THE PRODUCTIVITY EFFECT OF EMPLOYEE PROFIT SHARING PROGRAMS: A METHODOLOGY NOTE

Yu Peng Lin, University of Detroit Mercy

ABSTRACT

While the adopting years of employee profit sharing programs are known, multivariate regression models are often adopted in examining the performance impact of the program either in the context of cross-sectional or longitudinal data sets. In this study, we provide details of two methods – Difference-in-Differences (DID) and Endogenous Switching Model (ESM) and argue that the latter method is more general and more appropriate in estimating the impact. The reasons are two-fold. First, the program impact, if any, may be due to some factors other than the program. DID is one way of tackling this issue. Second, while a program dummy variable is utilized to capture the impact, DID nonetheless does not allow full interaction between the program and other exogenous variables. Hence, the estimation may be contaminated by sample selectivity. Generally, ESM can be very helpful in addressing this concern.

JEL Classification: J33

INTRODUCTION

There are a variety of employee profit sharing plans designed in an attempt to motivate employees. In the past decades, profit-sharing, in which an individual's compensation is tied to the overall performance of the firm, has become increasingly popular in modern corporations in the U.S. as well as in the rest of the world. Profit sharing is understood here to encompass any system which has a direct link between the profits of a company and the compensation of employees. Broadly speaking, profit sharing can be hypothesized to improve company performance through: (1) increasing worker effort; (2) improving the skills of the workforce; and /or (3) enhancing the flow of information within the organization (Kruse 1992)[1].

There are several alternatives of employee profit sharing arrangements that a firm can choose from, such as Employee Stock Ownership Plan (hereafter: ESOP) and Stock Options Program. These varieties of profit sharing programs share the same ingredient - granting employees a share of ownership or profit and expecting the grant to have a positive influence on employees' behavior which in turn leads to better firm performance. A majority of the compensation literature examines the association between such plans and firm performance (e.g. Blasi, Kruse, Sesil, Kroumova, and Carberry 2000, Kruse 1992 and 1993, Lazear 2000, Sesil, Kroumova, Blasi, and Kruse 2002) as well as tests the determinants of the programs. However, while showing an impact of the program, most of the literature relies on multivariate regression either in the context of cross-sectional data or longitudinal data sets (fixed-effect mostly) with the program measured via continuous or dichotomous variables. There is presently little research addressing the hidden limitations of such a methodology. While Difference-in-Differences (hereafter: DID) estimation is well known to and widely adopted in the analysis of policy changes (Wooldridge 2002), this methodology has little success finding its way to the studies of profit sharing programs. Further, DID does not allow full interaction between the program and other exogenous variables. Hence, the estimation may be contaminated by sample selectivity. We believe that an Endogenous Switching Model (hereafter: ESM) can be very helpful in addressing this concern and therefore, should be considered by researchers in the fields of employee compensation as well as policy change analysis. The purposes of this note are to provide details of DID and summarize the estimation procedures for ESM. The paper proceeds as follows: The next section covers a brief literature review. The section that follows proposes and summarizes the two estimating methods. The last section concludes.

LITERATURE REVIEW

While the previous literature has taken up various possible explanations for employee profit sharing programs and evaluated both theoretically and against some evidence pertaining to the way in which those plans are granted, the key question is whether these programs do indeed work as intended. Several papers took up the research effort and identified the performance impact of various employee profit sharing programs. In this section, a brief survey of past empirical studies regarding the association between the programs and firm performance is provided and the limitations of their methodologies are discussed.

The relationship between profit-sharing and company performance has been addressed in several studies. Among those, Kruse (1992) explores the relationship of profit sharing to productivity on a longitudinal data of U.S. public firms with known adopting year of profit sharing plans, concluding that the adoption of profit sharing is associated with a 2.8 ~ 3.5% productivity increase for manufacturing companies, and a 2.5 ~ 4.2% increase for non-manufacturing firms. However, the results were somewhat weakened when restricted to companies adopting profit sharing within the sample period. Cable and Wilson (1989) use a variety of specifications on a sample of 52 British firms. They conclude that profit sharing is associated with 3 ~ 8% higher productivity. Yet, they recognize the possibility that a longer-term feedback from performance to profit-sharing. By using an eight year panel data of 109 Japanese firms, Jones and Kato (1995) find that the introduction of an ESOP leads to a 4 ~ 5% increase in productivity. The findings of positive associations were also evidenced by Wadhwani and Wall's (1990) study of 101 British firms and Kruse's (1993) study of 500 U. S. public companies.

Black and Lynch (2000) show that stock options are associated with increased output by using a nationally representative sample of U.S. establishments surveyed in 1993 and 1996. In a similar vein, Kedia and Mozumdar (2002) find that firms grant options to retain key employees and that firms' use of options to retain key employees creates value and is associated with positive abnormal return in a sample of 200 large NASDAQ firms. Sesil et. al. (2002) document that the adoption of stock options program results in higher levels of value added per employee. By examining the non-executive employee stock options holding, grants, and exercises for 756 U.S. firms during 1994 ~ 1997, Core and Guay (2001) find that the employment of stock options program improves firm value. The positive impact of the adoption of stock options programs is also documented in Sesil and Lin (2011).

Most of the above-mentioned studies used longitudinal regression analysis (fixed-effect mostly) to examine the program impact. While the methodology is feasible, there are two major limitations that may largely weaken the plausibility of their estimation results. First, is there a pre-existing firm performance premium of adopters over non-adopters? If so, the positive impact may be over-estimated. Second, perhaps the most serious, is the positive impact emerges because of the adoption or due to some self-selection behavior? That is, such a positive association can mean various things: it may be true that profit sharing leads to better performance, but it may also be sound that better performing firms tend to adopt the plans. Or there may be some third factors that account for both observations (adoption and better performance) at the same time. While a few of previous studies tried to tackle these issues by, for instance, applying a Heckman selection bias correction (Heckman 1979), these attempting efforts are at best incomplete. We argue that in order to fully understand the true impact of profit sharing plans, formal estimating procedures are needed. In what follows, a discussion of DID estimation and an argument that DID can help to address the afore-mentioned first limitation is provided. Then, we propose and summarize an ESM which, we believe, is a more general and appropriate method in analyzing the impact of labor-related compensation programs as well as in examining the effects of policy changes.

DIFFERENCE-in-DIFFERENCES (DID) and ENDOGENOUS SWITCHING MODEL (ESM)

As discussed in the second section, current literature provides some supporting evidence on the impact of profit sharing programs. Since the introduction of profit sharing may be correlated not just with observable characteristics, but also with unobservable firm characteristics, such an analysis is increasingly benefited by using panel date sets where the same cross-sectional units are observed at different points in time and in turn, fixed effect models are often applied in examining the impact (e.g. Jones and Kato 1995, Kruse 1992, Sesil and Lin 2011)[2].

The fixed-effect estimator is the only consistent estimator when the expected value of the firm-specific error component, conditioned on observables, differs across firms. This is true if the adoption decision is correlated with an unobservable firm characteristic that also influences productivity. This correlation leads to heterogeneity bias in cross-section regression or in random-effects error component specification. Firm specific fixed effect controls for any time invariant heterogeneity of the firms. Other factors with the potential to impact performance include the existence of human resource policies and practices and pension plans. Firm-level fixed effects will capture the differences associated with the impact of these other factors assuming they are time invariant[3].

However, there are typically two major concerns stemming from the fixed-effect estimation. First, the simple comparison of the means of the outcome variable in the treatment and control groups is justified on the grounds that the randomization guarantees they should not have any systematic differences in any pre-treatment variables. Here we define the treatment group the firms that adopt profit sharing plans and the control group the ones that do not have the programs in place. The treatment would be the employment of profit sharing programs. Yet, the randomization is often a very difficult claim to make especially in non-experimental data as it is rarely possible to do this perfectly in which case observed differences between treatment and control groups may be the result of some other omitted factors. One can generally address this concern by applying the Difference-in-Differences (DID) estimation. Second, while longitudinal data set suits the purpose better, some concerns with selectivity inevitably remain. The issue of possible selectivity is a key and serious concern for the kind of analysis conducted in the literature. Unfortunately, to the best of our knowledge, there exist no profit sharing studies trying to formally tackle this issue. In what follows, a detailed elaboration of the DID technique and a summary of a formal procedure for estimating an ESM is provided. In the section that follows, the information of the adopting year of employee profit sharing plans is assumed available to the researchers[4].

Difference-in-Differences (DID) Estimation

The center idea of the DID research design is to use data on the treatment and control groups before the treatment to estimate the "pre-treatment" difference in the impact variable (i.e., firm performance) between the two groups and then compare it with the difference after the receipt of treatment. The treatment is the employment of profit sharing programs. Hence, the effect of the adoption of the program on productivity is the treatment effect. Naturally the two groups need to be similar in some key aspects such as size and industry identifications. Further, any changes other than the adoption decision should affect the two groups in a similar way. Hence, the validity of the DID estimator relies on the assumption of similar underlying trends in the outcome variable (i.e., firm performance) experienced by treatment and control groups prior to adoption.

However, in a non-experimental data, it is very probable that the observed differences between treatment and control groups in the outcome variable may be the result of some other omitted factors. Hence, using a multivariate regression may over-estimate the program impact. Figure 1 illustrates this idea.

Figure 1 suggests that if one used a panel data of adopters and non-adopters all together, one would estimate the treatment effect as A – the validity of this estimate is exclusively based on the assumption that the only reason for observing a difference in the outcome variable between the two groups is the receipt of the treatment. However, there may be a time-invariant difference in overall means between the two groups, which is shown as C. Hence, clearly, A is not the exact treatment effect. The true treatment effect should be B.

The DID research design relies on the idea of examining the outcome variable for similar groups that do not receive the treatment. Hence, the DID estimation of the treatment effect is to use data on treatment and control group before the treatment to estimate the "pre-treatment" difference (i.e., C) between the two groups and then compare this with the difference after the receipt of treatment (i.e., A). Their difference would be the true treatment effect (i.e., B = A – C in Figure 1). Let's define $Q_{i,t}$, to be the mean of firm performance in group *i* at time *t*. For the control group, *i = c*; for the treatment group, *i = tm*; *t = 0* for the pre-treatment periods; *t = 1* for the post-treatment periods. Therefore, the pre-treatment difference is $Q_{tm,0}$ - $Q_{c,0}$ and the post-treatment difference is $Q_{tm,1}$- $Q_{c,1}$. Again, if one just uses post-treatment data, the estimated treatment effect is $Q_{tm',1}$ - $Q_{c',1}$ (i.e., A in Figure 1), which is not the true impact. In contrast the DID estimator will take the "pre-treatment" difference between the treatment and control group as C and estimate the treatment effect as B. In other words, the true treatment effect is ($Q_{tm,1}$ - $Q_{c,1}$) - ($Q_{tm,1}$ - $Q_{c',1}$). We can rewrite this formula as($Q_{tm,1}$ - $Q_{tm,0}$) - ($Q_{c,1}$ - $Q_{c,0}$), and this is the DID estimator. The key assumption underlying the DID estimation is that the average change (or the trend of change) in the outcome variable is presumed to be the same for both the non-adopting peers (control firms) and, counterfactually, the adopting firms if they had not adopted. This assumption is never testable but with multiple observations, we can get some idea of its plausibility.

Again assuming the adopting year of profit sharing is known, following Wooldridge (2002), the regression-based DID estimation is to estimate the following model:

$$Q_{i,t} = \beta_0 + \beta_1 * d2_i + \beta_2 * dB_t + \beta_3 * d2_i * dBt + \varepsilon_{i,t} \text{..............................(1)}$$

$i = firm,\ t = time$

where $Q_{i,t}$ is firm performance. For now and for simplicity, we assume no other control variables. $d2i$ is a dummy variable taking the value 1 if the firm is in the treatment group and 0 if it is in the control group. Hence, $d2i$ is actually the adopter-non-adopter identification variable. dBt equals unity in the post-treatment period and 0 in the pre-treatment period. The period dummy dBt captures some aggregate factors that affect *Q* over time in the same way for both groups. It helps to remove any effect of post-adoption factors that are common to both groups. summarizes the way that both groups are influenced by time. There may also be time-invariant differences in overall means between the groups, but this aspect is captured by. Most importantly, the interaction term of $d2_i$ and dB_t is simply a dummy variable equals to unity for those observations in the treatment group "AND" in the post-treatment periods; zero otherwise. This is actually how we code the "treatment" group in the DID framework. Without other factors in the regression, would be the DID estimator in the simplest form (i.e., ($Q_{tm,1}$ - $Q_{tm,0}$) - ($Q_{c,1}$ - $Q_{c,0}$)). As other explanatory variables are added to equation (1), the estimate of no longer has the simple form, but its interpretation is similar. A good example can be found in Wooldridge (2006).

When the empirical model is specified as in regression (1), clearly, this design is most plausible when the untreated comparison group is very similar to the treatment group in all aspects except for the influence under study. A situation favorable to this design is one in which the comparison group both before and after has a distribution of the outcome variable close to that of the treatment group during the pre-treatment periods. This validity condition is reflected in the "common trend" assumption of the DID estimation design.

Endogenous Switching Model (EAM)

While the DID estimation helps to estimate the true treatment effect of the program, the possibility of selectivity inevitably exists. In this section we address this issue by deriving a formal estimation procedure that helps to correct the estimation bias associated with selectivity.

Sample selection bias refers to the problems where the dependent variable is observed only for a restricted, nonrandom sample. In the current framework, one observes a firm's productivity within the program only if the firm has adopted the program. Conversely, one observes a firm's non-adoption productivity only if the firm does not belong to the program[5]. Further, profit sharing plans may in fact have a true effect on productivity, but this effect cannot be generalized and it varies across firms. The reason is that the incentives to adopt such a program may be strongest where it is expected to have the most impact. Consider, for instance, within a particular organizational characteristic. In the current work, we consider the selection bias in which the incentives to adopt profit sharing programs are linked to some firm-specific characteristic that, in particular, are not constant over time, so that fixed effect models cannot fully eliminate it. One can attempt to examine the program impact by using an endogenous switching model which allows full interaction between the program and other firm characteristics.

Although selection bias is commonly confronted in cross-sectional studies (Heckman 1979), it is less frequently considered to be a concern in a panel data estimation. It is partially due to the perception that fixed effects estimation will eliminate most forms of unobserved heterogeneity. This is especially true while the selectivity is only through firm fixed effects (Verbeek and Nijman 1992). However, some other forms of selection bias might not be eliminated. Specifically, the selection bias through the firm/time specific factors (hence, non-time invariant) will contaminate the fixed effect estimations (Wooldridge 2002). Consider the following fixed effect model with a selection mechanism.

$$Q_{1,it} = \beta_{1,i}^{'} X_{it} + \mu_{1,i} + \varepsilon_{1,it} \quad \text{.....(2)}$$

$$Q_{0,it} = \beta_{0,i}^{'} X_{it} + \mu_{0,i} + \varepsilon_{0,it} \quad \text{.....(3)}$$

$$S_{it}^{*} = \gamma^{'} Z_{it} + \theta i + \eta_{it} \quad \text{.....(4)}$$

$$S_{it} = \begin{cases} 1 \ \textit{if} \ S_{it}^{*} > 0 \\ 0 \ \textit{otherwise} \end{cases} \quad \text{.....(5)}$$

Where i = *firm;* t = *time*; $Q_{1,it}$ is the productivity of firm i at time t if it adopts profit sharing programs and $Q_{0,it}$ is the productivity when the firm does not employ the program. Hence, Equations (2) and (3) are production functions of adopters (i.e., treatment group) and non-adopters (i.e., control group), respectively. Note that, in any firm year, only the output with or without the program can be observed but not both. X_{it} is a vector of firm characteristics, including production inputs and year dummies. Equations (4) and (5) constitute the adoption decision. Z_{it} is a vector of exogenous variables related to the adoption decision. The total benefit from the adoption is given by (4), so that a firm adopts if (4) > 0 and does not do so otherwise. This adoption decision is expressed in (5), where S_{it} is a dummy variable denoting the decision. μ_i and

θ_i are firm specific effects. ε_{it} and η_{it} are error terms and represent firm/time-specific effects. We assume these are *i.i.d.* (independently and identically distributed) drawings from a multivariate normal distribution.

Following Wooldridge (2002) and Vella and Verbeek (1998), we allow the co-variances $\sigma_{j,\mu\theta}$ and $\sigma_{j,\varepsilon\eta}$ ($j=0,1$ corresponding to the adoption decision) to be non-zero; zero for all other co-variances. These non-zero co-variances suggest that the random components in the productivity equations (*2*) and (*3*) are potentially correlated with those in the selection equation. This generates the potential endogeneity of adoption decision in the productivity equation, which leads to a selectivity problem. Notice that under common regularities, the consistency of the fixed effect estimator requires $E(\varepsilon_{it} \mid X_{it},S_{it}) = 0$. That is, $\sigma_{\varepsilon\eta} = 0$. Thus fixed effects estimation cannot produce consistent estimates if the selection is operating through the firm/time-specific effects ($_{it}$) (Jakubson 1991).

To evaluate the benefit of the program that has already been employed, "the treatment effect on the treated" (Maddala 1983, Heckman 1997) is considered. Specifically, one can compare the output $Q_{1,it}$ in the program and the expected potential outcome without the program, which is defined as $E(Q_{1,it}|S_{it}=1) - E(Q_{0,it}|S_{it}=1)$ (Maddala 1983). Clearly in this model, the program effect does not show up as a dummy variable, but rather in the fact that the constant term and betas may differ from the adopters to non-adopters. Hence, this model essentially allows a full set of interactions between program status and the independent variables. In what follows, the Heckman's method is applied and one or more correction terms are added to the production equations using an estimated version of the adoption equation[6].

The production equations (2) and (3) cannot in general be consistently estimated by ordinary least squares using the observed firm output. The trouble occurs since

$$E(\mu_{1,i}|S_{it}=1)\neq 0,\quad E(\mu_{1,i}|S_{it}=0)\neq 0,\quad E(\varepsilon_{1,it}|S_{it}=1)\neq 0, \text{ and } E(\varepsilon_{0,it}|S_{it}=0)\neq 0$$

Hence, the conventional Heckman two-stage technique is followed to obtain consistent estimation. The idea of the procedure is to find the expression for the means of $E(\mu_{j,i}|S_{it})$ and $E(\varepsilon_{j,it}|S_{it})$, where $j = 0, 1$ and adjust the error terms so that they will have zero means. The first stage in our estimation is by Probit and the second is by OLS. Therefore, conditional on adoption status, the adopters' and non-adopters production equations, respectively, are

$$Q_{1,it} = \beta_{1,i}^{'}X_{it} + \sigma_{1,\mu\beta}C_i + \sigma_{1,\varepsilon\eta}C_{it} + \varphi_{1,it} \quad\text{.........(2')}$$

$$Q_{0,it} = \beta_{0,i}^{'}X_{it} + \sigma_{0,\mu\theta}C_i + \sigma_{0,\varepsilon\eta}C_{it} + \varphi_{0,it} \quad\text{.........(3')}$$

where $E(\varphi_{j,it}|S_{it}) = 0$

Note that $E[\mu_{j,i}|S_i] = \sigma_{j,\mu\theta}C_i$ and $E[\varepsilon_{j,it}|S_i] = \sigma_{j,\varepsilon\eta}C_{it}$ are the two conditional expectations of the error terms (Nijman and Verbeek 1992, Vella and Verbeek 1998) with

$$C_i = \frac{T_i}{\sigma_\eta^2 + T_i\sigma_\theta^2}E(\theta_i + \eta_{it}|S_i) \quad\text{.........(6)}$$

$$C_{it} = \frac{1}{\sigma_\eta^2}[E(\theta_i + \eta_{it}|S_i) - \frac{T_i\sigma_\theta^2}{\sigma_\eta^2 + T_i\sigma_\theta^2}E(\theta_i + \eta_{it}|S_i)] \quad\text{.........(7)}$$

$$E(\theta_i + \eta_{it} | S_i) = \int_{-\infty}^{\infty} [\theta_i + E(\eta_{it} | S_i, \theta_i)] f(\theta_i | S_i) d\theta_i \qquad (8)$$

And

$$E(\eta_{it} | S_i, \theta_i) = (2S_{it} - 1)\sigma_\eta \frac{\phi(b_{it})}{\Phi(b_{it})} \qquad (9)$$

$$f(\theta_i | S_i) = \frac{\prod_{t=1}^{T_i} \Phi(b_{is})\sigma_\theta^{-1}\phi(\theta_i / \sigma_\theta)}{\int \prod_{s=1}^{T_i} \Phi(b_{is})\sigma_\theta^{-1}\phi(\theta_i / \sigma_\theta) d\theta_i} \qquad (10)$$

where $\phi(.)$ *and* $\Phi(.)$ are the p.d.f and c.d.f of N(0,1);

$b_{it} = (2S_{it} - 1)(\gamma' Z_{it} + \theta_i) / \sigma_\eta$

Notice that (6) and (7) are functions of the parameters in the program adoption decision mechanism (i.e. equations (4) and (5)) only. Consequently, though *Ci* and *Cit* are not observed, they can be consistently estimated by replacing the unknown parameters by their estimates obtained from the random effects Probit model in (5) with the whole sample. The resulting estimated correction terms can be substituted in (2') and (3'). The significance of these terms is the test for selection bias. After obtaining consistent estimates of the β's, these are plugged into equations (2') and (3'). One can then get estimates of $\hat{Q}_1$ and $\hat{Q}_0$ and calculate the difference as the program impact by using $\overline{X}$ as the vector of average values of the explanatory variables.

CONCLUSION

Assuming the adopting years of employee profit sharing programs are known, in this study, two methods – Difference-in-Differences and Endogenous Switching Model are proposed and summarized. The purposes are two-fold. First, while Fixed-effect models are plausible in accessing the program impact on firm performance, the estimation may in inevitably capture the productivity differences due to factors other than the program. Indeed, if there were performance premium enjoyed by the adopters prior to the adoption decision, the Fixed-effect estimation may over-estimate the impact. DID is a more appropriate estimation method and it can deliver a more reliable result.

Second, while the DID estimation helps to estimate the true treatment effect of the program, the issue of selectivity inevitably exists. Profit sharing plans may in fact have a true effect on productivity, but this effect cannot be generalized and it varies across firms. The reason is that the incentives to adopt such a program may be strongest where it is expected to have the most impact. This concern can be addressed by an Endogenous Switching Model in which it allows full interaction between the program and other firm characteristics. The estimation obtained from the model in fact serves the purpose much better than Fixed-effect and DID.

This research suggests at least two avenues for future research. First, ESM can be greatly applied to policy analysis as well. The decision of implementing a policy (e.g., labor-related policies, macroeconomic policies) is most likely endogenous in the sense that the adoption decision is with the expectation that it has the most impact given the firm's, society's, or country's characteristics. ESM incorporates the decision mechanism and produces more precise estimate of the policy effect. Second, a firm's adoption of a particular compensation program is possibly not only endogenously related to its characteristics but also associated with the outcome variable. Under this scenario, the ESM should be considered under a system of simultaneous equations.

However, there are few limitations/complexities in applying ESM. First, to evaluate a firm's adoption decision, one needs to identify appropriate determinants which, by theory, need to be unrelated to the independent variables in the production functions. Finding those determinants can be a challenge. Second, to resolve the above, Instrumental Variable (IV) technique can be applied.

ENDNOTES

1. We do not plan to provide an exhaustive review of employ profit sharing plans. Interested readers should refer to Alan Blinder's "*Paying for Productivity*" (1990) and Lazear and Gibbs (2008) for more details.

2. For the details of fixed effect estimation, interested reader should refer to Wooldridge (2002).

3. According to Cole (1989), once a human resource practice is adopted it is extremely unlikely to be discontinued.

4. While the lack of appropriate data poses the main limitation of empirical studies in this subject, there are indeed some public or proprietary data available for the proposed methodologies in this note. For instance, the ESOP adopting year data used in Kruse (1993), the stock option programs adoption information employed in Sesil, Kroumova, Blasi, and Kruse (2002)

5. While self-selection may create a bias, it does not necessarily do so. Heckman and Robb (1985) review several models using panel data in which a bias does not exist under certain decision rules and error processes.

6. Instead of using a two-step OLS estimation, the model can also be estimated by the Full Information Maximum Likelihood (FIML) (Wooldridge, 2002).

REFERENCES

Addison, John (2009). *The Economics of Codetermination: Lessons from the German Experience*, Palgrave Macmillan.

Blasi, Joseph, Kruse, Douglas, Sesil, James, Kroumova, Maya, & Carberry, Ed (2000). *Stock Options, Corporate Performance, and Organizational Change*, National Center for Employee Ownership press.

Black, Sandra & Lynch, Lisa (2000). What's Driving the New Economy: The Benefits of Workplace Innovation, *NBER working paper #7479*.

Cable, John & Wilson, Nicholas (1989). Profit-Sharing and Productivity: An Analysis of UK Engineering Firms. *The Economic Journal* 99(396), 366.

Card, David & Krueger, Alan (1994). Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania. *American Economic Review* 84(4), 772.

Cole, Robert (1989). *In Strategies for Learning*, Berkeley, Calf.: University of California Press.

Conte, Michael & Svejnar, Jan (1990). "The Performance Effects of Employee Ownership Plans", in *Paying for Productivity* edited by A. S. Blinder Washington, D.C., Brookings Institution 143 ~ 182.

Core, John & Guay, Wayne (2001). Stock Option Plans for Non-Executive Employees. *Journal of Financial Economics* 61, 253.

Fitzroy, Felix & Kraft, Kornelius (1987). Profitability and Profit Sharing. *Journal of Industrial Economics* 35(2), 113.

Heckman, James (1979). Sample Selection Bias as a Specification Error, *Econometrica* 47(1), 153.

Heckman, James & Richard Robb (1985). "Alternative Methods for Evaluating the Impact of Treatment on Outcomes" in *Longitudinal Analysis of Labor Market Data* edited by J. Heckman and B. Singer, Cambridge University Press.

Ittner, Christopher, Lambert, Richard, & Larcker, David (2003). The Structure and Performance Consequences of Equity Grants to Employees of New Economy Firms. *Journal of Accounting and Economics* 34 (1-3), 89.

Jakubson, George (1991). Estimation and Testing of the Union Wage. Effect Using Panel Data, *The Review of Economics Studies* 58(5), 971.

Jensen, Michael & Meckling William (1976). Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure, *Journal of Financial Economics* 3(4), 305.

Jensen, Michael & Murphy, Kevin (1990). Performance Pay and Top-Management Incentives, *Journal of Political Economy* 98(2), 225.

Jones, Derek & Kato, Takao (1995). The Productivity Effects of Employee Stock-Ownership Plans and Bonuses: Evidence from Japanese Panel Data. *The American Economic Review* 85(3), 391.

Kedia, Simi & Mozumdar, Abon (2002). Performance Impact of Employee Stock Options, working paper.

Kruse, Douglas (1992): Profit sharing and Productivity: Microeconomic Evidence from the United States. *The Economic Journal* 102(410), 24.

Kruse, Douglas (1993). Does Profit Sharing Affect Productivity?. *NBER working paper #4542*.

Lazear, Edward (1986). Salaries and Piece Rates, *Journal of Business* 59(3), 405.

Lazear, Edward (2000). Performance Pay and Productivity. *American Economic Review* 90(5), 1346.

Lazear, Edward & Gibbs, Michael (2008). *Personnel Economics in Practice*, Wiley.

Maddala, G. S. (1983). *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge University press.

Nijman, Theo & Verbeek, Marno (1992). Nonresponse in Panel Data: The Impact on Estimates of a Life Cycle Consumption Function. *Journal of Applied Econometrics* 7(3), 243.

Sesil, James, Kroumova, Maya, Blasi, Joseph, & Kruse, Douglas (2002). Broad-Based Employee Stock Options in US 'New Economy' Firms. *British Journal of Industrial Relations* 40(2), 273.

Weitzman, Martin & Kruse, Douglas (1990). Profit sharing and Productivity, in *Paying for Productivity* edited by A. S. Blinder, Washington, D.C., Brookings Institution, 95 ~ 140.

Wadhwani, Sushil & Wall, Martin (1990). The Effects of Profit-Sharing on Employment, Wages, Stock Returns and Productivity: Evidence from UK Micro-Data. *The Economic Journal* 100(399).

Vella, Francis & Verbeek, Marno (1998). Whose Wages do Unions Raise? A Dynamic Model of Unionism and Wage Rate Determination for Young Men. *Journal of Applied Econometrics* 13(2), 163.

Vella, Francis & Verbeek, Marno (1999). Two-step Estimation of Panel Data Models with Censored Endogenous Variables and Selection Bias. *Journal of Econometrics* 90(2), 239.

Vella, Francis & Verbeek, Marno (1999). Estimating and Interpreting Models with Endogenous Treatment Effects, *Journal of Business & Economics Statistics* 17(4), 473.

Verbeek, Marno & Nijman, Theo (1992). Testing for Selectivity Bias in Panel Data Models. *International Economic Review* 33(3), 681.

Wooldridge, J. M. (2002). Econometric Analysis of Cross Section and Panel Data, the MIT press.

Wooldridge, J. M. (2006). Introductory Econometrics: A Modern Approach, 3rd ED., Thomson South-Western.

FIGURE 1: AN ILLUSTRATION OF THE DID ESTIMAITON