NOTES

RANDOMIZATION AS AN INSTRUMENTAL VARIABLE

James J. Heckman*

Abstract—This paper discusses how randomized social experiments operate as an instrumental variable. For two types of randomization schemes, the fundamental experimental estimation equations are derived from the principle that experiments equate bias in control and experimental samples. Using conventional econometric representations, I derive the orthogonality conditions for the fundamental estimation equations. Randomization is a multiple instrumental variable in the sense that one randomization defines the parameter of interest expressed as a function of multiple endogenous variables in the conventional usage of that term. It orthogonalizes the treatment variable simultaneously with respect to the other regressors in the model and the disturbance term for the conditional population. However, conventional "structural" parameters are not in general identified by the two types of randomization schemes widely used in practice.

Randomized social experiments are now coming into widespread use. Their limitations and benefits are also beginning to be understood. Papers by Burtless (1995), Heckman (1992), Heckman and Smith (1995a) and Moffitt (1992) among others clarify the behavioral and statistical assumptions underlying the experimental method.

This paper contributes to this literature and considers the social experiment as an instrumental variable. It develops the point that under the assumptions that justify its application, widely-used randomization schemes do not achieve their results by producing exogeneity of the treatment with respect to the population error term, as that term is ordinarily used in econometrics. Rather, these randomizations operate by *balancing* or equating the bias in the sample of persons randomized into a program with the bias in the sample of persons randomized out of the program. Randomization creates independence of the treatment effect with respect to other regressors and with respect to the error term in conditional populations. One randomization generates a multiple instrumental variable. Treatment effects as functions of an arbitrarily large number of endogenous variables can be identified from one randomization. However, treatment effects are usually not "structural parameters" in the conventional use of that term.

I develop these points for two distinct economic models: (a) a common effect model (treatment has the same effect on everyone with the same observed X characteristics) and (b) a variable effect model (treatment has different effects on everyone with the same observed X characteristics). The latter model is also known as a random effects model. The former is the one most widely used in applied work. Heckman and Robb (1985) and Heckman (1992) demonstrate the value in distinguishing between these two models in devising strategies for evaluating social programs.

I first consider randomization administered at the stage where persons apply to and are accepted into a social program and are then randomized out of the program. Randomization administered at that stage is widely used. Under the conditions specified in Heckman (1992) and Heckman and Smith (1995), this randomization identifies the mean gain to participating in the program for those who would usually participate in it. This mean gain is sometimes called the effect of treatment on the treated. I also consider samples produced by randomizing eligibility for the program. Before doing this, I briefly state the evaluation problem.

Received for publication August 23, 1993. Revision accepted for publication September 1, 1995.

* University of Chicago and the American Bar Foundation.

This research was supported by NSF SES-93-34568 and a grant from the Russell Sage Foundation. Heckman is Henry Schultz Distinguished Service Professor in the Department of Economics, Director, Center for Social Program Evaluation, Harris School, University of Chicago and Fellow, American Bar Foundation. I thank an anonymous referee for comments.

I. The Evaluation Problem

The evaluation problem is a missing data problem. Persons may be in either one of two states but not both at the same time. The states are denoted "0" and "1" respectively. Outcomes are (Y_0, Y_1). Let $d = 1$ if a person is in state "1"; $d = 0$ otherwise. The outcome observed for an individual, Y, is

$$Y = dY_1 + (1 - d)Y_0.$$

This is an instance of the Roy model (1951) or a switching model (see Goldfeld and Quandt (1972)). Statisticians call this the "Rubin model" after one clear exposition of Fisher's model of experiments set forth by Rubin (1978). If "0" is the no program state, and "1" is the program state, the gain to participating in the program is

$$\Delta = Y_1 - Y_0.$$

If, contrary to hypothesis, we could simultaneously observe Y_1 and Y_0 for the same person, there would be no evaluation problem. One could construct Δ for everyone.

To cast the model into familiar econometric notation, write the two population model of Fisher (1951), Cox (1958), Roy (1951) or Rubin (1978) as a function of observables (X) and unobservables (U_1, U_0):

$$Y_1 = g_1(X) + U_1 \tag{1a}$$

$$Y_0 = g_0(X) + U_0 \tag{1b}$$

where

$$E(U_1) = 0 = E(U_0).$$

It is assumed that g_1 and g_0 are nonstochastic functions. For the familiar case of linear regression, the g functions specialize to

$$g_1(X) = X\beta_1$$

$$g_0(X) = X\beta_0.$$

There are many forms of the evaluation problem depending on what feature of the missing data one seeks to construct. The most common form of the problem is cast in terms of means. One mean receives the most attention:

The Mean Effect of Treatment on the Treated

$$\begin{aligned} E(Y_1 - Y_0 | X, d = 1) &= E(\Delta | X, d = 1) \\ &= g_1(X) - g_0(X) + E(U_1 - U_0 | X, d = 1). \end{aligned} \tag{2}$$

This mean answers the question "how much did persons participating in the program benefit compared to what they would have experienced without participating in the program?'' This parameter is the gross gain to participants from the program. When compared with costs, this parameter is informative on the question of whether or not an existing program's benefits exceed its costs. It is a non-standard parameter from the vantage point of conventional econometrics because it combines ''structure'' (the g_0 and g_1 functions) with the means of error terms (U_0 and U_1).[1]

A second mean also receives some attention in the literature—the effect of randomly selecting persons from the general population into the program:

Mean Effect of Treatment Randomly Applied to the Population

$$E(Y_1 - Y_0 | X) = E(\Delta | X) = g_1(X) - g_0(X) + E(U_1 - U_0 | X). \tag{3}$$

This mean answers the question of how much the average outcome would be affected if participation in a program were universal, assuming that there are no general equilibrium effects. Alternatively, this parameter is the effect of taking a person from the general population at random and moving him or her from ''0'' to ''1.'' Further discussion of these parameters and their relationship to the traditional parameters of cost–benefit analysis is presented in Heckman and Smith (1995b). Although the assumption of separability between X and U is conventional in econometrics, it is not required to define $E(Y_1 - Y_0 | d = 1, X)$ or $E(Y_1 - Y_0 | X)$ nor is it necessary to assume such separability in deriving estimates from experiments.

For certain purposes, it is also of interest to inquire about distributions of gains:

$$F(\Delta | d = 1, X) \tag{4a}$$

or

$$F(\Delta | X) \tag{4b}$$

but it has been shown that social experiments, unaccompanied by further assumptions, cannot recover these distributions (see Heckman (1992) or Clements, Heckman and Smith (1993, revised, 1995)). Under ideal conditions, social experiments recover

$$F(Y_0 | d = 1, X)$$

and

$$F(Y_1 | d = 1, X)$$

if randomization is administered at a stage of the application and acceptance decision at which persons would ordinarily be accepted into programs, and there is no attrition from the program.

The evaluation problem arises from the fact that ordinary observational data do not provide sample counterparts for the missing counterfactuals. For means, experiments supplement observational data by providing the information needed to form the sample counterpart of

$$E(Y_0 | d = 1, X).$$

More generally, social experiments supplement observational data and produce the information needed to form the empirical distribution counterpart of

$$F(Y_0 | d = 1, X).$$

Randomization provides the sample counterparts to these population objects if randomization bias induced by the process of experimentation is assumed to be unimportant (Heckman (1992)).

II. Randomization Balances Bias

First consider randomization at the stage where persons apply to and are accepted into a social program. Nowhere is it assumed that $E(U_1|X) = 0$ or $E(U_0|X) = 0$. Thus X can fail to be exogenous in the conventional sense of that term. Yet randomized trials that do not disrupt the program, and are not subject to attrition or non-compliance, produce the data that can be used to consistently estimate parameter (2). This highlights both the unusual nature of that parameter and the benefits of randomization.

To establish how randomization identifies (2), it is instructive to introduce new variables denoted by $*$. $d^* = 1$ denotes the event: ''in the presence of randomization a person would have participated in the program except possibly for being randomized out.'' We may also define Y_1^* and Y_0^* to be the outcomes observed under a regime of randomization. Absence of randomization bias for the mean gain is defined as

$$E(Y_1 | d = 1, X) = E(Y_1^* | d^* = 1, X)$$

and (A-1)

$$E(Y_0 | d = 1, X) = E(Y_0^* | d^* = 1, X)$$

i.e., randomization does not alter the outcome mean gain for the program being evaluated for all values of X.

Randomization operates *conditionally* on $d^* = 1$. This is appropriate because parameter (2) is defined conditionally. In the subpopulation for which $d^* = 1$, randomization operates by selecting persons into the program by a random device. $R = 1$ if a person is randomized into the program; $R = 0$ if a person is randomized out of the program. I assume that if $R = 1$, persons accept admission into the program and if $R = 0$, they do not obtain program services.

As a consequence of assumption (A-1), and the additional assumption that equates $R = 1$ or $R = 0$ with receipt of program services, using the definition $Y = Y_1 d + Y_0(1 - d)$,

$$E(Y | d^* = 1, R = 1, X) = E(Y_1 | d = 1, X) = g_1(X) + E(U_1 | d = 1, X) \tag{5a}$$

$$E(Y | d^* = 1, R = 0, X) = E(Y_0 | d = 1, X) = g_0(X) + E(U_0 | d = 1, X) \tag{5b}$$

[1] Heckman and Robb (1985, 1986) present conditions for identifying this parameter using instrumental variables in nonexperimental settings. Their conditions apply to the general ''variable treatment effect case'' of equations 1(a) and 1(b). See also Heckman (1995) for the implicit behavioral assumptions invoked in using instrumental variables to estimate parameter (2) when responses to treatment are heterogeneous.

Randomization creates data that can be used to estimate counterfactual (5b). Conditional mean (5a) can be consistently estimated using ordinary observational data. If there is no randomization bias, and R is synonymous with receipt of services, both experiments and observational data are equally informative about (5a). I assume no randomization bias and for notational simplicity henceforth equate d and d^*, Y_1 and Y_1^*, and Y_0 and Y_0^*.

Subtract (5b) from (5a) to obtain

$$\begin{aligned} &E(Y|d = 1, R = 1, X) - E(Y|d = 1, R = 0, X) \\ &= g_1(X) - g_0(X) + E(U_1 - U_0|d = 1, X) \\ &= E(\Delta|d = 1, X). \end{aligned} \tag{6}$$

This can be consistently estimated using sample counterparts to population means. If some of the X variables are continuous, a nonparametric kernel estimator for pointwise means can be constructed using conventional methods. (See, e.g., Härdle (1990)). Nowhere is it necessary to assume that

$$E(U_1|X) = 0 \quad \text{or} \quad E(U_0|X) = 0.$$

In fact, it is clear from the definition of Δ that in general

$$E(U_1 - U_0|\, d = 1, X) \neq 0.$$

To place randomization into a more familiar-looking instrumental variables framework, define $\tilde{U}_1$ as U_1 conditional on $d = 1$ and X and define $\tilde{U}_0$ as U_0 conditional on $d = 1$ and X. Then Y conditional on $d = 1$ and X may be written as $\tilde{Y}$, and

$$\tilde{Y} = g_0(X) + [g_1(X) - g_0(X)]R + \{\tilde{U}_1 - \tilde{U}_0\}R + \tilde{U}_0. \tag{7}$$

In this notation, it is not assumed that $E(\tilde{U}_0|d = 1, X) = 0$ nor is it assumed that $E(\tilde{U}_1|d = 1, X) = 0$.

Using definition (2), and defining the mean-adjusted errors $\tilde{U}_0^* = \tilde{U}_0 - E(U_0|d = 1, X)$ and defining $\tilde{U}_1^* = \tilde{U}_1 - E(U_1|d = 1, X)$, and defining $\varphi(X) = g_0(X) + E(U_0|d = 1, X)$ we may rewrite equation (7) as

$$\tilde{Y} = \varphi(X) + E(\Delta|X, d = 1)R + \tilde{U}_0^* + (\tilde{U}_1^* - \tilde{U}_0^*)R. \tag{8}$$

(Observe that * as used here is completely distinct from its use above.) Conditioning on X and $d = 1$, $\varphi(X)$ is an intercept and (8) is a simple univariate regression defined for each value of X, given $d = 1$. Randomization makes R independent of $(\varphi(X), \tilde{U}_0^*, \tilde{U}_1^*)$ conditional on $d = 1$ and X. The orthogonality conditions produced by randomization are

$$E[\tilde{U}_0^* + (\tilde{U}_1^* - \tilde{U}_0^*)R|R] = 0 \tag{9a}$$

and

$$E[U_0^* + (\tilde{U}_1^* - \tilde{U}_0^*)R] = 0. \tag{9b}$$

For all X given $d = 1$, (9a) identifies $E(\Delta|X, d = 1)$ and (9b) identifies $g_0(X) + E(U_0|d = 1, X)$ but not its individual components. Randomization makes R orthogonal to $\tilde{U}_0$, $\tilde{U}_1 - \tilde{U}_0$, $g_1(X)$ and $g_0(X)$. It does not make $g_0(X)$ or $g_1(X)$ orthogonal to U_0 or $U_1 - U_0$.

Thus experiments do not in general identify $g_0(X)$, since in general $E(\tilde{U}_0|X, \mathrm{d} = 1) \neq 0$. From this, it follows that experiments of the type discussed in this section do not in general identify the structural parameters of the original equation but they identify parameter (2) provided that (*A*-1) is valid and persons assigned to treatment receive it and persons denied treatment do not. This point is obvious once it is recognized that randomization at the stage where persons have applied to and been accepted into a program generates samples conditional on variables that are, in general, endogenous in the conventional usage of that term.

These expressions simplify when there is a common effect model. In that case $U_1 = U_0$, and $\tilde{U}_0 = \tilde{U}_1 \stackrel{\text{def}}{=} \tilde{U}$. Then

$$E(\Delta|X, d = 1) = g_1(X) - g_0(X)$$

and equation (7) may be written as

$$\begin{aligned} \tilde{Y} = {} & [g_0(X) + E(\tilde{U}|d = 1, X)] + E(\Delta|X, d = 1)R \\ & + [\tilde{U} - E(\tilde{U}|d = 1, X)]. \end{aligned}$$

Letting $\tilde{U}^* = \tilde{U} - E(\tilde{U}|d = 1, X)$, orthogonality condition (9a) becomes

$$E(\tilde{U}^*|R) = 0.$$

Again notice that in general $g_0(X)$ cannot be separated from $E(U|d = 1, X) = E(\tilde{U}|d = 1, X)$.

A familiar form of the common effect model writes

$$g_0(X) = X\beta_0$$

$$g_1(X) = X\beta_1$$

so in the common effect model

$$E(\Delta|X, d = 1) = X(\beta_1 - \beta_0)$$

where the conditioning on $d = 1$ is often left implicit. An even more familiar form of the common effect model writes

$$g_0(X) = X\beta_0$$

$$g_1(X) = X\beta_0 + \alpha.$$

This is the dummy endogenous variable model (Heckman (1978)). In this case

$$\tilde{Y} = X\beta_0 + R\alpha + \tilde{U}. \tag{10}$$

Randomization ensures that R is independent of both $\tilde{U}$ and X. It does not ensure that X is independent of $\tilde{U}$. The orthogonality between R and X induced by an experiment implies that any dependence between X and $\tilde{U}$ does not affect the identifiability of α. Randomization creates an orthogonal regressor model for the subpopulation defined conditional on $d = 1$. Since R is independent of X and $\tilde{U}$, α is identified even if X is not orthogonal to $\tilde{U}$, and β_0 is not identified.

III. Discussion

Observe that one randomization identifies an entire function $E(\Delta|X, d = 1)$ over the support of X (i.e., the values of X where this parameter is defined). In principle, $E(\Delta|X, d = 1)$ can be an infinite-dimensional function of X. Hence in this sense randomization is a multiple instrumental variable.

Observe further that randomization enriches the support of X in the following way. Suppose in the population that

$$\text{Support } (X|d = 1) \neq \text{Support } (X|d = 0).$$

Then in the subset of X values for which there is no overlap, observational methods cannot obtain comparisons for all X values and $E(\Delta|X,d = 1)$ cannot be identified for all X.[2] Randomization creates a balanced support set because

$$\text{Support } (X|d = 1, R = 1) = \text{Support } (X|d = 1, R = 0).$$

Unless

$$\text{Support } (X|d = 1) \subset \text{Support } (X|d = 0),$$

randomization enlarges the support set over which $E(\Delta|X, d = 1)$ can be defined and estimated. However, unless $0 < \Pr(d = 1|X) < 1$, for all X, randomization does not identify $E(\Delta|X, d = 1)$ for all possible values of X. (See, e.g. Rosenbaum and Rubin (1983).) An extreme example of the benefit of randomization in enlarging the support set occurs when for certain values of X, the event $d = 0$ does not occur, i.e.,

$$\Pr(d = 0|X) = 0$$

but $d = 1$ occurs with positive probability for all values of X:

$$0 < \Pr(d = 1|X) < 1.$$

In this case, randomization expands the support of X given $d = 1$ to the entire support of X. It permits identification of $E(\Delta|X, d = 1)$ for all possible values of X.

Observe that, in general, experiments defined conditional on $d = 1$, do not identify $E(\Delta|X)$, the effect of selecting a person at random from state "0" and moving the person to "1". However, if the common effect model is assumed so $U_1 - U_0 = 0$, experiments conducted on populations defined conditional on $d = 1$ recover

$$E(\Delta|X),$$

the effect of selecting a person from the general population and placing them in the program. For in that case,

$$E(\Delta|X) = E(\Delta|d = 1, X).$$

(See Heckman (1992) or Heckman and Smith (1995).)

In one case where responses to treatment are heterogeneous, and randomization is administered to those for whom d would have been "1" in the absence of randomization, the parameter (3) is nonetheless identified. Suppose at the time agents enroll in the program they forecast *their* gain to be the total population mean gain. Then clearly (2) equals (3) and the experiment identifies both parameters (Heckman and Robb (1985)). However, in general, if $U_0 \neq U_1$, $E(\Delta|X) \neq E(\Delta|d = 1, X)$. See Heckman (1995) or Heckman and Smith (1995) for more discussion of this case.

[2] Heckman and Roselius (1994); Heckman, Ichimura, Smith and Todd (1994a,b, revised, 1995), Heckman, Ichimura and Todd (1993a,b, revised, 1995a,b) and Heckman, Ichimura and Todd (1995, revised 1996) document that failure of a common support condition is a major component of what is traditionally called selection bias.

IV. Randomization of Eligibility

Randomization of eligibility for a program is sometimes proposed as a less disruptive alternative to randomization of admission among accepted applicants (Heckman (1992), Heckman and Smith (1993), Angrist and Imbens (1991)). In this section, I show that this type of randomization can be placed in an instrumental variable framework. Consider a population of persons ordinarily eligible for a program. For simplicity, this conditioning is kept implicit. Let $e = 1$ if a person is kept eligible after randomization; $e = 0$ if the person loses eligibility. Assume that assignment to eligibility does not disturb the underlying stochastic structure and that it is independent with respect to the outcome measures:

$$(Y_0, Y_1, d, X) \perp\!\!\!\perp e. \tag{A-2}$$

Assuming $P(d = 1|X) \neq 0$,

$$\frac{E(Y|e = 1, X) - E(Y|e = 0, X)}{P(d = 1|X)} = E(\Delta|d = 1, X). \tag{11}$$

To prove this use the law of iterated expectations to obtain

$$\begin{aligned} E(Y|e = 1, X) = {} & E(Y_1|d = 1, e = 1, X)P(d = 1|e = 1, X) \\ & + E(Y_0|d = 0, e = 1, X)P(d = 0|e = 1, X) \end{aligned} \tag{12a}$$

and

$$\begin{aligned} E(Y|e = 0, X) = {} & E(Y_0|d = 1, e = 0, X)P(d = 1|e = 0, X) \\ & + E(Y_0|d = 0, e = 0, X)P(d = 0|e = 0, X). \end{aligned} \tag{12b}$$

From (A-2)

$$P(d = 1|e = 1, X) = P(d = 1|e = 0, X) = P(d = 1|X)$$

so that the result follows by subtracting 12(b) from 12(a) and dividing by $P(d = 1|X)$ provided $P(d = 1|X) \neq 0$. Replacing population moments by sample moments, (11) is a version of Bloom's estimator for attrition from a program. (See Angrist and Imbens (1991) and Heckman, Smith and Taber (1994, revised 1995) for a discussion of Bloom's estimator.)

I now present an instrumental variables interpretation of this estimator. Using equations (1a) and (1b), the law of iterated expectations and (A-2), the notation introduced in section II,

$$\begin{aligned} Y = {} & \varphi(X) + [E(\Delta|d = 1, X)]P(d = 1|X)e + \tilde{U}_0^* \\ & + (\tilde{U}_1^* - \tilde{U}_0^*)P(d = 1|X)e \end{aligned} \tag{13}$$

where

$$\tilde{U}_0^* = \tilde{U}_0 - E(\tilde{U}_0|d = 1, X)$$

$$\tilde{U}_1^* = \tilde{U}_1 - E(\tilde{U}_1|d = 1, X)$$

and

$$\varphi(X) = g_0(X) + E(U_0|d = 1, X).$$

Observe that by random assignment of e,

$$e \perp\!\!\!\perp \{\tilde{U}_0^* + e[\tilde{U}_1^* - \tilde{U}_0^*]P(d = 1|X)\}$$

so that orthogonality (really independence) is an immediate consequence of the randomization. Again, there is no requirement that X be independent or orthogonal with respect to $\tilde{U}_1$ or $\tilde{U}_0$.

Under standard conditions one can consistently estimate $E(\Delta|d = 1, X)P(d = 1|X)$. Assuming that one can consistently estimate $P(d = 1|X)$, one can estimate $E(\Delta|d = 1, X)$, the effect of treatment on the treated, by dividing the IV estimator of the product of the two terms by $P(d = 1|X)$. Note that this randomization identifies $E(\Delta|d = 1, X)$ but not $E(\Delta|X)$, except in the special cases where the two parameters are the same.

V. Concluding Remarks

This paper considers randomization as an instrumental variable. Two types of randomizations are considered: (a) randomization of eligibility for a program and (b) randomization of admission into the program among eligible persons who would ordinarily be admitted into the program. The second type of randomization is widely used in conducting social experiments. Using a conventional separable-in-the errors representation of the equations, I have shown the orthogonality conditions that are produced by the two types of randomization schemes and how they identify a central parameter in program evaluation studies—the effect of treatment on the treated—parameter (2). One randomization serves to identify this parameter as a function of multiple endogenous variables as conventionally defined in econometrics.

Balancing the bias in experimental and control samples is the fundamental source of identification from experiments. Such balancing in no way depends on separability of errors from equations as conventionally assumed in econometrics (as in equations (1a) and (1b)) nor does it require that the X be either independent or orthogonal with respect to the U. The method of moments analogs to (6) or (10) can be implemented nonparametrically. The balancing conditions are the basic estimating equations for experiments from which the orthogonality conditions of this paper have been derived.

The fact that parameter (2) is not conventional has been the source of some confusion. It combines both structural portions (the g_0 and g_1 of equation (1a) and (1b), respectively) with conditional means of the errors (the U_1 and U_0). Experiments conducted at a stage where persons would ordinarily enter a program are not designed to consistently estimate the g_0 and g_1 functions and in general they do not. Experiments make the treatment variable orthogonal to the error and the other regressors thus separating the estimation of treatment effects from the estimation of the other parameters of the model. Thus the data produced from social experiments are often not informative on the "deep structural parameters" of interest to many economists.

Only under special conditions does either type of randomization discussed in this paper identify parameter (3)—the effect of moving a randomly selected person in the general population from state "0" to state "1." This is an intrinsically more difficult parameter to estimate using social experiment because in most societies people cannot be forced to participate in programs against their will. It is more difficult to estimate $E(Y_1|d = 0, X)$ than $E(Y_0|d = 1, X)$ if responses to treatments are heterogeneous, and are partly anticipated at the time decisions to enroll in the program are made. By the same token, if there is attrition from the program, it may also be difficult to estimate (2) except under special conditions discussed in Bloom (1984), Heckman, Taber and Smith (1994, revised 1995), or Hotz and Sanders (1994).

Other types of randomization might be used besides the two types considered in this paper. For example, if interest centers on estimating

$$Y = g(X) + U$$

and X is not independent of U, X might be experimentally varied as in the negative income tax experiments or in the electricity experiments. In this case, experimental variation in X can be used in the conventional way to produce an instrument for an endogenous variable. See the discussion in Heckman (1992).

REFERENCES

Angrist, Joshua, and Guido Imbens, "Sources of Identifying Information in Selection Models," NBER Technical Working Paper 117 (1991).

Bloom, Howard, "Accounting For No-Shows in Experimental Evaluation Designs," *Evaluation Review* 82(2) (1984), 225–246.

Burtless, Gary, "The Case for Randomized Field Trials in Economic and Policy Research," *Journal of Economic Perspectives* 9 (June 1995), 63–84.

Clements, Nancy, James Heckman and Jeffrey Smith, "Making the Most Out of Social Experiments," first draft, 1993, under revision, *Review of Economic Studies* (1995).

Cox, David, *The Planning of Experiments* (New York: Wiley, 1958).

Fisher, Ronald, *The Design of Experiments*, 6th edition (London: Oliver and Boyd, 1951).

Goldfeld, Steven, and Richard Quandt, *Nonlinear Methods in Econometrics* (Amsterdam: North Holland, 1972).

Härdle, Wolfgang, *Applied Nonparametric Regression* (Cambridge: Cambridge University Press, 1990).

Heckman, James, "Randomization and Social Program," in C. Manski and I. Garfinkle (eds.), *Evaluating Welfare and Training Programs* (Cambridge: Harvard University Press, 1992).

———, "Instrumental Variables: A Study of Implicit Behavioral Assumptions in One Widely-Used Estimator," unpublished manuscript, University of Chicago (Aug. 1995).

Heckman, James, and Rebecca Roselius, "Evaluating the Impact of Training on the Earnings and Labor Force Status of Women: Better Data Help A Lot," unpublished manuscript, University of Chicago (Sept. 1994).

Heckman, James, and Richard Robb, "Alternative Methods for Evaluating the Impact of Interventions," in *Longitudinal Analysis of Labor Market Data* (New York: Wiley, 1985).

———, "Alternative Methods for Solving the Problem of Selection Bias in Evaluating the Impact of Treatments on Outcomes," in Howard Wainer (ed.), *Drawing Inferences from Self-Selected Samples* (Berlin: Springer, Verlag, 1986).

Heckman, James, Jeffrey Smith and Christopher Taber, "Accounting for Dropouts in Evaluations of Social Experiments," under revision, this REVIEW (1994, revised, 1995).

Heckman, James, and Jeffrey Smith, "Assessing the Case for Randomized Evaluations of Social Programs," in K. Jensen and P. K. Madsen (eds), *Measuring Labour Market Outcomes*, Ministry of Labor, Copenhagen, Denmark, 1993.

———, "Assessing the Case for Social Experiments," *Journal of Economic Perspectives* 9 (June 1995), 84–110.

———, "Evaluating the Welfare State," Frisch Symposium, Oslo, Norway (Mar. 1995b).

Heckman, James, Hidehiko Ichimura, Jeffrey Smith and Petra Todd, "Nonparametric Characterization of Selection Bias Using Experimental Data: A Study of Adult Males in JTPA, Part I: Definitions, Applications and Empirical Results" (Oct. 1994a, revised June 1995).

———, "Nonparametric Characterization of Selection Bias Using Experimental Data: A Study of Adult Males in JTPA. Part II, Theory and Methods and Monte Carlo Evidence" (Oct. 1994b, revised June, 1995).

———, "Matching as an Econometric Evaluation Estimator: Theory and Evidence on Its Performance Applied to the JTPA Program, Part I: Theory and Methods," unpublished manuscript, University of Chicago (Sept. 1993, revised Oct. 1995).

———, "Matching as an Econometric Evaluation Estimator: Theory and Evidence Applied to the JTPA Program, Part II. Empirical Evidence," unpublished manuscript, University of Chicago (Sept. 1993 revised Oct. 1995).

———, "Interpreting Standard Measures of Selection Bias," unpublished manuscript, University of Chicago (June 1995, revised Jan. 1996).

Hotz, V. Joseph, and Seth Sanders, "Bounding Treatment Effects in Controlled and Natural Experiments Subject to Post-Randomization Treatment Choice," Population Research Center, University of Chicago (1994).

Moffitt, Robert, "Evaluation of Program Entry Effects," in C. Manski and I. Garfinkle (eds.), *Evaluating Welfare and Training Programs* (Cambridge: Harvard University Press, 1992).

Rosenbaum, Paul, and Donald Rubin, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika* 70 (1984), 41–55.

Rubin, Donald, "Bayesian Inference for Causal Effects: The Role of Randomization," *Annals of Statistics* 6 (1) (1978), 34–58.

Roy, Andrew D., "Some Thoughts on the Distribution of Earnings," *Oxford Economics Papers* 3 (1951), 135–146.

THE EFFECT OF NEWS ON BOND PRICES: EVIDENCE FROM THE UNITED KINGDOM, 1900–1920

Douglas W. Elmendorf, Mary L. Hirschfeld and David N. Weil*

Abstract—We study the relationship of news to bond prices. We select a set of major news events based solely on their significance as judged by historians, and examine the corresponding bond price movements. The variance of holding returns is higher for weeks with important news than for weeks without such news, and the probability of a very large return (in absolute value) is higher for "news" weeks than for "non-news" weeks. The magnitude of these differences, however, suggests that much of the variability in bond prices cannot be explained by news, though important caveats about our measurement of news apply.

In traditional economic models, asset prices are determined by expectations of economic variables, and so changes in asset prices are caused by the arrival of news that changes these expectations. The high volatility of asset prices has led some observers to question this view, however. Recent theoretical work has proposed models in which asset prices move for reasons other than the arrival of news (see, for example, DeLong et al. (1990), Cutler, Poterba and Summers (1990), and Romer (1993)).

Our goal in this paper is to link news to asset prices in a systematic way. We select a set of major news events based solely on their significance as judged by historians, and examine the corresponding bond price movements. Our procedure does not constitute a formal test of the hypothesis that news causes all bond price movements. Rather, we make a serious effort to find relevant news and to see what fraction of movements it can explain. This approach is most closely related to Niederhoffer (1971), who examines the relation between the size of newspaper headlines and daily stock returns, and to Cutler, Poterba and Summers (1989), who look for specific news stories coincident with large stock price movements and also examine how prices change on important news days listed in an almanac.[1]

Received for publication December 10, 1992. Revision accepted for publication September 12, 1994.

* Board of Governors, Federal Reserve System; Occidental College; and Brown University and NBER, respectively.

We thank David Cutler, Greg Duffee, Karen Dynan, Greg Mankiw, Sabine Miltner, Louise Sheiner, Lawrence Summers, and the referees for helpful comments. The views expressed in this paper are our own and do not necessarily reflect the views of the Federal Reserve Board or its staff.

[1] In other related work, Shiller (1987) surveys investors at the time of large stock price movements, and concludes that "no news story or rumor . . . was responsible for investor behavior" in October 1987 or during previous large price declines. Roll (1984) compares the variance of orange juice futures prices on days with newspaper stories relevant to the orange market to the variance on days without such stories. He concludes that most price volatility cannot be explained by either these stories or relevant quantitative variables. See also Frankel and Meese (1987), French and Roll (1986), and Roll (1988).

Section I of this paper discusses the data and section II describes our methodology. In section III, we see how well news can explain bond price fluctuations in our data set. The final section summarizes our results and offers some conclusions.

I. Data

Bond Prices

We study weekly prices of British government consols from 1900 to 1920. This choice of time period and asset presents several advantages. First, World War I and the series of international crises leading up to it are a rich source of news. Second, the path of interest rates before, during, and after wars is interesting in its own right. Most economic theories predict that temporary government spending should raise both real and nominal interest rates. Previous research suggests that data for the United States do not support this proposition, but very long-run data for Britain do (Barro (1987) and Evans (1985)). Finally, Britain had the largest and most liquid capital market in the world in the early 20th century.[2] We examine government consols because they carry no individual company risk and because their long time horizon means that, of all assets, their prices should be the least responsive to transient money market conditions and short-term policies, which are the most difficult phenomena to measure with our techniques.

The holding period return on consols is defined as

$$h_t = 100 * \left[\frac{c_t + (p_t - p_{t-1})}{p_{t-1}} \right],$$

where p_t is the average of closing bid and asked prices for each Friday (or for the previous day if the market was closed on Friday), and c_t is the coupon payment (if any) received during that week. The quarterly coupon payment on consols was 0.69 pounds through 1902 and 0.63 pounds thereafter; this coming change in the coupon was known before our sample period began. The analysis excludes data from August through December 1914 when the market was closed entirely, and from December 1914 through November 1915, when there was a binding floor on the consol price.

[2] In 1913, Britain's holdings of foreign assets exceeded the combined overseas holdings of France, Germany, Holland, and the United States. Further, the world markets for money and gold were centered in London (Floud (1981) and De Cecco (1984)).

LINKED CITATIONS

Randomization as an Instrumental Variable
James J. Heckman
The Review of Economics and Statistics, Vol. 78, No. 2. (May, 1996), pp. 336-341.
Stable URL:
http://links.jstor.org/sici?sici=0034-6535%28199605%2978%3A2%3C336%3ARAAIV%3E2.0.CO%3B2-V

[Footnotes]

[1] **Assessing the Case for Social Experiments**
James J. Heckman; Jeffrey A. Smith
The Journal of Economic Perspectives, Vol. 9, No. 2. (Spring, 1995), pp. 85-110.
Stable URL:
http://links.jstor.org/sici?sici=0895-3309%28199521%299%3A2%3C85%3AATCFSE%3E2.0.CO%3B2-K

References

The Case for Randomized Field Trials in Economic and Policy Research
Gary Burtless
The Journal of Economic Perspectives, Vol. 9, No. 2. (Spring, 1995), pp. 63-84.
Stable URL:
http://links.jstor.org/sici?sici=0895-3309%28199521%299%3A2%3C63%3ATCFRFT%3E2.0.CO%3B2-Q

Assessing the Case for Social Experiments
James J. Heckman; Jeffrey A. Smith
The Journal of Economic Perspectives, Vol. 9, No. 2. (Spring, 1995), pp. 85-110.
Stable URL:
http://links.jstor.org/sici?sici=0895-3309%28199521%299%3A2%3C85%3AATCFSE%3E2.0.CO%3B2-K

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

The Central Role of the Propensity Score in Observational Studies for Causal Effects
Paul R. Rosenbaum; Donald B. Rubin
Biometrika, Vol. 70, No. 1. (Apr., 1983), pp. 41-55.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28198304%2970%3A1%3C41%3ATCROTP%3E2.0.CO%3B2-Q

Bayesian Inference for Causal Effects: The Role of Randomization
Donald B. Rubin
The Annals of Statistics, Vol. 6, No. 1. (Jan., 1978), pp. 34-58.
Stable URL:
http://links.jstor.org/sici?sici=0090-5364%28197801%296%3A1%3C34%3ABIFCET%3E2.0.CO%3B2-X

NOTE: *The reference numbering from the original has been maintained in this citation list.*