We follow that intuition in Table 4, where we instrument with adjusted co-worker measures that are calculated only using co-workers who do not share the same (1-digit) occupation as the reference worker. If workers primarily interact with coworkers who perform similar functions, then peer effects in the participation decision should not be a factor here. It's also less likely that job-searchers would choose employment at a given establishment based on the characteristics of prospective coworkers in distant occupational categories. The cost of using these adjusted co-worker measures is that it reduces the power of the instruments by reducing the amount of data being used, diminishing the first stage R-squareds and increasing the standard errors. The APEs estimated with the adjusted instruments are smaller than their counterparts in Table 3. This suggests that some of the effects observed in Table 3 may have been due to peer effects, although the relatively high levels of imprecision make comparisons of this sort difficult. In our preferred specifications (including broad area and industry controls), the match rate APEs are positive, although due to large standard errors they are only statistically significant when both coworker and labor market instruments are employed. And these results are not robust to the exclusion of the area and industry controls.

Another avenue for the instruments to be invalid is if they are themselves correlated with the unobserved characteristics $a_{jk}$ and $c_k$. For example, in the case of the coworker instruments, positive assortativeness in the matching of workers to employers could generate a direct correlation between the characteristics of other workers at the establishment and the unobserved savings propensities $a_{jk}$. In the case of the labor market instruments, sectoral differences in plan generosity may proxy for sectoral differences in the establishment savings cultures $c_k$. Indeed, the significance of area and industry sectoral controls at a broad level suggests that such a dynamic may, in fact, be present. A fundamental problem is that equation (2) describes a relationship in which causality runs in both directions; this creates the possibility for feedbacks between the instruments and the dependent variable.

An identification strategy that alleviates these concerns is to focus on the variation in $M_k$ inherent in equation (3b). In this equation, it's easier to think of the right-hand side variables as being exogenous determinants of the match rate. To pursue this strategy, we control for $T_k$ and instrument for $M_k$ with $X_{\sim jk}$. Note that $O_k$ is not available as an instrument in this construct: it enters the model only through the determination of the control variable $T_k$. The results, shown in Table 5, reinforce those of Table 4. Without broad industry and area controls, the APEs of the employer match are statistically insignificant. But when the broad controls are included, the match rate APEs are positive, large, and significant. Using the adjusted coworker instruments, we estimate an APE of .261, indicating that a one standard deviation increase in the match rate yields an increase in plan participation of 12.9 percentage points. This specification also yields a significantly negative APE on the first stage residual, indicating a negative association between the match rate and the unobserved worker and establishment characteristics. These results indicate that the results in column (2) of Tables 3 and 4 were driven by variations in $M_k|T_k$. Accordingly, they support the notion that employers steepen the match profile to elicit additional plan participation when participation is otherwise low, causing cross-sectional estimates to be biased downward. These conclusions should be tempered, though, by their sensitivity to the inclusion/exclusion of area and industry controls.

## Conclusion

This study makes two key contributions to the literature on the effects of retirement benefit plan design. First, it exploits an underutilized source of data: the National Compensation Survey, a nationally representative survey that combines accurate, detailed information about 401(k) plans with participation rates collected at the detailed job level. Using this data helps shed light on the divergent results produced by the various non-representative samples exploited elsewhere. Second, the study implements several strategies for identifying the causal effects of employer matches on plan participation. These strategies provide several plausible corrections for the endogeneity of the match provisions and are placed within the context of a coherent model or employer and employee behavior.

The results of the study provide further evidence that employer matches provide a powerful incentive for employees to participate in their 401(k) plans, and that the level of the matches offered matters significantly. A one standard deviation increase in the match applied to the first dollar is found to raise participation by as much as 12.9 percentage points. This effect is similar in magnitude to the 11 point effect of the institution of an automatic enrollment provision as estimated found by Madrian and Shea (2001). A lesson that might be drawn from this result is that, while automatic enrollment provisions are a promising avenue for encouraging many workers to save for their retirements, the marginal incentives like employer matches that have been more traditionally offered may be even more important to some workers. In this sense, the study is in agreement with Duflo, *et al* (2006), who provide evidence that individuals respond to marginal saving incentives, but that important characteristics of the incentive program (e.g., level of complexity and accessibility) and the individual (e.g., income level and marital status) greatly affect the extent of the response.

A second lesson of the study is that the endogeneity of the match rates observed in practice is not so straightforward to characterize. It is often assumed that sorting effects predominate, so that generous matches are positively associated with workers with high predispositions to participate. But such a viewpoint overlooks another impetus for employers to offer generous matches: they may do so to encourage saving among workers with low predispositions to participate. In fact, whether it is motivated by a paternalist impulse or a desire to meet certain legal requirements imposed by ERISA regulations, this remedial impetus appears to be prevalent enough so that cross-sectional estimates of match effects are biased downwards, not upwards. In this sense, the findings corroborate those of Even and MacPherson (2005).

Several caveats to these results are worth noting. First, the identifying assumptions of the various strategies implemented in the study are in some cases fairly ambitious. While a fairly extensive set of controls is included for other correlates of plan participation, significant unobserved variation remains, and correlation between the instruments and these unobserved factors could affect the results. The study tries to address this concern by providing several alternative approaches and by elucidating the identification strategies within a model framework, but different readers may have lingering reservations about the results. Second, the results do not appear to be particularly robust to the exclusion of industry and area dummies. We believe that such variables do belong in the model as controls for sectoral variations in unobserved characteristics such as cultural attitudes toward saving, but the sensitivity of the empirical model to their inclusion might also be interpreted as a weakness. Finally, the results of this study leave several questions unaddressed. For example, increases in the participation of workers in their employer-provided 401(k) plan do not necessarily imply increases in retirement saving, either within the plans themselves or considering all savings vehicles together. But attempts to determine such impacts on the ultimate levels of retirement preparations should build into their consideration our main finding above: that matching does matter for at least some key components of employee saving.

**References**

Bayer, Patrick J., B. Douglas Bernheim and John Karl Scholz. 1996. "The Effects of Financial Education in the Workplace: Evidence from a Survey of Employers." Working Paper No. 5655. Cambridge, MA: National Bureau of Economic Research.

Bernheim, B. Douglas and Daniel M. Garrett. 2003. "The Effects of Financial Education in the Workplace: Evidence from a Survey of Households." *Journal of Public Economics* 87(7-8): 1487-1519.

Beshears, John, James J. Choi, David Laibson and Brigitte C. Madrian. 2007. "The Impact of Employer Matching on Savings Plan Participation Under Automatic Enrollment." Working Paper No. 13352. Cambridge, MA: National Bureau of Economic Research.

Bound, John, David A. Jaeger and Regina Baker. 1995. "Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variable Is Weak." J*ournal of the American Statistical Association* 90: 443-450.

Chan, Sewin and Ann Huff Stevens. 2008 "What You Don't Know Can't Help You: Pension Knowledge and Retirement Decision-Making." *Review of Economics and Statistics* 90(2): 253-266.

Choi, James J., David I. Laibson and Brigitte C. Madrian. 2004. "Plan Design and 401(k) Savings Outcomes." Working Paper No. W10486. Cambridge, MA: National Bureau of Economic Research.

Duflo, Esther and Emmanuel Saez. 2002. "Participation and Investment Decisions in a Retirement Plan: the Influence of Colleagues' Choices." *Journal of Public Economics* 85(1): 121-148.

Duflo, Esther, William Gale, Jeffrey Liebman, Peter Orszag, and Emmanuel Saez. 2006. "Saving Incentives for Low- and Middle-Income Families: Evidence from a Field Experiment with H&R Block." *Quarterly Journal of Economics* 121(4): 1311-1346.

Engelhardt, Gary and Anil Kumar. 2007. "Employer matching and 401(k) saving: Evidence from the health and retirement study." *Journal of Public Economics* 91(10): 1920-1943.

Even, William E. and David A. Macpherson. 2005. "The Effects of Employer Matching in 401(k) Plans." *Industrial Relations: A Journal of Economy and Society* 44(3): 525-549.

Herz, Diane E., Joseph R. Meisenheimer and Harriet G. Weinstein. 2000. "Health and Retirement Benefits: Data from Two BLS Surveys." *Monthly Labor Review* 123:3-20.

Holmer, Martin, Asa Janney and Bob Cohen. 2008. "Pensim Overview." *Mimeo*, U.S. Department of Labor, July 2008. http://www.polsim.com/overview.pdf.

Huberman, Gur, Sheena S Iyengar, and Wei Jiang. 2007. "Defined Contribution Pension Plans: Determinants of Participation and Contributions Rates." *Journal of Financial Services Research* 31(1):1-32.

Ippolito, Richard A. 2002. "Stayers as „Workers' and „Savers': Toward Reconciling the Pension-Quit Literature." *Journal of Human Resources* 37(2): 275-308.

Iyengar, Sheena S., Wei Jiang, and Gur Huberman, 2004. "How Much Choice Is Too Much?: Contributions to 401(k) Retirement Plans," In *Pension Design and Structure: New Lessons from Behavioral Finance*, edited by Olivia Mitchell and Stephen Utkus, pp. 83-96. Oxford, UK: Oxford University Press.

Madrian, Brigitte and Dennis Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics* 116(4): 1149-1187.

Mitchell, Olivia S., Stephen P. Utkus and Tongxuan Yang. 2005. "Turning Workers into Savers? Incentives, Liquidity, and Choice in 401(k) Plan Design." Working Paper No. W11726. Cambridge, MA: National Bureau of Economic Research.

Munnell, Alicia H., Annika Sunden and Catherine Taylor. (2001/2) "What Determines 401(k) Participation and Contributions?" *Social Security Bulletin* 64(3): 64-75.

Papke, Leslie. 2003. "Individual financial decisions in retirement saving plans: the role of participant-direction." *Journal of Public Economics* 88(1-2): 39-61.

Papke, Leslie and Jeffrey M. Wooldridge. 1996. "Econometric Methods for Fractional Response Variables With an Application to 401(k) Plan Participation Rates." *Journal of Applied Econometrics* 11(6): 619-632.

______________________________. 2007. "Panel Data Methods for Fractional Response Variables with an Application to Test Pass Rates." *Journal of Econometrics* 145(1-2): 121-133.

Shea, John. 1997. "Instrument Relevance in Multivariate Linear Models: A Simple Measure." *Review of Economics and Statistics* 79(2): 348-352.

Wooldridge, Jeffrey M. 2005. "Unobserved Heterogeneity and Estimation of Average Partial Effects," in *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, edited by D.W.K. Andrews and J.H. Stock, pp. 27-55. Cambridge: Cambridge University Press.

**Table 1: Summary Statistics for NCS Data Sample**
**(2,708 jobs in 587 establishments)**

| Variable | Mean | Standard Deviation |
|---|---|---|
| **Participation Rate** | 0.72 | 0.30 |
| **Match Provisions** | | |
| First Dollar Match Rate | 76.68 | 39.19 |
| Log First Dollar Match Rate | 4.21 | 0.49 |
| Percent of Salary Matched | 5.08 | 1.95 |
| Total potential match | 3.55 | 1.94 |
| Log Total potential match | 1.12 | 0.58 |
| **Other Plan Provisions** | | |
| Control over Investment of of Employee Contributions | 0.84 | 0.37 |
| Control over Investment of of Employee Contributions | 0.74 | 0.44 |
| Availability of Loans | 0.69 | 0.46 |
| Automatic Enrollment | 0.06 | 0.24 |
| **Compensation** | | |
| Total Compensation | 33.10 | 23.68 |
| Wage | 22.56 | 16.45 |
| Health Cost | 2.21 | 1.19 |
| DB Cost | 0.54 | 1.40 |
| DB Coverage | 0.41 | 0.49 |
| Other DC | 0.21 | 0.41 |
| **Data Details** | | |
| Year=2003 | 0.54 | 0.50 |

**Table 2: Direct Cross-Sectional (BQMLE) Estimates of Average Partial Effects on 401(k) Participation**
**(standard errors in parentheses)**

| | Base Model | | Total Potential Match and Compensation Components Included | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| **Log of First-Dollar Match** | 0.0567 | 0.0580 | 0.0396 | 0.0456 |
| | (0.0191) | (0.0191) | (0.0237) | (0.0235) |
| **Log of Potential Match** | | | 0.0225 | 0.0162 |
| | | | (0.0208) | (0.0205) |
| **<u>Other Plan Provisions</u>** | | | | |
| **Automatic Enrollment** | 0.0775 | 0.0712 | 0.0801 | 0.0728 |
| | (0.0439) | (0.0440) | (0.0439) | (0.0443) |
| **Investment Choice (Own Contribs)** | -0.0790 | -0.0698 | -0.0829 | -0.0721 |
| | (0.0318) | (0.0317) | (0.0314) | (0.0317) |
| **Investment Choice (Employer Contribs)** | 0.0318 | 0.0250 | 0.0376 | 0.0263 |
| | (0.0303) | (0.0300) | (0.0302) | (0.0302) |
| **Loan Availability** | -0.0103 | -0.0212 | -0.0137 | -0.0211 |
| | (0.0212) | (0.0213) | (0.0214) | (0.0213) |
| **<u>Job-Level Attributes</u>** | | | | |
| **Compensation** | 0.0078 | 0.0066 | 0.0072 | 0.0064 |
| | (0.0018) | (0.0017) | (0.0018) | (0.0017) |
| **Compensation Squared** | -0.000024 | -0.000020 | -0.000022 | -0.000021 |
| | (0.000005) | (0.000005) | (0.000005) | (0.000005) |
| **<u>Imputed Demographics</u>** | | | | |
| **Average Age** | 0.0053 | 0.0055 | 0.0053 | 0.0056 |
| | (0.0025) | (0.0024) | (0.0025) | (0.0024) |
| **Percent College Grad** | -0.0578 | 0.0795 | -0.0581 | 0.0785 |
| | (0.0666) | (0.0730) | (0.0669) | (0.0732) |
| **Percent White** | 0.3429 | 0.1950 | 0.3526 | 0.1949 |
| | (0.1406) | (0.1443) | (0.1406) | (0.1443) |
| **Percent Male** | 0.0196 | -0.0141 | 0.0237 | -0.0148 |
| | (0.0330) | (0.0335) | (0.0325) | (0.0333) |
| **<u>Compensation Components</u>** | | | | |
| **Wage** | -0.0026 | -0.0020 | -0.0024 | -0.0027 |
| | (0.0019) | (0.0017) | (0.0019) | (0.0018) |
| **Defined Benefit Cost** | -0.0028 | -0.0027 | -0.0025 | -0.0024 |
| | (0.0017) | (0.0016) | (0.0016) | (0.0015) |
| **Health Cost** | 0.0343 | 0.0329 | 0.0331 | 0.0330 |
| | (0.0091) | (0.0090) | (0.0090) | (0.0092) |
| **Other DC Plan Present** | 0.0850 | 0.0824 | 0.0862 | 0.0828 |
| | (0.0245) | (0.0231) | (0.0241) | (0.0233) |
| **<u>Other Controls</u>** | | | | |
| **Union, Occupation Controls** | Yes | Yes | Yes | Yes |
| **Estab Size Controls** | Yes | Yes | Yes | Yes |
| **Broad Region and Industry Controls** | No | Yes | No | Yes |

**Table 3: Instrumental Variables (BQMLE) Estimates of Average Partial Effects on 401(k) Participation**
**Instruments Include All Co-Workers**
**(standard errors in parentheses)**

| | Coworker Instruments | | Labor Market Instruments | | Coworker and Labor Market Instruments | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **Average Partial Effects of Match** | | | | | | |
| **Log of First-Dollar Match** | 0.0914 | 0.1994 | 0.1337 | 0.4264 | 0.0679 | 0.2226 |
| | (0.0463) | (0.0621) | (0.0569) | (0.1399) | (0.0366) | (0.0537) |
| **First-Stage Residual** | -0.0495 | -0.1583 | -0.0918 | -0.3858 | -0.0256 | -0.1835 |
| | (0.0461) | (0.0628) | (0.0576) | (0.1411) | (0.0367) | (0.0551) |
| **First-Stage Coefficients** | | | | | | |
| **Co-Workers' Compensation** | 0.0085 | 0.0074 | | | 0.0078 | 0.0072 |
| | (0.0022) | (0.0023) | | | (0.0022) | (0.0023) |
| **Co-Workers' Compensation Squared** | -0.000048 | -0.000046 | | | -0.000044 | -0.000044 |
| | (0.000013) | (0.000014) | | | (0.000013) | (0.000014) |
| **Co-Workers' Age** | -0.0110 | -0.0222 | | | -0.0125 | -0.0104 |
| | (0.0051) | (0.0050) | | | (0.0051) | (0.0052) |
| **Co-Workers' Percent Male** | -0.3181 | -0.2803 | | | -0.2726 | -0.3024 |
| | (0.0611) | (0.0686) | | | (0.0616) | (0.0685) |
| **Co-Workers' Percent White** | -0.4718 | -0.5542 | | | -0.6168 | -0.7267 |
| | (0.2784) | (0.2857) | | | (0.2790) | (0.2869) |
| **Co-Workers' Percent College Graduate** | 0.0863 | 0.1079 | | | 0.0562 | 0.1018 |
| | (0.0828) | (0.0895) | | | (0.0828) | (0.0893) |
| **DC Proportion of Compensation among Other Employers in Local Area** | | | 0.6889 | -0.0189 | 0.1930 | -0.6459 |
| | | | (1.9909) | (2.0373) | (1.9862) | (2.0299) |
| **DC Proportion of Compensation among Other Employers in Industry** | | | 8.1949 | 7.2916 | 7.0890 | 8.1102 |
| | | | (1.4581) | (1.6888) | (1.5313) | (1.7474) |
| **First-Stage Diagnostics** | | | | | | |
| **Partial R-Squared** | 0.0201 | 0.0146 | 0.0134 | 0.0079 | 0.0290 | 0.0236 |
| **Adjusted F-Test (p-value)** | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| **Other Controls** | | | | | | |
| **Union, Occupation Controls** | Yes | Yes | Yes | Yes | Yes | Yes |
| **Estab Size Controls** | Yes | Yes | Yes | Yes | Yes | Yes |
| **Broad Region and Industry Controls** | No | Yes | No | Yes | No | Yes |

**Table 4: Modified Instrumental Variables Analysis (Coworker Instruments Exclude Workers in Same Broad Occupation)**
**(standard errors in parentheses)**

| | Coworker Instruments | | Coworker and Labor Market Instruments | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| **<u>Average Partial Effects of Match</u>** | | | | |
| **Log of First-Dollar Match** | 0.0013 | 0.1306 | 0.0307 | 0.2161 |
| | (0.0570) | (0.0703) | (0.0406) | (0.0600) |
| **First-Stage Residual** | 0.0424 | -0.0880 | 0.0126 | -0.1757 |
| | (0.0568) | (0.0709) | (0.0409) | (0.0613) |
| **<u>First-Stage Coefficients</u>** | | | | |
| **Co-Workers' Compensation** | 0.0063 | 0.0051 | 0.0056 | 0.0047 |
| | (0.0021) | (0.0022) | (0.0022) | (0.0022) |
| **Co-Workers' Compensation Squared** | -0.000041 | -0.000037 | -0.000037 | -0.000035 |
| | (0.000014) | (0.000014) | (0.000014) | (0.000014) |
| **Co-Workers' Age** | -0.0081 | -0.0050 | -0.0091 | -0.0067 |
| | (0.0044) | (0.0045) | (0.0044) | (0.0045) |
| **Co-Workers' Percent Male** | -0.2264 | -0.1849 | -0.1891 | -0.2003 |
| | (0.0533) | (0.0586) | (0.0537) | (0.0585) |
| **Co-Workers' Percent White** | -0.3919 | -0.4555 | -0.4850 | -0.5872 |
| | (0.2475) | (0.2532) | (0.2472) | (0.2540) |
| **Co-Workers' Percent College Graduate** | 0.1084 | 0.1266 | 0.0746 | 0.1202 |
| | (0.0734) | (0.0784) | (0.0736) | (0.0782) |
| **DC Proportion of Compensation among Other Employers in Local Area** | | | 0.4962 | -0.3351 |
| | | | (1.9909) | (2.0332) |
| **DC Proportion of Compensation among Other Employers in Industry** | | | 7.2170 | 7.7851 |
| | | | (1.5297) | (1.7404) |
| **<u>First-Stage Diagnostics</u>** | | | | |
| **Partial R-Squared** | 0.0146 | 0.0101 | 0.0239 | 0.0185 |
| **Adjusted F-Test (p-value)** | 0.000 | 0.000 | 0.000 | 0.000 |
| **<u>Other Controls</u>** | | | | |
| **Union, Occupation Controls** | Yes | Yes | Yes | Yes |
| **Estab Size Controls** | Yes | Yes | Yes | Yes |
| **Broad Region and Industry Controls** | No | Yes | No | Yes |

**Table 5: Instrumental Variables (BQMLE) Estimates of Average Partial Effects on 401(k) Participation**
**With Total Potential Match as a Control**
**(standard errors in parentheses)**

| | Coworker Instruments | | | |
|---|---|---|---|---|
| | Instruments Include All Workers | | Instruments Exclude Workers In Same Broad Occupation | |
| | (1) | (2) | (3) | (4) |
| **Average Partial Effects of Match** | | | | |
| **Log of First-Dollar Match** | -0.0700 | 0.1544 | -0.0648 | 0.2606 |
| | (0.0777) | (0.0811) | (0.0979) | (0.1132) |
| **First-Stage Residual** | 0.0901 | -0.1279 | 0.0844 | -0.2345 |
| | (0.0775) | (0.0814) | (0.0980) | (0.1139) |
| **First-Stage Coefficients** | | | | |
| **Co-Workers' Compensation** | 0.0049 | 0.0063 | 0.0040 | 0.0052 |
| | (0.0016) | (0.0017) | (0.0016) | (0.0017) |
| **Co-Workers' Compensation Squared** | -0.000035 | -0.000040 | -0.000031 | -0.000035 |
| | (0.000010) | (0.000010) | (0.000010) | (0.000010) |
| **Co-Workers' Age** | 0.0006 | 0.0015 | 0.0017 | 0.0033 |
| | (0.0038) | (0.0039) | (0.0033) | (0.0034) |
| **Co-Workers' Percent Male** | -0.1378 | -0.1149 | -0.0915 | -0.0634 |
| | (0.0462) | (0.0518) | (0.0403) | (0.0442) |
| **Co-Workers' Percent White** | -0.8642 | -0.7812 | -0.6198 | -0.5422 |
| | (0.2102) | (0.2150) | (0.1867) | -(0.5422) |
| **Co-Workers' Percent College Graduate** | -0.0398 | -0.1081 | -0.0221 | -0.0906 |
| | (0.0625) | (0.0676) | (0.0554) | (0.0592) |
| **DC Proportion of Compensation among Other Employers in Local Area** | | | | |
| **DC Proportion of Compensation among Other Employers in Industry** | | | | |
| **First-Stage Diagnostics** | | | | |
| **Partial R-Squared** | 0.0146 | 0.0131 | 0.0096 | 0.0088 |
| **Adjusted F-Test (p-value)** | 0.000 | 0.000 | 0.000 | 0.000 |
| **Other Controls** | | | | |
| **Union, Occupation Controls** | Yes | Yes | Yes | Yes |
| **Estab Size Controls** | Yes | Yes | Yes | Yes |
| **Broad Region and Industry Controls** | No | Yes | No | Yes |