AN ECONOMETRIC ANALYSIS OF AN INSTITUTIONAL FORECASTING MODEL FOR U.S. ARMY ENLISTMENTS

by

J. Kevin Berner

B.S., United States Military Academy
(1979)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

MASTER OF SCIENCE
IN ECONOMICS

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
July 1986

Signature of Author____
rtment of Economics
July 14, 1986

Certified by_ 6-18-86
Jerry A. Hausman
Thesis Supervisor

Accepted by____
Richard S. Eckaus
Chairman, Departmental Committee on Graduate Students

AN ECONOMETRIC ANALYSIS OF AN INSTITUTIONAL FORECASTING MODEL FOR U.S. ARMY ENLISTMENTS

by

J. Kevin Berner

Submitted to the Department of Economics on July 14, 1986 in partial fulfillment of the requirements for the Degree of Master of Science in Economics

ABSTRACT

This paper is an econometric analysis of a forecasting model for the number of high quality males expected to enlist in the U.S. Army. Prior to this work, the model has been estimated on a panel data set over the period October 1980 to June 1983. This paper reestimates the model over the period October 1980 to September 1984. A Chow test is performed under two alternative error structures, but the hypothesis that the model is stable is always rejected. A Scheffé multiple comparison procedure reveals that the rejection is not caused by an economically uninteresting contrast.

A test for the presence of an AR1 structure confirms that observations within recruiting battalions are serially dependent. Without specifying the precise nature of this autocorrelation, the model is estimated with a general random effects error structure using the two-stage instrumental variables technique. Several key variables have the incorrect sign. An explanation of this is offered, and it is recommended that the model not be used in policy applications until more work is done on the specification of the entire system of equations.

Thesis Supervisor: Dr. Jerry A. Hausman

Title: Professor of Economics

ACKNOWLEDGEMENTS

The author wishes to thank all the many people who were of assistance in the preparation of this paper. Jerry Hausman's instruction and comments were indispensable. Franklin Fisher's comments on an earlier draft were quite helpful. The cooperation of Dave Smith and the staff of the Office of Economic and Manpower Analysis at West Point is gratefully acknowledged. Glenn Sueyoshi was very generous with his assistance in overcoming computing difficulties. The comments and support of Anil Kashyap and Mike Edleson were much appreciated. And finally, I owe much gratitude to Leslie E. Papke for her inputs into every aspect of this work.

TABLE OF CONTENTS

I. INTRODUCTION

Since the United States relies on economic incentives to attract the large numbers of young men and women needed to fill its Armed Forces, the determinants of military labor supply are of considerable public policy concern. Of particular interest is whether the level of military compensation is sufficient to make the services attractive alternatives to civilian employment for a sufficient number of high quality young men and women. Forecasting models are currently in use by economists to inform policy makers for the U.S. Army. This paper is an econometric analysis of the forecasting model for Army high quality enlistments proposed by Daula and Smith (1985).

After a short review of the problem of military labor supply in Section II, the work of Daula and Smith is examined more closely in Section III. The distinction between their structural model of enlistments and their forecasting model is drawn, and the forecasting model is analyzed in considerable detail in Sections IV and V.

In Section IV the stability of the forecasting model over fluctuating macroeconomic conditions is addressed and tested. Daula and Smith estimate the model over a period which is dominated by a deep recession. Using data provided the United States Military Academy, the model is reestimated over a second period of macroeconomic recovery, and a Chow test is performed. This test rejects the hypothesis that the model is stable. An alternative error structure is proposed and estimated, but the Chow test still rejects this hypothesis. A Scheffé multiple comparison procedure provides some insight into this rejection.

In Section V the presence of serial correlation is discussed as a reasonable consequence of the institutional behavior which motivates the structural model of Daula and Smith. A test verifies the presence of a first order autoregressive process in the data. A general error structure is proposed, and the forecasting specification is then estimated using the Two-Stage Instrumental Variables technique.

The paper concludes with a summary and some recommendations to improve this forecasting model.

II. THE MILITARY LABOR SUPPLY PROBLEM REVIEWED

Economists first turned their attention to the issue of military labor supply in the late 1960's to advise policy makers about theoretical and quantitative matters concerning the economic costs of the draft. The earliest papers were essentially feasibility studies designed to estimate the cost of maintaining sufficient numbers of personnel to meet the requirements of our large standing Armed Forces using voluntary, economic incentives rather than the compulsary selective service.[1]

These studies concluded that the budgetary cost of replacing conscription with an all-volunteer force would not be prohibitively expensive, increasing personnel costs by approximately $4 billion per year in 1967 dollars. Encouraged by these studies, and amid considerable political pressures, Congress enacted the necessary legislation to require the Department of Defense to maintain its 2.65 million member Armed Forces as an all-volunteer force.

One of the best of these early studies is that of Fisher (1969). His work is explicit in its treatment of such issues as the distribution of civilian employment opportunities and compensating wage differentials. An important contribution of this work is his characterization

[1] For examples of these see Altman (1967), Altman and Fechter (1967), Oi (1967), and Weisbrod and Hansen (1967).

of the individual decision making process. The individual compares discounted cash flows from military service and some best civilian alternative, and enlists if the military compensation exceeds the civilian compensation. Fisher accounts for personal tastes by adding a compensating wage differential to the civilian wage.

Holding tastes constant, an individual with more lucrative civilian alternatives will not enlist. This description of the military labor market is not Walrasian because prices do not fluctuate to equate supply and demand. Of course, most labor markets have wages which are rigid downward. But the military labor market is characterized by a completely rigid military wage which changes only once a year. It is the quality of recruits which fluctuates to clear the market.

The Armed Forces are limited to a certain end strength by law. The Department of Defense (DOD) establishes some measure of quality control by pre-enlistment testing, denying enlistment to those who do not meet certain minimum qualifications, and restricting the numbers of recruits who score below average.

This problem of demand constraints complicates the economics of military labor supply.[2] Fisher's treatment of demand constraints, while brief, has greatly influenced subsequent studies. Since his work, most researchers have

[2] It should be noted that female recruits are always considered to be demand constrained because the legal restrictions on the types of jobs in which they may serve severely limit their numbers.

attempted to abstract from the problem of demand constraints by using some measure of high quality enlistments as the dependent variable in their regressions.[3] High quality recruits are assumed to be supply constrained, so that the occurrence of a high quality enlistment is considered to be a labor supply decision.

A second complication which is frequently addressed in the literature is how to appropriately define a high quality enlistment. There are two aspects of this debate. The first issue is to define when the enlistment occurs. An individual desiring to enlist can sign a contract to join the Army in one month and delay his or her accession into the Army for up to nine months under the Delayed Entry Program (DEP). One can measure the enlistment at the time the contract is signed or at the time of the accession. There is general agreement that the contract measure is the more appropriate one for the labor supply decision.[4] Accessions are very highly concentrated in the summer months, while contracts are less seasonal. Studies which use accessions, such as Ash, Udis, and McNown (1983), tend to confuse the seasonal effect of schooling with the labor supply effect which they are investigating. These studies produce unreliable estimates of unemployment elasticities, as discussed in Dale and Gilroy (1985).

---

[3] There are other reasons for this, such as recent evidence that high quality recruits make better soldiers. See Baldwin and Daula (1985).

[4] See, for example, Ash, Udis, and McNown (1983), and particularly Figures 1 and 2 in Dale and Gilroy (1985). Daula and Smith (1985) also show its importance in testing whether high quality enlistments are demand constrained.

The second aspect of this debate is the definition of "high quality." There are two popular definitions, the first being whether or not the recruit is a high school graduate. A high school senior who signs an enlistment contract under the DEP is usually considered to be a high school graduate. Before enlisting, individuals must take the Armed Services Vocational Aptitude Battery (ASVAB) of tests. The scores for the four tests on mathematical and verbal skills are combined to give the Armed Forces Qualification Test (AFQT). When an individual scores above average on the AFQT, he or she is considered to be in enlistment categories I-IIIA. This is an alternative definition of high quality. A conservative approach is to consider a person a high quality recruit only if they meet both criteria. This is becoming increasingly popular.[5]

Since the advent of the All-Volunteer Force in 1973, there have been several studies by economists to assess the progress of the services and to help establish appropriate pay levels by estimating forecasting equations. These studies vary in approach and emphasis, and sometimes arrive at very different conclusions. The usual approach is to regress some measure of high quality enlistments on a set of explanatory variables which always includes some measure of relative military-to-civilian pay and unemployment. These two variables are the most relevant for policy analysis. The dependent variable is usually an enlistment

[5] The most recent studies all include this in at least one specification. See Ash, Udis, and McNown (1983), Brown (1985), and Daula and Smith (1985).

rate, and the functional form of the specification is almost always log-linear. The earliest studies employed time series or cross-sectional data. More recently, panels have been constructed to follow enlistments in a geographical area over time.

The following elasticities of high quality enlistments bracket the existing literature.[6] The lowest relative pay elasticity is 0.88 in Fernandez (1979). This is one of the few relative pay elasticity estimates which is below 1.0. Goldberg (1982) has the highest estimate at 2.13. The lowest unemployment elasticity is also in Goldberg (1982), where he finds that Navy enlistments are perfectly inelastic with respect to unemployment.[7] Dale and Gilroy (1983) find the largest unemployment elasticity at 0.94. In general, high quality enlistments are considered to be quite elastic with respect to relative pay, but somewhat inelastic with respect to unemployment.

The Goldberg (1982) and Dertouzos (1983) studies begin to show interest in the effects of recruiting efforts on enlistment. In particular, Dertouzos (1983) points out that asymmetric recruiter incentives cause distortions in the labor supply decisions as viewed by the economist. Recruiters are given monthly objectives for high quality

---

[6] A more complete listing of these elasticities may be found in Table 4 of Daula and Smith (1985).

[7] It is perhaps not surprising that Navy enlistments are not much affected by unemployment. Oi (1967) points out that both before and during the Vietnam War, very few draftees were needed to meet Navy requirements. The Navy seems to have a long history of meeting its strength needs with volunteers. The lowest unemployment elasticity estimated with Army data is 0.133 in Ash, Udis, and McNown (1983).

enlistments. There are few rewards for exceeding these objectives, but repeated failures to meet them brings adverse consequences. The recruiter therefore has an incentive to exactly meet the objective each month, and to delay any additional enlistments until they can be applied against the next month's objective. If this is so, it casts considerable doubt on the assumption that high quality recruits are always supply-constrained.

Many of these studies fail to consistently estimate the parameters of their models. A typical example of this is Brown (1985). Brown constructs a panel which looks at enlistment by state by quarter from 1975-1982. His relative pay and unemployment elasticities are 1.0 and 0.5, respectively. These estimates suffer from several sources of inconsistency. First, Brown recognizes that the separate services compete among themselves for high quality recruits, so that his estimates are subject to simultaneity bias. Additionally, because recruiters must devote considerable amounts of time to enlist other categories of recruits, it is likely that the recruiting effort causes these other categories of enlistments to be jointly endogenous with high quality enlistments. Reasoning that the specification of the entire system is beyond the scope of his work, Brown does not correct for this. Even so, he should have used a limited information approach to estimating one equation in the system.

A second source of inconsistency is errors in variables. Brown admits that the relative pay variable is constructed with considerable error.[8] The traditional solution for this problem in simultaneous equations treats an exogenous variable measured with error as an endogenous variable (Hausman, 1977). This approach requires that the equation be "conditionally" overidentified, or that some other instrument be available from outside the system. Griliches and Hausman (1986) calculate the inconsistent probability limit of the fixed effects estimator in the presence of errors in variables in panel data, indicating that Brown's fixed effects estimates are inconsistent even if there was no simultaneity bias.

[8] This problem is common to every study of this topic.

III. THE DAULA AND SMITH STUDY

The work of Daula and Smith (1985) makes three important contributions to the theory of military labor supply. First, they broaden the specification of the enlistment equation to account for the institutional behavior of the U.S. Army Recruiting Command (USAREC). Second, they test the fifteen year old assumption that high quality recruits are supply constrained and reject it quite strongly. Third, they are able to develop consistent parameter estimates for their structural model.

Daula and Smith (hereafter DS) are able to incorporate institutional behavior by constructing a panel of monthly data which has as its units of observation the recruiting battalions of USAREC. The recruiting effort is captured by such explanatory variables as the recruiting objectives for the various categories of enlistments, the number of recruiters which the Army employs in a given battalion, and the amount of advertising which is done in national and local media. The effectiveness of advertising has long been in doubt. Although recruiters are often told by recruits that the Army's national advertising is quite effective, no enlistment study has been able to show this. By disaggregating the data, and measuring national advertising in impressions in the media rather than in dollar expenditures, DS find that national advertising is quite significant.

To test the hypothesis that high quality recruits are supply constrained, DS look at the effects of Army recruiting goals. If asymmetric incentives for recruiters distort the enlistment figures seen by the economist, then when a recruiting battalion meets or exceeds its enlistment goal, that goal will be a significant variable. Conversely, when a recruiting battalion does not meet its goal, the goal will not be significant. DS split the sample according to whether the goal was met, and find that this is exactly the case. This leads them to a switching model of enlistment which they propose as their structural model. When the enlistment goal is not met, the battalion is considered to be supply constrained, and the economist actually observes labor supply behavior.[9] When the goal is met or exceeded, then the battalion is considered to be demand constrained, and the economist observes the outcome of an enlistment production process.

The third contribution of DS is that in addressing simultaneity directly, they consistently estimate the parameters of their structural equation. DS allow for the joint endogeneity of other service enlistments as well as that of the enlistments of other categories of Army enlistments. Estimation is carried out using limited information instrumental variables techniques. Additionally, DS allow for the relative pay variable to be constructed with error. As the equation is conditionally

[9] DS use Category I-IIIA high school graduate and high school senior contracts as their dependent variable.

underidentified in the language of Hausman (1977), they use an instrument from outside the system to identify the equation. The Durbin instrument is used for relative pay. This instrument is obtained by assigning to each observation of the relative pay variable its rank order.

Sample selectivity is introduced by the switching nature of the structural model. Consistent parameter estimates are obtained by using the two-step estimation procedure of Heckman (1979). This procedure does not give a consistent estimate of the asymptotic covariance matrix of the estimator. But DS are able to obtain consistent standard errors with the method of Lee, et al. (1980).[10] For more details on the specification of this structural model, the reader is referred to DS (1985).

The structural model is useful for understanding the economics of military labor supply, but to be useful for policy applications, the model must have some forecasting ability. There are two forecasts which are of interest. Given that DS find evidence that the enlistment goals are significant variables in the structural model, it would be useful to have a method of establishing these for USAREC. DS find that their structural model estimated with fixed effects is the best forecasting model for use in establishing enlistment goals. This allows the commander of USAREC to set challenging but realistic goals for the recruiting battalions. Although it is the most accurate of the DS models, it still has a very large percentage

[10] Amemiya (1985) addresses these problems.

prediction error. While they recognize that this may be a good management tool, they recommend that other factors be considered when establishing these enlistment goals.

The second, and perhaps more important, forecast is how many total enlistments the Army can expect. The structural model of DS does not forecast this well. They note that this is probably because as economic conditions deteriorate, and more recruiting battalions become supply constrained, the elasticities of total enlistments changes, while the model was estimated by imposing constant elasticities. A within-sample forecasting experiment reveals that this is a reasonable explanation.

DS then turn to other specifications in the paper to find a better forecasting model. Using an ex-post forecasting test, they settle on one particular specification as their best model for forecasting levels of enlistments. The model was estimated over the subsample October 1980 to June 1982, and forecasts were obtained for each of the next two years. The forecasts of this best-fit equation had 0.0 and 2.6 percent prediction errors, respectively. The model seems to forecast quite well.

The remainder of this paper examines this forecasting equation more closely. The specification of the model is similar to the structural model. As discussed earlier, DS use as their measure of high quality recruits the number of male high school seniors and high school graduates in enlistment categories I-IIIA who sign a contract to enlist in the Army. The explanatory variables may be grouped into four categories: recruiting competition variables, recruiting effort variables, socioeconomic variables (including relative pay), and dummy variables to control for certain events. Each category is discussed in turn.

The recruiting competition variables measure other categories of enlistment which are considered to be simultaneously determined with the dependent variable. These are: CNHSG, the count of Army male category I-IIIA non-high school graduate (or senior) enlistments; COTH, the count of Army male enlistments in other categories; and CDOD, the count of male category I-IIIA high school senior or graduates who enlist in other DOD services. CNHSG and COTH are considered to be jointly determined with CGRAD since they represent competing demands on the time of recruiters. CDOD is considered to be endogenous because all the services are competing for high quality recruits from the same population of eligible males. Accordingly, the signs of these variables should be negative.

The recruiting effort variables include those variables over which USAREC has discretionary control. These are: MGRAD, the mission (or objective) for Army male category I-IIIA high school senior or graduate enlistments; RECR, the number of production recruiters assigned to a recruiting battalion; REXP, the number of those production recruiters with at least nine months of recruiting experience in the battalion; LADV, the dollar expenditure disbursed by the recruiting battalion for advertising in the local media; and NADV, the number of media impressions in the national electronic and print media.[11] These variables are measured so that an increase in the recruiting effort should, ceteris paribus, cause an increase in the number of high quality enlistments, so that their coefficients should be positive.

The socioeconomic variables are: RPAY, the relative military-to-civilian wage; UR, the unemployment rate; QMA, the estimated number of males in the recruiting battalion area who qualify to enlist in the military and are category I-IIIA high school seniors or graduates; MIN, the percentage of minorities in QMA; and VOTE, the percentage of the vote in the 1980 Presidential election which was cast Republican (this is included as a measure of pro-military feeling). Relative pay and the local unemployment rate are measured so that increases relect an improvement in the recruiting environment. There is a strong a priori prediction from the theory of compensating wage differentials that these variables should have a positive coefficient.[12] Likewise QMA and VOTE should have positive coefficients. Other cross-sectional studies have found negative coefficients for MIN. DS believe that this may be caused by unmeasured differences across regions which are correlated with the racial distribution.

The event variables are: Q2, Q3, Q4, to control for the quarter of the fiscal year (Q4 is the months July,

---

[11] This resource is not controlled by local recruiters, but rather at Headquarters, USAREC.
[12] See, for example, Marshall (1952), pp. 547-570, or Ehrenberg and Smith (1985), Ch. 8.

August, and September and more contracts are signed in this quarter than in any other); ACF, NCVP, BILL, control for experiments with the Army College Fund, Non-contributory Veteran's Educational Assistance Program (VEAP), and the Mini GI Bill, respectively; and BONC, BON8K,and BON84K, to control for experiments with the enlistment bonus (These are for the control group, the $8000 bonus group and the $8000 or $4000 option bonus group).

There are 23 explanatory variables. The specification is log-linear, so the natural logarithms of the dependent variable, the recruiting competition variables, the recruiting effort variables and the socioeconomic variables are used. For more detail conerning the data, the reader is referred to the Data Appendix.

As discussed earlier, the recruitng competition vairables and RPAY are such that plim $(X'\epsilon/T) \neq 0$. Therefore, estimation is accomplished using the following instrumental variables.

Variable	Instrument
CNHSG	MNHSG, the Army recruiting mission for this category of enlistment.
COTH	MOTH, the Army recruiting mission for this category of enlistment.
CDOD	DRECR, the number of other service DOD recruiters stationed in the battalion's geographical area.
RPAY	RANK, the Durbin instrument.

DS estimate this model using two-stage least squares (2SLS), under the usual assumptions for simultaneous equations. Namely, for the system of equations,

$$YB + Z\Gamma = U, \tag{1}$$

we estimate the first equation,

$$Y\beta_1 + Z\Gamma_1 = \mathbf{u}_1 . \tag{2}$$

From Hausman (1983), the assumptions are:

(A1) B is nonsingular.

(A2) Z has full column rank equal to s.

(A3) The rows of U are iid. U has mean zero and non-singular covariance matrix $\Sigma \otimes I_T$ (where T is the number of observations on each equation).

Assumption (A3) is of great importance. In particular, it implies that the first column of U, $\mathbf{u}_1$, has mean zero and covariance matrix $E(\mathbf{u}_1\mathbf{u}_1') = \sigma_{11} I_T$. That is, this assumes that the disturbances in the first equation are independent and identically distributed.

After a nomalization, it is possible to rewrite equation (2) as

$$\mathbf{y}_1 = Y_1\beta_1 + Z_1\Gamma_1 + \mathbf{u}_1 = X_1\delta_1 + \mathbf{u}_1 . \tag{3}$$

Under assumptions (A1) to (A3) 2SLS is consistent and asymptotically efficient in the class of all instrumental variable estimators. Then

$$\delta_{2SLS} = (X_1'Z(Z'Z)^{-1}Z'X_1)^{-1}X_1'Z(Z'Z)^{-1}Z'\mathbf{y}, \tag{4}$$

$$\text{and, } \operatorname{avar}(\delta_{2SLS}) = \sigma_{11}(X_1'Z(Z'Z)^{-1}Z'X_1)^{-1} . \tag{5}$$

DS estimate the model using equations (4) and (5). Table 1 presents the results of their regression over the period October 1980 to June 1983 (hereafter, the first subsample) in Column (1). Due to some differences in the data used by DS and in this paper, the forecasting equation is reestimated over the first subsample and the results are reported in Column (2)[13]. Column (3) contains the results of estimating the equation over the entire period October 1980 to September 1984. A quick comparison of the first and second columns shows that the differences in the data do not greatly affect the parameter estimates.

---

[13] Of the 56 recruiting battalions, DS exclude the Miami, FL and San Juan, PR battalions because some observations on certain variables were missing. Since then, the problems with the Miami battalion have been solved, and it is included in all regressions in this paper, except Column (1), Table 1. The Beckley, WV and Nashville, TN battalions are excluded due to some reporting problems in the second subsample. The regressions in this paper, therefore, represent $n = 53$ recruiting battalions. There are $p_1 = 33$ observations in the first subsample, $p_2 = 15$ in the second subsample, for a total of $p = 48$ observations on each recruiting battalion. This gives $T = np = 2544$ total observations.

TABLE 1
COMPARISON OF 2SLS RESULTS
DEPENDENT VARIABLE: ARMY HIGH QUALITY ENLISTMENTS

EXPLANATORY VARIABLE	(1)	(2)	(3)
CONSTANT	-1.43	0.045691	0.643801
	(0.494)	(0.471654)	(0.381101)
MGRAD	0.406	0.459668	0.528874
	(0.024)	(0.035818)	(0.032441)
CNHSG	0.113	0.152388	0.146508
	(0.032)	(0.041742)	(0.033736)
COTH	-0.018	-0.072518	-0.127871
	(0.055)	(0.046274)	(0.037998)
CDOD	-0.416	-0.379968	-0.290989
	(0.180)	(0.190302)	(0.131403)
RPAY	0.494	0.510924	0.350580
	(0.115)	(0.148222)	(0.093104)
UR	0.562	0.564204	0.403671
	(0.068)	(0.089477)	(0.046716)
QMA	0.142	0.083903	-0.015259
	(0.086)	(0.053187)	(0.032633)
MIN	-0.050	-0.109718	-0.076149
	(0.015)	(0.017116)	(0.012815)
RECR	0.585	0.495687	0.500391
	(0.099)	(0.089526)	(0.074777)
REXP	0.028	-0.033590	-0.073017
	(0.042)	(0.048124)	(0.040714)
LADV	0.035	0.018934	-0.013896
	(0.017)	(0.018263)	(0.014498)
NADV	0.089	0.064675	0.074324
	(0.020)	(0.015469)	(0.010237)
VOTE	0.144	0.051182	0.014436
	(0.087)	(0.080190)	(0.062458)
Q2	0.072	0.086659	0.038335
	(0.025)	(0.025175)	(0.018968)
Q3	0.096	0.162540	0.083755
	(0.020)	(0.023288)	(0.017646)
Q4	0.244	0.250962	0.227759
	(0.049)	(0.035715)	(0.024011)
ACF	0.083	0.039954	0.050919
	(0.026)	(0.034313)	(0.029016)
NCVP	0.015	-0.003425	0.016121
	(0.034)	(0.035469)	(0.033753)
BILL	0.031	0.068269	0.082759
	(0.036)	(0.039920)	(0.037658)
BONC	0.007	0.045772	0.024682
	(0.036)	(0.028887)	(0.019773)
BON8K	-0.002	0.007989	0.025873
	(0.047)	(0.042965)	(0.025385)
BON84K	0.049	0.063925	0.074571
	(0.043)	(0.039080)	(0.025179)
Observations	1782	1749	2544
Std error of Regr.	0.292	0.304	0.290

(1)Model (2), Table 3, of Daula and Smith (1985).
(2)Results using additional data on the first subsample.
(3)Results using additional data on the expanded sample.

The first subsample is dominated by the worst recession in the U.S. economy since the Great Depression. Additionally, the 12 percent 1981 military pay raise served to make relative military-to-civilian pay quite high. Examining Column (2) of Table 1, one finds that RPAY and UR are both positive and quite significant. Consistent with past results, the unemployment elasticity of 0.564 is relatively inelastic. But the relative pay coefficient of 0.511, while positive, is also relatively inelastic. This is because a large portion of the recruiting battalions were demand constrained over the first subsample. When DS estimate this same model over supply constrained observations, the elasticity of enlistment with respect to relative pay is found to be 1.89. This is the appropriate measure of this elasticity.

The effect of recruiting competition does not appear to be very strong. Two of the three recruiting competition variables have a negative coefficient, reflecting the effects of the competition, although they are not precisely estimated. The other variable, CNHSG, is positive and significant. One possible explanation for this is that over the sample period there was some sort of group dynamic which dominated the effects of competition. As an example of this, one might think of several friends visiting the recruiting station together.

QMA has the correct sign but is not measured with enough precision to be significantly different from zero.

The evidence that the coefficient on MIN is negative is quite strong in these data, supporting earlier cross sectional evidence. It is possible that high quality minorities are in high demand in the private sector and in universities, and therefore show less interest in enlisting in the Army. Evidence from other services might help to explain this result. The coefficient for VOTE is negative, but insignificant. This reflects that individuals vote for more reasons than merely pro-military feeling, and that the overlap between voters and those who enlist is not perfect.

The recruiting effort variables REXP and LADV are not significant. However, both RECR and NADV are positive and highly significant. It is not surprising to find that production recruiters produce recruits. Given that no other study has found a significant effect for advertising, it is clear that media impressions is a superior measure. We also see that MGRAD, a measure of the Army's demand for enlistments, is highly significant in this forecasting specification.[14]

---

[14] Certainly, a large portion of this significance results from the fact that recruiting missions are heavily weighted toward past performance, so that MGRAD closely follows lagged values of the dependent variable. MGRAD is considered to be predetermined for our purposes. For comparison, when this specification of the structural model, corrected for selectivity, is estimated over supply constrained observations only, MGRAD is insignificant.

IV. THE STABILITY OF THE MODEL SINCE 1983:4

Since the first subsample is dominated by a recession, one might wonder how robust the forecasting specification is to fluctuating macroeconomic conditions. The second subsample is dominated by the 1983 to 1984 recovery. Improving economic conditions worsen the recruiting environment. Estimation of this model over the entire sample would give some confidence in the specification. As Column (3) of Table 1 shows, the model seems to be stable over the entire period.

This notion of the stability of the model can be tested in a precise statistical manner with a Chow test. Chow (1960) shows that one can test for the equality of the coefficients of two regressions by appropriately stacking the model, estimating a restricted and unrestricted coefficient vector, and obtaining the restricted and unrestricted sum of squares, RSS and USS, respectively. For an ordinary least squares regression, and under the null hypothesis that the two coefficient vectors are equal, the statistic $F = ((RSS-USS)/q)/(USS/(T-k))$ has an exact Fisher's $F(q,T-k)$ distribution.[15]

When using instrumental variables regression, however, this familiar form of the F-test is not appropriate. Because IV does not minimize the sum of squared residuals

[15]See, for example, Chow (1960), Judge, et al.(1982), pp.189-203, and problem 7.8 of section 3.7 in Theil (1971).

as OLS does, it is possible for the RSS to be less than the USS.[16] 2SLS does, however, minimize the sum of squared residuals after they have been projected into the space spanned by the instruments, as shown in Hausman (1983). A Chow test may still be performed as long as the restricted and unrestricted residuals are projected onto the column space of the same instruments. This is most easily accomplished using the Wald form of the test. Dropping the subscript for the first equation and using superscripts to denote subsamples, we can stack equation (3) as

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}^1 \\ \mathbf{y}^2 \end{bmatrix} = \begin{bmatrix} X^1 & 0 \\ 0 & X^2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\delta}^1 \\ \boldsymbol{\delta}^2 \end{bmatrix} + \begin{bmatrix} \mathbf{u}^1 \\ \mathbf{u}^2 \end{bmatrix} = X\boldsymbol{\delta} + \mathbf{u} \tag{6}$$

Using the stacked model (6) is equivalent to estimating the model on each subsample separately. Under the assumptions (A1) to (A3), the unrestricted coefficient vector is

$$\boldsymbol{\delta}_{2SLS} = (X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'\mathbf{y}, \tag{7}$$

$$\text{and } (T^{1/2})(\boldsymbol{\delta}_{2SLS} - \boldsymbol{\delta}) \overset{A}{\sim} N(0,\ \sigma_{11}(X'P_Z X)^{-1}), \tag{8}$$

where $Z = \begin{bmatrix} Z^1 & 0 \\ 0 & Z^2 \end{bmatrix}$ and $P_Z = Z(Z'Z)^{-1}Z'$. We can then test the general linear hypothesis $R\boldsymbol{\delta} = 0$, where $R = (I_q \mid -I_q)$. Under this null hypothesis, we have from equation (8) that

$$(T^{1/2})R(\boldsymbol{\delta}_{2SLS} - \boldsymbol{\delta}) \overset{A}{\sim} N(0,\ \sigma_{11}R(X'P_Z X)^{-1}R'). \tag{9}$$

---

[16] This occurs in the present case, where RSS = 212.672 and USS = 241.718, yielding an "F-statistic" of -15.85.

Assuming an equal variance for both periods, the Wald form of this Chow test is

$$W = (1/\sigma_{11})(R\delta)'(R(X'P_Z X)^{-1}R')^{-1}(R\delta), \tag{10}$$

and this statistic is asymptotically distributed as Chi-Squared with q degrees of freedom.

The unrestricted coefficient vector[17] is presented in Table 2 with like coefficients for the two periods presented side by side. The associated 2SLS standard errors are presented in the row immediately beneath the coefficients in Table 2. An asymptotic Chi-Squared (19) test of the null hypothesis that the entire coefficient vector is stable over the two subsamples yields a test statistic of 143.78. This rejects the null hypothesis at any reasonable significance level. It might be important to know, however, whether the estimated coefficients of principal policy concern are stable with this forecasting specification. An asymptotic Chi-Squared (6) test that the coefficients for MGRAD, RPAY, UR, RECR, LADV, and NADV are stable yields a test statistic of 75.60. Again, this rejects the null hypothesis at any reasonable level of

---

[17]Four of the dummy variables represent events which do not occur in the second subsample. Including them in the specification causes a singular submatrix in (Z'Z), and so they were excluded. This does not greatly affect any of the results, as is evident by comparing the unrestricted coefficients and 2SLS standard errors from Period 1 in Table 2 to the corresponding estimates in Column (2) of Table 1.

significance.[18] The data are quite emphatic in rejecting the stability of the model.

Recall that assumption (A3) of the simultaneous equations model is that $E(u_1 u_1') = \sigma_{11} I_T$. This does not seem like a reasonable assumption for our data. Panel data is characterized by a block diagonal error structure. Judge, et al.(1985) show that under the usual panel data assumptions,

$$\text{(A3)}' \quad E(u_1 u_1') = \Omega = \sigma_{tp}^2 I_{np} + \sigma_\alpha^2 (I_n \otimes J_p), \tag{11}$$

where n=the number of observational units in the panel, p=the number of observations on each unit in the panel, and J_p is a (pxp) matrix of ones.[19]

---

[18] The P-values for these two test statistics are 1.54E-09, and 2.88E-14, respectively. The P-value is defined as the area under the Chi-Squared(q) probability density function from the test statistic to positive infinity. It is the significance level at which the test would not reject the null hypothesis.

[19] There is reason to believe that assumption (A3)' is not quite correct, either. This is discussed in the next section.

TABLE 2
CHOW TEST RESULTS

THE UNRESTRICTED COEFFICIENT VECTOR
(alternative measures of standard errors are computed with the usual and block diagonal error structures, respectively)

Var	Period 1	Period 2	Var	Period 1	Period 2
CONST	0.157292	1.353330	REXP	-0.036279	-0.217446
	(0.468186)	(0.723751)		(0.049798)	(0.106423)
	(1.229700)	(1.367728)		(0.158207)	(0.248922)
MGRAD	0.478541	1.031634	LADV	0.018708	-0.030570
	(0.031090)	(0.125876)		(0.018744)	(0.030362)
	(0.068248)	(0.457530)		(0.041001)	(0.052244)
CNHSG	0.179552	0.160957	NADV	0.073252	0.147508
	(0.038524)	(0.087588)		(0.014606)	(0.033045)
	(0.091552)	(0.203622)		(0.018900)	(0.060239)
COTH	-0.105370	-0.429156	VOTE	0.041444	0.160442
	(0.048749)	(0.124052)		(0.081907)	(0.125395)
	(0.132129)	(0.359600)		(0.326273)	(0.355102)
CDOD	-0.443735	-0.556237	Q2	0.076829	-0.188889
	(0.205961)	(0.238294)		(0.026070)	(0.045398)
	(0.710864)	(0.858041)		(0.050530)	(0.098215)
RPAY	0.592384	0.139317	Q3	0.167113	-0.179541
	(0.125859)	(0.166443)		(0.023958)	(0.043499)
	(0.900758)	(1.344239)		(0.058783)	(0.070492)
UR	0.617460	0.178196	Q4	0.265048	0.096566
	(0.076936)	(0.071362)		(0.035882)	(0.038727)
	(0.359016)	(0.415252)		(0.116970)	(0.071785)
QMA	0.096380	-0.166334	BONC	0.025417	-0.102281
	(0.058262)	(0.057036)		(0.025832)	(0.036941)
	(0.266922)	(0.256038)		(0.052639)	(0.066273)
MIN	-0.109935	0.033943	BON8K	-0.011998	-0.120708
	(0.018377)	(0.028566)		(0.042057)	(0.041129)
	(0.058432)	(0.052063)		(0.099445)	(0.074445)
RECR	0.519276	0.614787	Observations		2544
	(0.087506)	(0.167586)	Residual Sum of Squares		241.7
	(0.294993)	(0.492578)	Std Err of Regression		0.311

Under assumptions (A1), (A2), and (A3)', 2SLS is still consistent, but it is no longer asymptotically efficient in the class of instrumental variables estimators. The coefficient vector is unaffected by (A3)', but the new asymptotic covariance matrix becomes

$$\text{avar}(\delta_{2SLS}) = (X'P_Z X)^{-1} X'Z(Z'Z)^{-1} Z'\Omega Z(Z'Z)^{-1} Z'X(X'P_Z X)^{-1} \qquad (12)$$

The problem here is to estimate the matrix $Z'\Omega Z$. In a generalization of the White (1980) Conditional-Heteroskedasticity Consistent Covariance Matrix Estimator, Theorem 6.3 in White (1984) shows that the following is a consistent estimator for the (2sx2s) matrix $Z'\Omega Z$:

$$V = (1/n) \sum_{t=1}^{n} Z_t{}'\epsilon_t \epsilon_t{}'Z_t . \qquad (13)$$

Here Z_t is a (px2s) matrix, ϵ_t is a (px1) vector of fitted 2SLS residuals, s is the number of instrumental variables, and t=1,...,n indexes recruiting battalions.

The estimated asymptotic standard errors from equation (12) are shown as the second set of standard errors in Table 2. These are, as expected, uniformly larger than the incorrect 2SLS standard errors, so that the 2SLS standard errors lead one to believe that the coefficients are more precisely estimated than, in fact, they are.

Tests of the stability of the entire coefficient vector and the subvector of principal policy interest result in Chi-Squared test statistics of 166.46 (19 df) and 13.28 (6 df), respectively. The first rejects at any reasonable significance level. The second rejects at the 0.05 level, but not at the 0.01 level.[20]

Given the marginal nature of this last rejection, it would be useful to have some insight into what caused it. The Scheffé multiple comparison procedure, which is discussed in Savin (1980), allows the researcher to investigate which contrasts might be responsible for the rejection of a null hypothesis. For the unrestricted (38x1) 2SLS coefficient vector, $\boldsymbol{\delta}$, we are testing the hypothesis H_0: $\mathbf{R}\boldsymbol{\delta} - \mathbf{r} = \boldsymbol{\Theta} = \mathbf{0}$, where

$$R_{(6 \times 38)} = [j_2' \ \ j_6' \ \ j_7' \ \ j_{10}' \ \ j_{12}' \ \ j_{13}']', \tag{14}$$

$\mathbf{r} = \mathbf{0}$, and j_i is the (1x38) row vector with a 1 in the i^{th} position, and a -1 in the $(i+19)^{th}$ position. There are q=6 restrictions. Using the estimated coefficient, define $\mathbf{h} = \mathbf{R}\hat{\boldsymbol{\delta}} - \mathbf{r}$. Then the Wald test statistic is

$$W = \mathbf{h}'[R(VCOV)R']^{-1}\mathbf{h}, \tag{15}$$

where VCOV is the estimated asymptotic covariance matrix in equation (12).[21] The acceptance region of this test is

[20] The P-values for these tests are 1.54E-09 and 0.039.
[21] Multiplication by T, which is usually seen in the asymptotic form of the Wald test statistic, is implicit in the estimation of VCOV.

$W \leq$ Chi-Squared$_\alpha(q)$. For $\alpha = 0.05$ and $q = 6$, we accept the hypothesis if $W \leq 12.5916$.

Let L be the set of linear combinations $\emptyset = \mathbf{a}'\boldsymbol{\theta}$ for $\boldsymbol{\theta} \in \mathbf{R}^q$. $\emptyset$ is estimated as $\mathbf{a}'\mathbf{h}$, and the variance of $\emptyset$ as $\mathbf{a}'R(VCOV)R'\mathbf{a}$. Then the Scheffé theorem states that the probability is asymptotically $1-\alpha$ that simultaneously for all $\emptyset$ in L,

$$\mathbf{a}'\mathbf{h}-S(\mathbf{a}'R(VCOV)R'\mathbf{a})^{1/2} \leq \emptyset \leq \mathbf{a}'\mathbf{h}+S(\mathbf{a}'R(VCOV)R'\mathbf{a})^{1/2}, \quad (16)$$

where $S = (\text{Chi-Squared}_\alpha(q))^{1/2}$. That is, the Wald test accepts H_0 if and only if for all $\emptyset$ in L, the large sample Scheffé interval covers zero.

Savin (1980) shows that we can find the linear combination which is most likely to reject H_0 as

$$\mathbf{a}_0 = [R(VCOV)R']^{-1}(R\hat{\boldsymbol{\delta}}-\mathbf{r})/S. \quad (17)$$

This choice of $\mathbf{a}_0$ is normalized so that $\mathbf{a}_0'[R(VCOV)R']\mathbf{a}_0=1$. For the case of the hypothesis which we are investigating, we compute this vector to be

$$\mathbf{a}_0 = [-4.67 \quad 0.64 \quad 3.05 \quad -1.56 \quad 13.54 \quad 17.53]'. \quad (18)$$

A comparison of $\mathbf{a}_0$ and the results reported in Table 2 prompts the following conclusions.

(i). It is not the case that one must take some perverse, economically uninteresting linear combination of the parameter estimates in order to reject the hypothesis that the model is stable. a_0 does not give particularly heavy weight to any one of the parameter estimates.

(ii). Very little weight is given to RPAY in this linear combination which is most likely to reject the hypothesis. While the coefficient on RPAY drops drastically for the second period, its standard error is so large that the change does not seem significant. This highlights the imprecision of the RPAY estimate.

(iii). The relatively large weights given to LADV and NADV compared to that given to MGRAD are difficult to interpret based on what is reported in Table 2. Both MGRAD and NADV are fairly precisely measured, so that one would think that changes in these coefficients should be given considerable weight. This is true for NADV, but not for MGRAD. LADV is imprecisely measured, so that one would expect it to get less weight than it actually receives. The reason for this discrepancy must be hidden in the covariances.

To summarize, a Chow test of the hypothesis that the model is stable over the two periods rejects that hypothesis. The rejection occurs when the usual error structure is assumed as well as when a block diaginal error structure characteristic of panel data is assumed.[22]

---

[22]Further, this result will be unchanged when the assumption of serial independence is relaxed in the next section, because the matrix Ω will still be block diagonal, leaving the matrix $V = Z'\Omega Z$ unchanged within the sample.

V. THE PROBLEM OF SERIAL CORRELATION

The last section analyzed the stability of the DS forecasting model under two alternative error structures: the standard assumption for a simultaneous equations model, which is incorrect for this model, and the more likely block diagonal error structure characteristic of panel data. But neither of these is completely correct since both assume serial independence.

To see that there is some form of serial correlation present, and to gain further understanding of the true error structure, recall that the structural model of DS is a switching model of enlistment. When the number of high quality enlistments does not exceed the mission for such enlistments, an observation is considered to be supply constrained, and the economist observes the outcome of independent labor supply decisions. However, when the opposite is true, the observation is considered to be demand constrained, and the economist observes the result of an enlistment production process. DS found evidence for the hypothesis of Dertouzos (1983) that asymmetric incentives for recruiters result in delaying the enlistment of some high quality recruits until those enlistments can be applied against a new objective.

This institutional behavior in the demand constrained environment destroys the serial independence of observations within a recruiting battalion. Imposing a first order autoregressive (AR1) process on the errors, and assuming that it is the same process for every battalion, leads to an estimate for the autoregressive parameter, ρ. Stacking the vector of 2SLS residuals and deleting the first observation from each battalion, the residuals are regressed on their lagged values to obtain $\hat{\rho} = 0.5592$, with an estimated standard error of (0.0167).

But the form of the serial correlaton is likely to be more complicated than AR1. One must account for the switching nature of the structural model. As discussed earlier, enlistments are very seasonal, arguing for a possible AR12 specification. Finally, the structural model of DS leads one to expect serial dependence within a recruiting battalion, but not across battalions.

These considerations lead to a general specification of the error structure as

$$E(u_1 u_1') = \Omega = \mathrm{diag}(\Omega_t), \quad t=1,\ldots,n, \tag{19}$$

where Ω_t is not necessarily equal to Ω_τ for $t \neq \tau$. Each Ω_t is a full (pxp) matrix which allows for a random effects specification of

$$u_{tp} = \epsilon_{tp} + \epsilon_\alpha + \theta, \tag{20}$$

where ϵ_{tp} is a truly random component of the error, ϵ_α is a component specific to the recruiting battalion, and θ is a time component of the error term which follows some autoregressive process.

The problem of instrumental variable estimation under nonstandard error structures has been addressed by Chamberlain (1982), White (1982,1984), and Cragg (1983). Chamberlain and White independently introduced a minimum distance estimator which is more efficient than 2SLS, and which is referred to here as the Two-Stage Instrumental Variables estimator, or 2SIV.[23] To see why 2SIV is more efficient than 2SLS, recall that 2SLS may be interpreted as a minimum distance estimator (Hausman, 1983). 2SLS is the estimator which minimizes the sum of squared residuals after they have been projected into the space spanned by the instruments. Mathematically,

$$\boldsymbol{\delta}_{2SLS} = \operatorname{argmin}\ (\boldsymbol{y}_1 - X_1\boldsymbol{\delta})'Z(Z'Z)^{-1}Z'(\boldsymbol{y}_1 - X_1\boldsymbol{\delta}). \qquad (21)$$

The first order condition from minimizing (21) gives the familiar expression for the 2SLS estimator,

$$\boldsymbol{\delta}_{2SLS} = (X_1{}'Z(Z'Z)^{-1}Z'X_1)^{-1}X_1{}'Z(Z'Z)^{-1}Z'\boldsymbol{y}, \qquad (22)$$

and the asymptotic covariance matrix of $\boldsymbol{\delta}_{2SLS}$ is given in equation (12). When $\Omega = \sigma_{11}I_T$, as is usually assumed in a simultaneous equations model, avar($\boldsymbol{\delta}_{2SLS}$) reduces to equation (5).

However, when $\Omega \neq \sigma_{11}I_T$, we can find a more efficient estimator than 2SLS. This is done with the following minimization problem:

[23]This is White's name for the estimator. Chamberlain calls it Generalized 2SLS, but they are the same estimator.

$$\delta_{SIV} = \underset{\delta}{\text{Min}}\ (\mathbf{y}_1 - X_1\delta)'ZPZ'(\mathbf{y}_1 - X_1\delta), \tag{23}$$

where P is some positive definite matrix to be chosen to minimize the asymptotic covariance matrix of the estimator. The first order condition is

$$\delta_{2SIV} = (X_1'ZPZ'X_1)^{-1}X_1'ZPZ'\mathbf{y}, \text{ with} \tag{24}$$

$$\text{avar}(\delta_{2SIV}) = (X_1'ZPZ'X_1)^{-1}X_1'ZPZ'\Omega ZPZ'X_1(X_1'ZPZ'X_1)^{-1}. \tag{25}$$

Denoting $V = Z'\Omega Z$, and choosing the optimal $P = V^{-1}$, expressions (24) and (25) can be rewritten as

$$\delta_{2SIV} = (X_1'ZV^{-1}Z'X_1)^{-1}X_1'ZV^{-1}Z'\mathbf{y}, \text{ and} \tag{26}$$

$$\text{avar}(\delta_{2SIV}) = (X_1'ZV^{-1}Z'X_1)^{-1}. \tag{27}$$

White (1984) proves that for block diagonal Ω which satisfy some regularity conditions, V is consistently estimated by equation (13). Because the instrument list is not being split into two subsamples for a Chow test, V is an (sxs) matrix in this case, rather than (2sx2s). When the instrument list does not contain a lagged dependent variable, 2SIV is consistent and asymptotically efficient in the class of all instrumental variable estimators.

There is no gain in efficiency when the model is just identified, as both 2SLS and 2SIV reduce to ordinary instrumental variable estimation, as shown in White (1982). White (1984) shows that for serial correlation of some unknown form, it is possible to expand the list of valid instrumental variables by taking measureable functions of the existing instruments. In theory, there is no limit to the number of additional instrumental variables that can be generated in this manner, but one should be careful that the new instruments are, in fact, asymptotically uncorrelated with the error terms. Given a particular list of valid instrumental variables, there is only one optimal way to combine them. If the error structure obeys the usual assumptions, then 2SLS is optimal. Otherwise, 2SIV is optimal.

In order to overidentify the model, seven additional instruments are generated. These are the squares of the instruments MGRAD, MNHSG, MOTH, and DRECR, and the square roots of the variables MGRAD, MNHSG, and MOTH.[24] Before proceeding with 2SIV estimation, the validity of these additional instruments is tested. This is done using a form of the Hausman Specification Test proposed in Hausman and Taylor (1981).[25]

Sufficient assumptions are made to just-identify the

---

[24]Actually, eight additional instruments were generated, but the eighth, the square root of the variable DRECR, caused the (Z'Z) matrix to become singular, so it was not used.

[25]The literature on this test is long. First proposed in Hausman (1978), relevant review and extension articles are Hausman and Taylor (1980,1981), Holly (1982), and Ruud (1984).

equation, and 2SLS gives the just-identified estimate δ_J. All the instruments are then included, and an overidentified 2SLS estimate, δ_O, is obtained. The covariance matrices of these estimators are of the form of equation (12) since the error structure is block diagonal. Then, defining $\mathbf{q} = (\delta_J - \delta_O)$, the test statistic

$$H = \mathbf{q}'(\ \mathrm{VCOV}(\delta_J) - \mathrm{VCOV}(\delta_O)\ )^{+}\mathbf{q} \tag{28}$$

is asymptotically distributed as Chi-Squared(7), where "+" denotes any generalized inverse. Performing this test results in a test statistic of 1.09, which is well below the critical value of 14.06 for the 0.05 significance level. While the test accepts that the new instruments are valid, it is not a very powerful test in this case. The coefficients do change considerably, but VCOV($\mathbf{q}$) is quite large. Since the inverse of VCOV($\mathbf{q}$) appears in the noncentrality parameter for this test, the noncentrality parameter will be much smaller for any given alternative hypothesis, which reduces the power of the test.[26]

Having established the validity of these additional instruments, the 2SIV estimates are computed for the entire sample. Table 3 presents the results. In Column (1) is 2SIV on the just-identified model.[27] Column (2) is 2SIV on

---

[26] Additionally, the test does not have unit power asymptotically, since it will not reject if we have used an invalid instrument to just-identify the equation.
[27] 2SLS on the just-identified model produces numerically identical estimated coefficients and standard errors, as expected.

the model with the seven additional instruments. Column (3) is 2SLS including the seven additional instruments, and is included for comparison. The standard errors presented in Column (3) are from equation (12).

Examining Column (1) of Table 3 we find that the estimated coefficients are numerically identical to those in Column (3) of Table 1, but the standard errors of 2SIV are large relative to the coefficients, so that only MGRAD, NADV, and Q4 are significantly different from zero. The incorrectly-computed standard errors of a packaged 2SLS program lead one to believe that the coefficients are estimated more precisely than they actually are. Column (1) of Table 3 represents the best possible estimates of the forecasting equation in DS when the true error structure is utilized, and no additional instruments are generated. Comparing Column (2) with Column (3) reveals that 2SIV is, in fact, more efficient than 2SLS when the error structure is as specified in equation (19).

TABLE 3
COMPARISON OF 2SIV AND 2SLS RESULTS
DEPENDENT VARIABLE: ARMY HIGH QUALITY ENLISTMENTS

EXPLANATORY VARIABLE	(1)	(2)	(3)
CONSTANT	0.643801	0.055891	-0.378554
	(1.249460)	(0.640497)	(0.756861)
MGRAD	0.528874	0.373392	0.387923
	(0.107629)	(0.051638)	(0.061154)
CNHSG	0.146508	-0.026504	-0.008359
	(0.118281)	(0.033369)	(0.041609)
COTH	-0.127871	-0.042960	-0.094132
	(0.103187)	(0.054747)	(0.079368)
CDOD	-0.290989	0.596021	0.424463
	(0.582122)	(0.183955)	(0.246096)
RPAY	0.350580	-0.615162	-0.033321
	(1.016985)	(0.262420)	(0.610740)
UR	0.403671	0.022270	0.196534
	(0.403671)	(0.094334)	(0.190613)
QMA	-0.015259	-0.188349	-0.115926
	(0.187360)	(0.072886)	(0.107777)
MIN	-0.076149	-0.037511	-0.048833
	(0.048251)	(0.025023)	(0.027220)
RECR	0.500391	0.017260	0.195670
	(0.384859)	(0.133877)	(0.189784)
REXP	-0.073017	0.058894	0.010517
	(0.158061)	(0.049875)	(0.083510)
LADV	-0.013896	0.005530	0.001484
	(0.022444)	(0.017014)	(0.018332)
NADV	0.074324	0.085534	0.063018
	(0.032836)	(0.012577)	(0.026483)
VOTE	0.014436	-0.140814	-0.001974
	(0.327161)	(0.133871)	(0.195367)
Q2	0.038335	0.014987	0.038271
	(0.032599)	(0.017801)	(0.023439)
Q3	0.083755	0.030542	0.064322
	(0.060179)	(0.018780)	(0.034841)
Q4	0.227759	0.106145	0.152038
	(0.110170)	(0.030526)	(0.055295)
ACF	0.050919	0.217587	0.137404
	(0.143764)	(0.047192)	(0.093027)
NCVP	0.016121	-0.030731	0.000544
	(0.082316)	(0.053077)	(0.058652)
BILL	0.082756	0.085963	0.046201
	(0.117567)	(0.069974)	(0.074297)
BONC	0.024682	0.033017	0.052114
	(0.042445)	(0.026420)	(0.036741)
BON8K	0.025873	-0.008735	0.065687
	(0.074617)	(0.047987)	(0.065630)
BON84K	0.074571	0.079342	0.099063
	(0.052258)	(0.026454)	(0.037337)
Observations	2544	2544	2544
Std error of regr	0.294	0.231	0.231

(1)2SIV/2SLS on the just-identified model.
(2)2SIV on the overidentified model.
(3)2SLS on the overidentified model (for comparison only).

In Column (2) the forecasting specification is estimated using the additional instrumental variables generated as suggested by White (1984). The increase in the precision of the estimates is quite striking. The variables MGRAD, CDOD, RPAY, QMA, NADV, Q4, ACF, and BON84K are all significantly different from zero. Unfortunately, three important variables, CDOD, RPAY, and QMA have the incorrect sign, and UR is still insignificant.

The negative sign on RPAY is particularly troublesome.[28] There is simply no economic justification for this, although one can propose several sociological rationales. One might explain this fact as an aberration of the sample[29] in the following manner. The 1981 military pay raise of 12% was successful in beginning to draw increasing numbers of high quality recruits into the Army. The recession of 1982 -1983 reinforced this pattern. Younger high school students watched as older high school students enlisted, and this precedent legitimized the Army as an option among high school students whom the Army would consider to be high quality recruits. Then, although the relative military-to-civilian wage began to fall in 1982, high quality enlistments continued to rise.

This argument turns on the inability of young men to

---

[28] In fact, as soon as the model is overidentified with any one of the additional instruments, this coefficient becomes negative.

[29] It is surprising to find, as the correlation matrix in the Data Appendix shows, that RPAY is negatively correlated with most of the measures of enlistments and enlistment missions.

make marginal calculations. While this is probably true in a very strict sense, and while it is also true that the decision to enlist in the Army is not and should not be a decision which is motivated by purely economic considerations, at some point any economic agent will recognize a better economic alternative. This explanation highlights the fact that one cannot forecast reliably with a model which has a negative coefficient on RPAY.

The insignificance of the UR coefficient is equally disappointing. The affect of the macroeconomy on recruiting is a very important input into policy decisions. Forecasting with a model which has an insignificant coefficient on unemployment does not seem appropriate.

There is one more possible source of inconsistency in the model as it is estimated in Table 3. Serial correlation in a simultaneous equations model when a lagged dependent variable is included in the list of instrumental variables produces inconsistent parameter estimates. Fair (1970) provides a method of obtaining consistent estimates under the assumption that the serial correlation is AR1, and the dependent variable appears with a one period lag as a predetermined variable in the instrument list. As discussed earlier, the mission for high quality recruits, MGRAD, is heavily weighted toward past performance in high quality enlistments, CGRAD, so that it is possible to view MGRAD as a lagged dependent variable, but it is not reasonable to assume that a one period lag structure is correct.

A more reasonable interpretation, however, is that there are other inputs into the determination of the high quality enlistment mission, so that MGRAD is simultaneously determined with CGRAD. DS show that MGRAD definitely belongs in the equation for CGRAD, and this causes the forecasting model to be underidentified.

An identifying instrument can come from many sources. One of these is a covariance restriction, as proposed by Hausman and Taylor (1983). Let the equation determining MGRAD be the k^{th} equation in the system. Then

$$MGRAD = f(CGRAD_{-1}, \ldots) + u_k. \tag{29}$$

It is necessary to correctly specify the other equations in the system in order to determine relative recursivity. If there were only two equations in the system, then they would be relatively recursive since CGRAD and $CGRAD_{-1}$ are different variables and CGRAD does not appear in the MGRAD equation. Even then, it would be unreasonable to assume that $\sigma_{1k} = 0$, since MGRAD is a policy variable closely linked to CGRAD.

The most promising solution to this problem of underidentification is to specify some of the exogenous variables which might appear in the other equations and which are excluded from the CGRAD equation. This can be done without completely specifying the other equations. Then the model can be identified, or perhaps even over-identified. Unfortunately, these data are currently unavailable so that one cannot add this additional structure to the model.

VI. SUMMARY AND CONCLUSIONS

This paper has examined a model for forecasting Army high quality enlistments proposed by Daula and Smith (1985). The model was estimated over an extended period with additional data provided by the United States Military Academy. The stability of the model was tested by performing a Chow test for the equivalence of the coefficient vector over the two subsamples of data corresponding to the period originally examined by DS and the period corresponding to the additional data. This tests rejects the hypothesis that the entire coefficient vector is stable, and also rejects the hypothesis that the subvector of principal policy interest is stable.

The error structure assumed by DS in their estimation was also examined and determined to be incorrect. The forecasting model was reestimated using 2SIV and a general block diagonal error structure which allows for serial correlation within the recruiting battalions. The results do not coincide with economic theory in that several key variables, including relative military-to-civilian pay, have the incorrect sign. Another very important variable, unemployment, is measured without much precision.

The current sample period is characterized by falling relative pay and rising enlistments. An extended sample period may provide data which are more in keeping with the strong predictions of economic theory.

Further consideration of the effects of serial correlation revealed that there is reason to believe that MGRAD is jointly endogenous with CGRAD. This causes the model to be underidentified. More work needs to be done on the specification of the other equations in the system of simultaneous equations which describe military labor supply. Only when data are collected on variables excluded from the CGRAD equation can sufficient structure be imposed on this forecasting model to obtain consistent estimates.

BIBLIOGRAPHY

AMEMIYA, Takeshi.(1985). Advanced Econometrics. Cambridge: Harvard University Press.

ALTMAN, Stuart H.(1967). "Earnings, Unemployment, and the Supply of Enlisted Volunteers." Journal of Human Resources Winter,38-59.

ALTMAN, Stuart H., and Alan E. Fechter.(1967). "The Supply of Military Personnel in the Absence of a Draft." Papers and Proceedings of the American Economic Association 57,16-31.

ASH, Colin, Bernard Udis, and Robert F. McNown.(1983). "Enlistments in the All-Volunteer Force: A Military Personnel Supply Model and its Forecasts." American Economic Review 73,145-155.

BALDWIN, Robert H. and Thomas V. Daula.(1985). "Army Recruit Attrition and Force Manning Costs: Methodology and Analysis." in Ronald G. Ehrenberg ed. Research in Labor Economics. Vol 7. Greenwich, CT: Jai Press.

BROWN, Charles.(1985). "Military Enlistments: What Can We Learn from Geographic Variation?" American Economic Review 75,228-234.

CHAMBERLAIN, Gary.(1982). "Multivariate Regression Models for Panel Data." Journal of Econometrics 18,5-46.

CHAMBERLAIN, Gary.(1984). "Panel Data." Ch.22 in Handbook of Econometrics, VOL II, ed. Zvi Griliches and Michael Intriligator. New York: North Holland.

CHOW, Gregory C.(1960). "Tests of Equality Between Sets of Coefficients in Two Linear Regressions." Econometrica 28,591-605.

CRAGG, J.G.(1983). "More Efficient Estimation in the Presence of Heteroskedasticity of Unknown Form." Econometrica 51,751-763.

DALE, Charles and Curtis Gilroy.(1983). "The Effects of the Business Cycle on the Size and Composition of the U.S. Army." Atlantic Economic Journal 11,42-53.

______.(1985)."Enlistments in the All-Volunteer Force: Note." American Economic Review 75,547-551.

DAULA, Thomas V., and D. Alton Smith.(1985). "Estimating Enlistment Models for the U.S. Army." United States Military Academy mimeo.

DERTOUZOS, James N.(1983). "Recruiter Incentives and Enlistment Supply." Rand Working Draft 1896-MRAL.

EHRENBERG, Ronald G., and Robert S. Smith.(1985). Modern Labor Economics: Theory and Public Policy, 2d ed. Glenview,IL: Scott,Foresman, and Co.

FAIR, Ray C.(1970). "The Estimation of Simultaneous Equation Models With Lagged Endogenous Variables and First Order Serially Correlated Errors." Econometrica 38,507-516.

FERNANDEZ, Richard L.(1979). Forecasting Enlisted Supply: Projections for 1979-1990. Rand Report.

FISHER, Anthony C.(1969). "The Cost of the Draft and the Cost of Ending the Draft." American Economic Review 59,239-254.

GOLDBERG, Lawrence.(1982). Enlisted Supply: Past, Present, and Future. Center for Naval Analysis Report 1168.

GRILICHES, Zvi, and Jerry A. Hausman.(1986). "Errors in Variables in Panel Data." Journal of Econometrics 31,93-118.

HAUSMAN, Jerry A.(1977). "Errors in Variables in Simultaneous Equation Models." Journal of Econometrics 5,389-401.

HAUSMAN, Jerry A.(1978). "Specification Tests in Econometrics." Econometrica 46,1251-1271.

HAUSMAN, Jerry A.(1983). "Specification and Estimation of Simultaneous Equation Models." Ch.7 in Handbook of Econometrics, VOL I, ed. Zvi Griliches and Michael D. Intriligator. New York: North Holland.

HAUSMAN, Jerry A., and William E. Taylor.(1980). "Comparing Specification Tests and Classical Tests." MIT mimeo number 266.

HAUSMAN, Jerry A., and William E. Taylor.(1981). "A Generalized Specification Test." Economics Letters 8,239-245.

HAUSMAN, Jerry A. and William E. Taylor.(1983). "Identification in Linear Simultaneous Equations Models with Covariance Restrictions: An Instrumental Variable Interpretation." Econometrica 51,1527-1549.

HECKMAN, James J.(1979). "Sample Selection as a Specification Error." Econometrica 47,153-161.

HOLLY, Alberto.(1982). "A Remark on Hausman's Specification Test." Econometrica 50,749-759.

JUDGE, George G., et al.(1982). Introduction to the Theory and Practice of Econometrics. New York: John Wiley and Sons.

JUDGE, George G., et al.(1985). The Theory and Practice of Econometrics, 2d ed. New York: John Wiley and Sons.

LEE, Lung-Fei, et al.(1980). "Asymptotic Covariance Matrices of Two-Stage Probit and Two-Stage Tobit Methods for Simultaneous Equations Models with Selectivity." Econometrica 48,491-504.

MARSHALL, Alfred.(1952). Principles of Economics, 8th ed. London: Macmillan.

OI, Walter.(1967). "The Economic Cost of the Draft." Papers and Proceedings of the American Economic Association 57,39-62.

PARKS, Richard W.(1967). "Efficient Estimation of a System of Regression Equations When Disturbances Are Both Serially and Contemporaneously Correlated." Journal of the American Statistical Association 62,500-509.

RUUD, Paul A.(1984). "Tests of Specification in Econometrics." Econometric Reviews 3,211-242.

SAVIN, N. E.(1980). "The Bonferroni and the Scheffé Multiple Comparison Procedures." Review of Economic Studies 47,255-273.

THEIL, Henri.(1971). Principles of Econometrics. New York: John Wiley and Sons.

WEISBROD, B. A., and W. L. Hansen.(1967). "Economics of the Military Draft." Quarterly Journal of Economics 81,395-421.

WHITE, Halbert.(1980). "A Heteroskedasticity-Consistent Covariance Matrix Estimatator and a Direct Test for Heteroskedasticity." Econometrica 48,817-838.

WHITE, Halbert.(1982). "Instrumental Variables Regression with Independent Observations." Econometrica 50,483-499.

WHITE, Halbert.(1984). Asymptotic Theory for Econometricians. New York: Academic Press.

DATA APPENDIX

These data were provided by Dave Smith and the United States Military Academy. I have used the same data which Daula and Smith used in their study, plus the data compiled by Dave Smith for 15 additional months spanning July 1983 to September 1984. As explained in the section on the Stability of the Model, I have excluded the San Juan, PR, Beckley, WV, and the Nashville, TN recruiting battalions from the data set because some of the variables are missing or are not reported correctly. Therefore, fifty-three of the Army's fifty-six recruiting districts are included in the data. There are forty eight monthly observations beginning in October 1980 and ending in September 1984. This gives a total of 2544 observations.

The dependent variable is the number of male high school graduates and high school seniors who score in categories I-IIIA of the Armed Forces Qualtification Test (AFQT). These figures, as well as those of the other enlistment competition variables are from the Defense Manpower Data Center. They measure the number of enlistment contracts signed, rather than the number of accessions into the Army.

One problem encountered in constructing the data is that the boundries of recruiting districts do not match those of states or Standard Metropolitain Statistical Areas (SMSA's), which are the boundries usually used in reporting other data, particularly unemployment data, wages data, and demographic data. The method used to construct each variable is found in Daula and Smith (1985), and I do not repeat that here. In general, they disaggregate the original data to county level and then build recruiting districts by matching the counties to the district which contains it.

It is worthwhile to mention the variables and the reason for including them in the specification. The other jointly endogenous enlistment variables from the Army are: the number of male category I-IIIA non-high school graduates who enlist in the Army, and all other male Army enlistments. These two categories are included because Army recruiters are given enlistment goals for each of these categories, and time spent recruiting these categories of enlistments is time spent not recruiting high quality males. Female enlistments are excluded from the original data because they are considered always to be demand-constrained due to the legal restrictions on the number of positions in which they can serve. The last endogenous variable is the number of category I-IIIA high school graduates and seniors who enlist in the other services. This is included due to competition among the services for recruits from the same population of high quality young men.

The restrictions which permit the identification of the Army high quality enlistment equation are that the enlistment goals for the other two categories of Army recruits do not appear in the high quality enlistment equation, and that the number of other-service recruiters also does not appear. These variables are then used as instruments in the Army high quality enlistment equation. The enlistment goals for the other services might have been used to overidentify the Army high quality enlistment equation, but those data are not presently available.

As discussed earlier, the relative military pay variable is usually a source of inconsistency of parameter estimates in studies on enlistment. Included in the military pay variable is the value of pay and allowances that an individual can expect to receive during a three year enlistment. Average promotion rates are used and a weighted average of married and unmarried recruits is used, as some allowances are substantially greater for married soldiers. Finally, the average bonus for enlisting in shortage specialties is also included. This variable is measured in dollars per week for a three year enlistment, and is discounted to the present.

Measuring the civilian pay opportunities in a recruiting district is more problematical. Daula and Smith use average weekly earnings of production workers in manufacturing as reported in the Bureau of Labor Statistics (BLS) publication Employment and Earnings. This measures wages earned by workers of all ages in one of the most pecuniarily lucrative sectors of the economy, and so one expects that it overstates the wage opportunities available to recent high school graduates. To correct for this problem of measurement error, we instrument for the relative military pay variable using the Durbin instrument.

Unemployment data is also from the BLS. See Daula and Smith (1985), pp.22-23 for details on how this variable is constructed.

Three demographic variables are included. The first is the male seventeen to twenty-one year old population that scores above average on the AFQT and is physically qualified. The 1980 Census and the NLS Youth survey are used to construct this variable. The second variable is the percentage of minorities in the qualified military available population. The third is the percent Republican vote in the 1980 presidential election. This is a proxy for pro-military feeling in a particular recruiting district. Notice that these demographic variables vary in cross-section only; the panel is not long enough to experience any time series variation.

Recruiting effort is measured by four variables. The number of production recruiters and the percentage of those recruiters with at least nine months experience are included. The experience variable is included because it takes a new recruiter some time to learn the recruiting area. Advertising is measured as the expenditures on local advertising and by the number of impressions in the national electronic and print media.

Nine dummies are included in the regression to control for various effects. Dummies are included for the second, third, and fourth quarters of the fiscal year. One expects to find a larger number of enlistments in the fourth quarter which includes the summer months after graduation from high school. Dummies are also included to control for the type of educational benefit in force at the time of enlistment, and for experiments conducted by USAREC on enlistment bonuses.

Summary statistics and a correlation matrix for selected variables (in levels, not logs) is provided.

SUMMARY STATISTICS

	MEAN	STD DEV	MIN	MAX	VARIANCE
CGRAD	82.49	37.849	7.00	252.00	1432.58
MGRAD	72.34	34.814	2.00	198.00	1212.05
CNHSG	19.68	9.909	1.00	79.00	98.19
MNHSG	18.59	10.011	1.00	68.00	100.24
COTH	127.80	51.291	25.00	381.00	2630.86
MOTH	116.20	48.550	16.00	389.00	2357.13
CDOD	163.03	56.241	21.00	363.00	3163.12
DRECR	132.36	43.794	28.00	339.34	1917.98
RPAY	0.75	0.110	0.50	1.32	0.01
UR	8.50	2.428	3.20	19.27	5.89
RECR	88.81	26.729	32.00	173.00	714.44
NADV	1547.19	1190.141	97.55	9549.91	1416436.27
VOTE	52.04	5.529	41.18	68.68	30.57

CORRELATION MATRIX

	CGRAD	MGRAD	CNHSG	MNHSG	COTH
CGRAD	1.00000				
MGRAD	0.82425	1.00000			
CNHSG	0.43093	0.39158	1.00000		
MNHSG	0.34840	0.36610	0.60227	1.00000	
COTH	0.36341	0.24653	0.35147	0.23374	1.00000
MOTH	0.03940	0.05089	0.22939	0.22660	0.74458
CDOD	0.71677	0.63018	0.50257	0.40321	0.46039
DRECR	0.31802	0.31003	0.43961	0.32464	0.43967
RPAY	-0.18566	-0.18709	-0.06214	-0.06905	0.10189
UR	0.41927	0.28785	0.09399	0.03694	0.16772
RECR	0.54656	0.58080	0.42158	0.33955	0.42063
NADV	0.27724	0.12978	0.27135	0.16879	0.34674
VOTE	-0.12849	-0.14166	-0.00357	-0.01371	-0.22824

	MOTH	CDOD	DRECR	RPAY	UR
MOTH	1.00000				
CDOD	0.29560	1.00000			
DRECR	0.49815	0.62374	1.00000		
RPAY	0.07529	-0.19158	-0.12285	1.00000	
UR	0.01803	0.30586	0.11084	-0.25489	1.00000
RECR	0.41806	0.69156	0.75625	-0.35903	0.20490
NADV	0.28063	0.31623	0.36658	0.06489	0.29862
VOTE	-0.20562	-0.17938	-0.21077	-0.14683	-0.05865

	RECR	NADV	VOTE
RECR	1.00000		
NADV	0.27711	1.00000	
VOTE	-0.18613	-0.10046	1.00000

All estimation and computation was accomplished using the Gauss statistical package for personal computers (version 1.44).