S E A R C H*

A Linear Regression Computer Package

by

Edward E. Leamer

Discussion Paper Number 89

March 1977

*Seeking Extreme and Average Regression Coefficient Hypotheses

Preview

SEARCH is designed to search sensibly and exhaustively over the set of linear regressions that may be computed from a given data set. It is intended to replace both the expensive haphazard *ad hoc* searches and also the theoretically questionable and non-exhaustive computer assisted searches such as stepwise regression and principal component regression.

The statistical philosophy underlying SEARCH has a Bayesian flavor, but researchers are not expected to be able to specify completely their prior distribution.

A user's guide and object or source decks are available on request.

Reader's Guide

A reader can get a fairly clear idea of the performance of SEARCH by reading only Sections I and VII of this document.

Acknowledgements

This report describes a regression computer package, designed by Edward E. Leamer with the assistance of Herman Leonard. The planning and programming of this package have been partly supported by grants from the National Science Foundation, GS-31929 and SOC76-08863. The programming has been done by Kathy Burgoyne, Herman Leonard and Robert Topel.

TABLE OF CONTENTS

I. Introduction

This program pools prior information about a linear regression parameter vector $\underset{\sim}{\beta}$ with the information generated by the observations of a normal linear regression process

$$\underset{\sim}{Y} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{u}, \tag{1}$$

where $\underset{\sim}{Y}$ is a (T x 1) vector of observations, $\underset{\sim}{X}$ is a (T x k) design matrix, $\underset{\sim}{\beta}$ is a (k x 1) vector of parameters and $\underset{\sim}{u}$ is a (T x 1) vector of errors normally distributed with mean $\underset{\sim}{0}$ and variance $\sigma^2\underset{\sim}{I}$.

Prior information may be, but is not necessarily describable in terms of a completely specified prior probability distribution. Three options are provided.

<u>OPTION 1</u> The program will allow the prior distribution to be the Student function

$$f_s(\underset{\sim}{\beta}|\underset{\sim}{b}^*,\underset{\sim}{N}^*,\nu^*) \;\alpha\; [\nu^* + (\underset{\sim}{\beta}-\underset{\sim}{b}^*)'\underset{\sim}{N}^*(\underset{\sim}{\beta}-\underset{\sim}{b})]^{-(q+\nu^*)/2} \tag{2}$$

where q is the rank of the kxk positive semi-definite matrix $\underset{\sim}{N}^*$ and if q=k

$$E(\underset{\sim}{\beta}) = \underset{\sim}{b}^* \qquad \nu^* > 1$$

$$Var(\underset{\sim}{\beta}) = \underset{\sim}{N}^{*-1}\,\nu^*/(\nu^*-2) \qquad \nu^* > 2$$

For this option, summary measures of the posterior distribution are provided by the program. As of September, 1976, OPTION 1 is available only for the Normal special case $\nu^* = \infty$, and only if s^2 is an adequate estimate of σ^2.

<u>OPTION 2</u> The prior is known to be uniform on the ellipsoids

$$(\underset{\sim}{\beta}-\underset{\sim}{b}^*)'\underset{\sim}{N}^*(\underset{\sim}{\beta}-\underset{\sim}{b}^*) = c \tag{3}$$

where $\underset{\sim}{N}^*$ is a kxk symmetric positive semi-definite matrix. In this case, posterior modes necessarily lie on the locus of tangencies between the prior ellipsoids and the likelihood ellipsoids. This locus is called the <u>contract curve</u> and the program provides a numerical description of it.

<u>OPTION 3</u> The vector $\underset{\sim}{R}\underset{\sim}{\beta}$ is known to have mean $\underset{\sim}{r}$ but no covariance matrix is available. Equivalently, the prior is known to be uniform on the ellipsoids

$$(\underset{\sim}{R}\underset{\sim}{\beta}-\underset{\sim}{r})'\underset{\sim}{V}^{*-1}(\underset{\sim}{R}\underset{\sim}{\beta}-\underset{\sim}{r}) = c \tag{4}$$

where $\underset{\sim}{R}(q\text{x}k)$ and $\underset{\sim}{r}(q\text{x}1)$ are given, but $\underset{\sim}{V}^{*-1}$ is any $(q\text{x}q)$ symmetric positive semi-definite matrix. In this case the posterior modes lie in a region which could be described as the union of contract curves generated by all $\underset{\sim}{V}^{*-1}$. The hull of this region is the set of constrained least-squares points formed by imposing the constraints $\underset{\sim}{G}\underset{\sim}{R}\underset{\sim}{\beta} = \underset{\sim}{G}\underset{\sim}{r}$ where $\underset{\sim}{G}$ is any $(r\text{x}q)$ matrix. This hull of estimates is described numerically by the program in terms of the vectors at which (a) β_i is maximized (b) β_i is minimized (c) $\underset{\sim}{\psi}'\underset{\sim}{\beta}$ is maximized (d) $\underset{\sim}{\psi}'\underset{\sim}{\beta}$ is minimized. In addition, for each of the coefficients β_i, the matrix $\underset{\sim}{G}\underset{\sim}{R}$ is restricted to have zeroes in the i<u>th</u> row -- the constraints do not involve β_i -- and $\underset{\sim}{G}$ is otherwise chosen to maximize β_i and to minimize β_i. The program provides this pair of extreme estimates also.

All three options are automatically available. The input must be sufficient to specify the Student prior used for OPTION 1. (As of September, 1976, $\nu^* = \infty$.) This prior is then broken down to form the inputs necessary for OPTION 2 and OPTION 3. The quadratic form (3) needed for OPTION 2 is taken from the Student distribution (2). For OPTION 3 this same quadratic form is rewritten in terms of (4) with $\underset{\sim}{V}^{*-1}$ a qxq invertible matrix.

II. Definitions

The data evidence is fully described by the sufficient statistics

T, the number of observations

k, the number of coefficient parameters

$\underset{\sim}{b}$, a (kx1) solution to the normal equations $\underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{b} = \underset{\sim}{X}'\underset{\sim}{Y}$.

s^2, an estimate of σ^2; $s^2 = (\underset{\sim}{Y}-\underset{\sim}{X}\underset{\sim}{b})'(\underset{\sim}{Y}-\underset{\sim}{X}\underset{\sim}{b})/(T-k^*)$, where k^* is the rank of $\underset{\sim}{X}'\underset{\sim}{X}$ and $s^2 \equiv 1$ if $k^* \geq T$.

$k^* = \text{rank } (\underset{\sim}{X}'\underset{\sim}{X})$

$\underset{\sim}{H}$, the precision matrix $\underset{\sim}{H} = s^{-2}\underset{\sim}{X}'\underset{\sim}{X}$.

The Student prior (2) is fully described by the parameters

$\underset{\sim}{b}^*$, the kx1 prior location

$\underset{\sim}{N}^*$, the kxk prior precision matrix

ν^*, the prior degrees of freedom, set to infinity in the September, 1976 version.

q, the rank of N^*

The likelihood function can be written as

$$L(\underset{\sim}{\beta},\sigma^2;\underset{\sim}{Y},\underset{\sim}{X}) \propto (\sigma^2)^{-T/2} \exp(-[(\underset{\sim}{Y}-\underset{\sim}{X}\underset{\sim}{b})'(\underset{\sim}{Y}-\underset{\sim}{X}\underset{\sim}{b}) + (\underset{\sim}{\beta}-\underset{\sim}{b})'\underset{\sim}{X}'\underset{\sim}{X}(\underset{\sim}{\beta}-\underset{\sim}{b})]/2\sigma^2) \tag{5}$$

For a given σ^2, the likelihood function is constant on the ellipsoids

$$(\underset{\sim}{\beta}-\underset{\sim}{b})'\underset{\sim}{X}'\underset{\sim}{X}(\underset{\sim}{\beta}-\underset{\sim}{b}) = c^2$$

The concentrated likelihood function makes use of the maximum likelihood estimate of σ^2

$$\hat{\sigma}^2(\underset{\sim}{\beta}) = (\underset{\sim}{Y}-\underset{\sim}{X}\underset{\sim}{\beta})'(\underset{\sim}{Y}-\underset{\sim}{X}\underset{\sim}{\beta})/T$$

to form

$$
\begin{aligned}
L^{m}(\beta;Y,X) &= L(\beta,\hat{\sigma}^2(\beta);Y,X) \\
&\propto [(Y-X\beta)'(Y-X\beta)]^{-T/2} \\
&= [(Y-Xb)'(Y-Xb) + (\beta-b)'X'X(\beta-b)]^{-T/2}
\end{aligned}
\tag{6}
$$

The <u>conditional likelihood function</u> makes use of the unbiased estimate of σ^2, $\sigma^2 = s^2$, to form

$$
\begin{aligned}
L^{c}(\beta;Y,X) &= L(\beta,s^2;Y,X) \\
&\propto \exp[-(\beta-b)'H(\beta-b)/2]
\end{aligned}
\tag{7}
$$

where $H = s^{-2}X'X$. An alternative is to use the maximum likelihood estimate $\hat{\sigma}^2(b)$, but s^2 is thought more accurately to reflect the variance of the sample information.

The <u>contract curve</u> is the locus of tangencies between the likelihood ellipsoids and the prior ellipsoids

$$
b^{**}(\rho) = (\rho N^{*}+H)^{-1}(\rho N^{*}b^{*}+Hb), \qquad 0 \leq \rho \leq \infty. \tag{8}
$$

III. Input Options

IIIa. Data Input Options

The data may be input directly; alternatively the program will read the data moments or the output of a regression package:

OPTION 1 Input: T, k, k^*, $\underset{\sim}{Y}$, $\underset{\sim}{X}$

OPTION 2 Input: T, k, k^*, $\underset{\sim}{X}'\underset{\sim}{X}$, $\underset{\sim}{X}'\underset{\sim}{Y}$, $\underset{\sim}{Y}'\underset{\sim}{Y}$

OPTION 3 Input: T, k, k^*, s^2, $s^2(\underset{\sim}{X}'\underset{\sim}{X})^{-1}$, $\underset{\sim}{b}$

The third option is not recommended for problems with high dimensionality (k), since it "doubles" the errors in inverting $\underset{\sim}{X}'\underset{\sim}{X}$.

IIIb. Prior Input Options

The prior is described in terms of the following parameters

ν^* = prior "degrees of freedom" [set to ∞]

$\underset{\sim}{b}^*$ = (kxl) prior location vector

$\underset{\sim}{N}^*$ = (kxk) prior precision matrix, symmetric positive semi-definite

q = rank of $\underset{\sim}{N}^*$ (dimensionality of prior information)

$\underset{\sim}{R}$ = (qxk) matrix of constraints implicit in prior

$\underset{\sim}{r}$ = (qxl) location of $\underset{\sim}{R}\underset{\sim}{\beta}$

$\underset{\sim}{V}^*$ = (qxq) symmetric positive definite matrix, the covariance of $\underset{\sim}{R}\underset{\sim}{\beta}$.

Three input options are available:

OPTION 1 INPUT: $\underset{\sim}{N}^*$, $\underset{\sim}{b}^*$, $\underset{\sim}{q}$.

From this input, the program will compute $\underset{\sim}{R}$, $\underset{\sim}{r}$, $\underset{\sim}{V}$ and for its own use $\underset{\sim}{N}^*\underset{\sim}{b}^*$.

(a) If $q = k$, $\underset{\sim}{R} = \underset{\sim}{I}_k$, $\underset{\sim}{r} = \underset{\sim}{b}^*$, $\underset{\sim}{V}^* = (\underset{\sim}{N}^*)^{-1}$.

(b) If $q < k$, it is desired to find a qxq invertible matrix $\underset{\sim}{V}^*$ and a qxk matrix $\underset{\sim}{R}$ such that $\underset{\sim}{N}^* = \underset{\sim}{R}'\underset{\sim}{V}^{*-1}\underset{\sim}{R}$. The program first finds a kxk invertible matrix $\underset{\sim}{C}$ such that $\underset{\sim}{C}'\underset{\sim}{N}^*\underset{\sim}{C} = \underset{\sim}{\Lambda} = \text{diag}\ \{\lambda_1, \lambda_2, \ldots, \lambda_q, 0, \ldots, 0\}$. Then making use of $\underset{\sim}{N}^* = \underset{\sim}{C}'^{-1}\underset{\sim}{\Lambda}\underset{\sim}{C}^{-1} = \underset{\sim}{R}'\underset{\sim}{V}^{*-1}\underset{\sim}{R}$, $\underset{\sim}{R}$ is chosen to be the first q rows of C^{-1}, $\underset{\sim}{V}^* = \text{diag}\ \{\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_q^{-1}\}$ and $\underset{\sim}{r} = \underset{\sim}{R}\underset{\sim}{b}^*$.

<u>OPTION 2</u> INPUT $\underset{\sim}{R}$, $\underset{\sim}{r}$, $\underset{\sim}{V}^*$, $\underset{\sim}{q}$

From this input the routine computes $\underset{\sim}{N}^* = \underset{\sim}{R}'\underset{\sim}{V}^{*-1}\underset{\sim}{R}$, $\underset{\sim}{N}^*\underset{\sim}{b}^* = \underset{\sim}{R}'\underset{\sim}{V}^{*-1}\underset{\sim}{r}$ and $\underset{\sim}{b}^*$ which is any solution to $\underset{\sim}{N}^*\underset{\sim}{b}^* = \underset{\sim}{R}'\underset{\sim}{V}^{*-1}\underset{\sim}{r}$. The particular solution $\underset{\sim}{b}^*$ is found implicitly by minimizing the data quadratic form $(\underset{\sim}{b}^*-\underset{\sim}{b})'\underset{\sim}{X}'\underset{\sim}{X}(\underset{\sim}{b}^*-\underset{\sim}{b})$ subject to the constraint $\underset{\sim}{N}^*\underset{\sim}{b}^* = \underset{\sim}{R}'\underset{\sim}{V}^{*-1}\underset{\sim}{r}$. Equivalently this value of $\underset{\sim}{b}^*$ is the limit of $\underset{\sim}{b}^{**}(\rho)$ as ρ goes to infinity.

<u>OPTION 3</u> INPUT q

This is the default option which sets $\underset{\sim}{V} = \underset{\sim}{I}_q$, $\underset{\sim}{r} = \underset{\sim}{0}$, and $\underset{\sim}{R} = (\underset{\sim}{I}_q, \underset{\sim}{0})$ and then calls option 2.

IV. Numerical Description of the Contract Curve

Posterior modes necessarily lie on a curve that is the locus of tangencies between the likelihood ellipsoids $(\beta-b)'H(\beta-b)$ and the prior ellipsoids $(\beta-b^*)'N^*(\beta-b^*)$:

$$b^{**}(\rho) = (H+\rho N^*)^{-1}(Hb+\rho N^* b^*), \quad 0 \leq \rho \leq \infty .$$

This contract curve is described by the computer program in terms of a set of points on the curve, and also in terms of a set of $k+1 - (k-q) - (k-k^*)$ "ideal" points. The ideal points do not lie on the curve but their convex hull does completely contain the curve. Furthermore, each ideal point is a weighted average of constrained estimates.

IVa Rotation Invariant Average Regressions (Ideal Points)

The ideal points discussed in Leamer and Chamberlain (1976) are formed by writing the contract curve (8) as a ratio of polynomials in ρ:

$$b^{**}(\rho) = \Sigma_{j=o}^{p} w_j \rho^j a_j / \Sigma_{j=o}^{p} w_j \rho^j, \qquad w_j \geq 0, \tag{9}$$

$p = k - (k-q) - (k-k^*)$. In words, the curve is a weighted average of the $p+1$ "ideal" points, a_j, $j=0,\ldots,p$. Each ideal point is a weighted average of constrained least squares points. If $X'X$ is invertible, and if N^* is a diagonal matrix with rank k, then there are k+1 ideal points. The jth ideal point is a weighted average of the $\binom{k}{j}$ regressions formed by omitting j variables from the regression. The ideal points are invariant to rotations of the parameter space. (See Leamer and Chamberlain (1976)).

The ideal points are computed in the following way. An invertible matrix C is found such that $C'N^*C = \Delta = \text{diag}\ \{\delta_1,\delta_2,\ldots,\delta_k\}$ and $C'HC = \Lambda = \text{diag}\ \{\lambda_1,\lambda_2,\ldots,\lambda_k\}$. The contract curve $b^{**}(\rho)$ is written

$$\begin{aligned} b^{**}(\rho) &= [C'^{-1}(\rho C'N^*C + C'HC)C^{-1}]^{-1}[\rho N^* b^* + Hb] \\ &= C\, c^{**}(\rho) \end{aligned}$$

where $\underset{\sim}{c}^{**}(\rho) = (\rho\underset{\sim}{\Delta} + \underset{\sim}{\Lambda})^{-1}(\rho\underset{\sim}{C}'\underset{\sim}{N}^{*}\underset{\sim}{b}^{*} + \underset{\sim}{C}'\underset{\sim}{H}\underset{\sim}{b})$

$$= (\rho\underset{\sim}{\Delta} + \underset{\sim}{\Lambda})^{-1}(\rho\underset{\sim}{z}^{*} + \underset{\sim}{z})$$

The program collects powers of ρ by forming the polynomials

$$P^{0}(\rho) = \Pi_{j=1}^{k}(\rho\delta_i + \lambda_i) = \Sigma_{j=o}^{k}\, q_j^{o}\, \rho^{j}$$

$$P^{i}(\rho) = P^{o}(\rho)/(\delta_i + \lambda_i) = \Sigma_{j=o}^{k-1}\, q_j^{i}\, \rho^{j}$$

An element of $\underset{\sim}{c}^{**}(\rho)$ can be written as

$$c_i(\rho) = [P^{o}(\rho)]^{-1}[P^{i}(\rho)][\rho z_i + z]$$

$$= [P^{o}(\rho)]^{-1}\, \Pi^{i}(\rho)$$

where

$$\Pi^{i}(\rho) = P^{i}(\rho)[\rho z_i^{*} + z_i] = \Sigma_{j=o}^{k}\, r_j^{i}\, \rho^{j}$$

$$= \Sigma_{j=1}^{k}\, q_j^{o}\, (r_j^{i}/q_j^{o})\rho^{j}.$$

This allows us finally to write

$$\underset{\sim}{b}^{**}(\rho) = \underset{\sim}{C}\,\underset{\sim}{c}^{**}(\rho) = \Sigma_{j=1}^{k}\, q_j^{o}\, \rho^{j}\, \underset{\sim}{a}_j \,/\, \Sigma_{j=1}^{k}\, q_j^{o}\, \rho^{j}$$

where

$$\underset{\sim}{a}_j = \underset{\sim}{C}\underset{\sim}{r}_j$$

and

$$r_j = \{r_j^{i}/q_j^{o}\}.$$

IVb <u>Points on the Contract Curve</u>

The contract curve is traced in both directions, from the prior point to the data point, in steps of equal likelihood relative to the metric of the ending point. Ten points are presented for each trace, and various statistics describing the points are computed. The process of computing these points will be discussed with reference to the trace in the data metric from the prior point to the data point.

The relative (conditional) likelihood of the prior point to the data point is

$$L^c(b^*;Y,X) = \exp[-(b^*-b)'H(b^*-b)/2]$$

where $H = X'X\, s^{-2}$. We wish to find 10 points on the contract curve "equally spaced" in terms of L in the interval $[L^c(b^*), 1]$. The ith point on the curve b_i^{**}, is selected to have likelihood $L^c(b_i^{**}) = L^c(b^*) + j^*.1^*[1-L^c(b^*)]$. Thus b_i^{**} satisfies the equation

$$\ell_i = -2\log L(b_i^{**}) = (b_i^{**}-b)'H(b_i^{**}-b).$$

At the same time b_i^{**} must lie on the contract curve (9):

$$[b^{**}(\rho_i)-b] = \Sigma w_j\, \rho_i^j(a_j-b)/\Sigma w_j\, \rho_i^j.$$

Substituting this expression into the preceding equation we obtain the polynomial in ρ_i:

$$\ell_i\, \Sigma_j\, \Sigma_{j'}\, w_j w_{j'}\, \rho_i^{j+j'} = \Sigma_j\, \Sigma_{j'}\, w_j w_{j'}\, \rho_i^{j+j'}(a_j-b)'H(a_{j'}-b). \quad (10)$$

The (only positive?) root of this polynomial, ρ_i^+ implies the ith point on the contract curve

$$b_i^{**} = b^{**}(\rho_i^+)$$

Statistics describing the points on the contract curve.

For each point β on each trace of the contract curve, a number of quantities are reported. These quantities indicate the distance from β to b^* and from β to b. A quadratic form measuring distance between x and y is

$$Q(x,y,A) = (x-y)'A(x-y).$$

The reported quantities are:

a) The Data Likelihood (relative to likelihood at b):

$$d_1 = \exp[-Q(\beta,b,H)/2].$$

b) The Prior Density (relative to the density at b^*):

$$d_2 = \exp[-Q(\beta,b^*,N^*)/2].$$

c) The Data Distance (relative to the distance from $\underset{\sim}{b}$ to $\underset{\sim}{b}^*$):

$$d_3 = Q(\underset{\sim}{\beta},\underset{\sim}{b},\underset{\sim}{H})/Q(\underset{\sim}{b}^*,\underset{\sim}{b},\underset{\sim}{H}).$$

This measure lies between zero and one. If it is close to zero it indicates that $\underset{\sim}{\beta}$ is much closer to $\underset{\sim}{b}$ than is $\underset{\sim}{b}^*$ (in the data metric).

d) The Prior Distance (relative to the distance from $\underset{\sim}{b}$ to $\underset{\sim}{b}^*$):

$$d_4 = Q(\underset{\sim}{\beta},\underset{\sim}{b}^*,\underset{\sim}{N}^*)/Q(\underset{\sim}{b},\underset{\sim}{b}^*,\underset{\sim}{N}^*).$$

(See comment for d_3.)

e) The Euclidean Distance to the Data Point (relative to the Euclidean distance from $\underset{\sim}{b}$ to $\underset{\sim}{b}^*$):

$$d_5 = Q(\underset{\sim}{\beta},\underset{\sim}{b},\underset{\sim}{I})/Q(\underset{\sim}{b},\underset{\sim}{b}^*,\underset{\sim}{I}).$$

The Euclidean measures are intended to reflect visual impressions of closeness.

f) The Euclidean Distance to the Prior Point (relative to the Euclidean distance from $\underset{\sim}{b}$ to $\underset{\sim}{b}^*$):

$$d_6 = Q(\underset{\sim}{\beta},\underset{\sim}{b}^*,\underset{\sim}{I})/Q(\underset{\sim}{b},\underset{\sim}{b}^*,\underset{\sim}{I})$$

g) RHO, the solution to equation (10) for this point on the contract curve. If it were known that $\sigma^2 = s^2$, and if the prior were normal with mean $\underset{\sim}{b}^*$ and known precision matrix $\underset{\sim}{N}^*$, then the posterior mean would occur at $\rho = 1$. Thus Rho is a description of the posterior distribution. If Rho exceeds one, the prior needs to be more precise or the data need to be less precise, in order for the given point to be the posterior location.

h) SIGMA1, the standard deviation factor corresponding to the point $\underset{\sim}{\beta}$. The given point is a posterior mode, if the prior covariance were known to be $(\text{SIGMA1})^2\underset{\sim}{N}^{*-1}$. SIGMA1 and RHO are related as follows

$$(\text{SIGMA1})^2 = (\underset{\sim}{Y}-\underset{\sim}{X}\underset{\sim}{\beta})'(\underset{\sim}{Y}-\underset{\sim}{X}\underset{\sim}{\beta})/(Ts^2(\text{RHO}))$$

i) FDATA, the F statistic for testing the coefficient vector β.

$$F = Q(\utilde{\beta},\utilde{b},\utilde{H})/k^*$$

j) CHIPRIOR, the prior chi-square

$$Q(\utilde{\beta},\utilde{b}^*,\utilde{N}^*)$$

A priori this quantity has a chi-square distribution with q degrees of freedom.

V. Bounds Over Constrained Least Squares Points

The prior has implicit in it a set of q uncertain constraints, $\underset{\sim}{R}\underset{\sim}{\beta} = \underset{\sim}{r}$. The bounds described in this section make use of priors in which some of these constraints (or linear combinations thereof) are known exactly, but otherwise the prior is diffuse. These estimates bound the posterior location given that the prior mean of $\underset{\sim}{R}\underset{\sim}{\beta}$ is $\underset{\sim}{r}$ but the covariance matrix of $\underset{\sim}{R}\underset{\sim}{\beta}$ is free. The posterior location given $\mathrm{Var}(\underset{\sim}{R}\underset{\sim}{\beta}) = \underset{\sim}{V}^*$, is

$$\underset{\sim}{b}^{**}(\underset{\sim}{V}^*) = (\underset{\sim}{H} + \underset{\sim}{R}'(\underset{\sim}{V}^*)^{-1}\underset{\sim}{R})^{-1}(\underset{\sim}{H}\underset{\sim}{b} + \underset{\sim}{R}'(\underset{\sim}{V}^*)^{-1}\underset{\sim}{r}).$$

For the linear combination of coefficients $\underset{\sim}{\psi}'\underset{\sim}{\beta}$, the program solves for the extreme values of $\underset{\sim}{b}^{**}$, the solutions to $\max_{\underset{\sim}{V}^*} \underset{\sim}{\psi}'\underset{\sim}{b}^{**}(\underset{\sim}{V}^*)$ and $\min_{\underset{\sim}{V}^*} \underset{\sim}{\psi}'\underset{\sim}{b}^{**}(\underset{\sim}{V}^*)$. This can also be described in terms of constrained regressions. Let $\hat{\underset{\sim}{\beta}}(\underset{\sim}{R},\underset{\sim}{r})$ be the least-squares estimate of $\underset{\sim}{\beta}$ given $\underset{\sim}{R}\underset{\sim}{\beta} = \underset{\sim}{r}$. The program computes the constrained estimates $\hat{\underset{\sim}{\beta}}$ which are solutions to

$$\max_{\underset{\sim}{M}} \underset{\sim}{\psi}'\underset{\sim}{\beta}(\underset{\sim}{M}\underset{\sim}{R},\underset{\sim}{M}\underset{\sim}{r}) \text{ and } \min_{\underset{\sim}{M}} \underset{\sim}{\psi}'\underset{\sim}{\beta}(\underset{\sim}{M}\underset{\sim}{R},\underset{\sim}{M}\underset{\sim}{r})$$

These bounds for $\underset{\sim}{\psi}'\underset{\sim}{\beta}$ possibly make use of prior information about $\underset{\sim}{\psi}'\underset{\sim}{\beta}$. In some cases, it may be interesting to compute a bound that does not make use of constraints involving $\underset{\sim}{\psi}'\underset{\sim}{\beta}$. Such a bound can be computed as above with $\underset{\sim}{R}$ and $\underset{\sim}{r}$ replaced by $\underset{\sim}{R}\underset{\sim}{M}_\psi$ and $\underset{\sim}{r}\underset{\sim}{M}_\psi$ where $\underset{\sim}{M}_\psi = \underset{\sim}{I} - \underset{\sim}{\psi}(\underset{\sim}{\psi}'\underset{\sim}{\psi})^{-1}\underset{\sim}{\psi}'$. The matrix $\underset{\sim}{R}\underset{\sim}{M}_\psi$ has rank q or q-1. Its rows are constructed to be orthogonal to $\underset{\sim}{\psi}$, but otherwise the rows of $\underset{\sim}{R}\underset{\sim}{M}_\psi$ span the same space as the rows of $\underset{\sim}{R}$. Notice that constraints of the form $\underset{\sim}{M}\underset{\sim}{R}\underset{\sim}{M}_\psi\, \underset{\sim}{\beta} = \underset{\sim}{M}\underset{\sim}{r}\underset{\sim}{M}_\psi$ cannot constrain $\underset{\sim}{\psi}'\underset{\sim}{\beta}$. Suppose there were a(1xq) $\underset{\sim}{M}$ such that $\underset{\sim}{M}\underset{\sim}{R}\underset{\sim}{M}_\psi\, \underset{\sim}{\beta} = \psi'\beta$; then $MRM_\psi = \psi'$; postmultiply by $\underset{\sim}{\psi}$ to obtain $0 = \underset{\sim}{M}\underset{\sim}{R}\underset{\sim}{M}_\psi\, \underset{\sim}{\psi} = \underset{\sim}{\psi}'\underset{\sim}{\psi}$, a contradiction.

These bounds require the vectors $\underset{\sim}{\psi}$ to be input except when $\underset{\sim}{\psi}$ is a coordinate vector, in which case the bounds are activated by a special command.

A two dimensional example is illustrated in Figure 2 on page 24. An ellipse of constrained estimates is generated by rotating a line through the origin and finding the tangency between the line and a likelihood ellipse. The coefficient β_1 is maximized at the point A. The constraint that implies this estimate is the ray through A and the origin. The point H represents an extreme value of some linear combination of the coefficients. The point H is computed either without constraint or subject to being on the ray through H. The point K is computed subject to $\beta_1 = 0$ <u>and</u> $\beta_2 = 0$, or it is computed given a suitable constraint ray out of the origin.

The mathematics that implies the extreme estimates will now be discussed. The constrained least-squares estimator subject to $\underset{\sim}{R}\underset{\sim}{\beta} = \underset{\sim}{r}$ with $\underset{\sim}{V} = \sigma^2(\underset{\sim}{X}'\underset{\sim}{X})^{-1}$ and $\underset{\sim}{b} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{Y}$ is

$$\hat{\beta} = \underset{\sim}{b} - \underset{\sim}{V}R'(\underset{\sim}{R}\underset{\sim}{V}\underset{\sim}{R}')^{-1}(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r}) \tag{11}$$

with variance

$$V(\underset{\sim}{\hat{\beta}}|\underset{\sim}{R}\underset{\sim}{\beta}=\underset{\sim}{r}) = \underset{\sim}{V} - \underset{\sim}{V}\ R'(\underset{\sim}{R}\underset{\sim}{V}\underset{\sim}{R}')^{-1}\underset{\sim}{R}\underset{\sim}{V} \tag{12}$$

A notational convenience is to define

$$\underset{\sim}{A} = \underset{\sim}{R}\underset{\sim}{V}\underset{\sim}{R}' \tag{13}$$

The following theorem is used to find the bounds.

<u>Theorem 1</u> Least-squares subject to the constraint $\underset{\sim}{M}\underset{\sim}{R}\underset{\sim}{\beta} = \underset{\sim}{M}\underset{\sim}{r}$ satisfies the ellipsoid equality

$$(\underset{\sim}{R}\underset{\sim}{\beta}-\underset{\sim}{f})'\underset{\sim}{A}^{-1}(\underset{\sim}{R}\underset{\sim}{\beta}-\underset{\sim}{f}) = (\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})'\ \underset{\sim}{A}^{-1}(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})/4 \tag{14}$$

where $\underset{\sim}{f} = (\underset{\sim}{R}\underset{\sim}{b}+\underset{\sim}{r})/2$. Conversely, for any $\underset{\sim}{\gamma} = \underset{\sim}{R}\underset{\sim}{\beta}$ on the ellipsoid, there is a constrained estimate $\underset{\sim}{\hat{\beta}}$ such that $\underset{\sim}{\gamma} = \underset{\sim}{R}\underset{\sim}{\hat{\beta}}$. The estimate is computed subject to the constraint $\underset{\sim}{M}\underset{\sim}{R}\underset{\sim}{\beta} = \underset{\sim}{M}\underset{\sim}{r}$, with

$$M = (Rb-\gamma)'A^{-1}$$

Proof. The constrained estimate is

$$\hat{\beta}(MR,Mr) = b - V\, R'M'(MAM')^{-1}\,(MRb-Mr).$$

Then $A^{-1}R(\hat{\beta}-f) = A^{-1}Rb/2 - M'(MAM')^{-1}(MRb-Mr) - A^{-1}r/2$

$$= [A^{-1} - 2M'(MAM')^{-1}M]\,(Rb-r)/2,$$

and $(R\hat{\beta}-f)'A^{-1}(R\hat{\beta}-f) = (R\hat{\beta}-f')A^{-1}\,A\,A^{-1}(R\hat{\beta}-f)$

$$= (Rb-r)'[A^{-1} - 4M'(MAM')^{-1}M$$

$$+ 4M'(MAM)^{-1}M](Rb-r)/4$$

$$= (Rb-r)'\,A^{-1}(Rb-r)/4$$

The converse of the theorem makes use of $M = (Rb-\gamma)'\,A^{-1}$ to determine the constrained estimate

$$\hat{\beta}(MR,Mr) = b - V\,R'A^{-1}(Rb-\gamma)$$

$$[(Rb-\gamma)'A^{-1}\,A\,A^{-1}\,(Rb-\gamma)]^{-1}$$

$$[(Rb-\gamma)\,R'\,A^{-1}\,(Rb-r)].$$

But the ellipsoid (14), $(\gamma-f)'\,A^{-1}(\gamma-f) = (Rb-r)\,A^{-1}(Rb-r)/4$, can be rewritten as

$$(Rb-\gamma)'\,A^{-1}(Rb-\gamma) = (Rb-\gamma)'\,A^{-1}(Rb-r),$$

and thus $\hat{\beta} = b - V\,R'\,A^{-1}(Rb-\gamma)$, which is the constrained estimate subject to $R\beta = \gamma$.

Theorem 2 The extreme values of $\psi'\beta$ on the ellipsoid $(\beta-f)'\,A^{-1}(\beta-f) = c$ occur at the points $\beta = f \pm A\,\psi\,\sqrt{c/\psi'A\psi}$.

Proof. Setting the derivatives of the Lagrangian to zero yields $0 = \psi + A^{-1}(\beta-f)\lambda$, where λ is the Lagrange multiplier. Thus $(\beta-f) = -A\,\psi\lambda^{-1}$;

$c = (\beta-f)'\,A^{-1}(\beta-f) = \lambda\,\psi'A\,\psi$, and $\lambda^2 = \psi'A\,\psi/c$.

<u>Theorem 3</u> Given constraints of the form $\underset{\sim}{M}\underset{\sim}{R}\beta = \underset{\sim}{M}\underset{\sim}{r}$, where $\underset{\sim}{R}$ and $\underset{\sim}{r}$ are given but $\underset{\sim}{M}$ is free to vary, the extreme values of $\underset{\sim}{\psi}'\underset{\sim}{\hat\beta}$ (where $\hat\beta$ is a constrained least-squares point) occur at the points

$$\underset{\sim}{\hat\beta} = \underset{\sim}{b} - \underset{\sim}{V}\,\underset{\sim}{R}'\,A^{-1}\,(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})/2 \pm \underset{\sim}{V}\underset{\sim}{R}'\,\underset{\sim}{A}^{-1}\,\underset{\sim}{R}\,\underset{\sim}{V}\,\underset{\sim}{\psi}\sqrt{c/m} \qquad (15)$$

where $c = (\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})'\,\underset{\sim}{A}^{-1}\,(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})/4$

$m = \underset{\sim}{\psi}'\,\underset{\sim}{V}\,\underset{\sim}{R}'\,A^{-1}\,\underset{\sim}{R}\,\underset{\sim}{V}\,\underset{\sim}{\psi}.$

<u>Proof</u>. Select a value of $\gamma(=\underset{\sim}{R}\beta)$ on the ellipsoid

$$(\underset{\sim}{\gamma}-\underset{\sim}{f})'\,\underset{\sim}{A}^{-1}(\underset{\sim}{\gamma}-\underset{\sim}{f}) = (\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})'\,\underset{\sim}{A}^{-1}(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})/4 = c \qquad (16)$$

where $\underset{\sim}{f} = (\underset{\sim}{R}\underset{\sim}{b}+\underset{\sim}{r})/2$. Least squares subject to $R\beta = \gamma$ is

$$\underset{\sim}{\hat\beta} = \underset{\sim}{b} - \underset{\sim}{V}\,\underset{\sim}{R}'\,\underset{\sim}{A}^{-1}\,(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma}).$$

The linear combination $\underset{\sim}{\psi}'\underset{\sim}{\hat\beta} = \underset{\sim}{\psi}'\underset{\sim}{b} - \underset{\sim}{\psi}'\underset{\sim}{V}\,\underset{\sim}{R}'\,A^{-1}(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma})$ where $\underset{\sim}{\gamma}$ is constrained by (16). Thus the extreme values of $\underset{\sim}{\psi}'\underset{\sim}{\hat\beta}$ occur when $\underset{\sim}{\gamma} = \underset{\sim}{f} \pm \underset{\sim}{A}\underset{\sim}{\psi}^* \sqrt{c/\underset{\sim}{\psi}^{*\prime}\underset{\sim}{A}\underset{\sim}{\psi}^*}$ where $\underset{\sim}{\psi}^* = \underset{\sim}{A}^{-1}\,\underset{\sim}{R}\,\underset{\sim}{V}\,\underset{\sim}{\psi}$. Thus $\underset{\sim}{\gamma} = (\underset{\sim}{R}\underset{\sim}{b}+\underset{\sim}{r})/2 \pm \underset{\sim}{R}\,\underset{\sim}{V}\underset{\sim}{\psi}\sqrt{c/m}$, and $(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma}) = (\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})/2 \pm \underset{\sim}{R}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{\psi}\sqrt{c/m}$.

In order to compute the variance of the constrained least squares estimator (15) it is necessary to find the matrix $\underset{\sim}{M}$ such that the constraint $\underset{\sim}{M}\underset{\sim}{R}\beta = \underset{\sim}{M}\underset{\sim}{r}$ implies the given estimate. Unfortunately the matrix $\underset{\sim}{M}$ is not unique as has already been suggested in the discussion of points B and C in Figure 1. (Suppose for example that the extreme vector is just the least-squares vector $\underset{\sim}{b}$. Then obviously $\underset{\sim}{M}$ equal to the zero matrix is suitable. But any $\underset{\sim}{M}$ such that $\underset{\sim}{M}\underset{\sim}{R}\underset{\sim}{b} = \underset{\sim}{M}\underset{\sim}{r}$ will also do.) In most cases, the rank of the matrix $\underset{\sim}{M}\underset{\sim}{R}$ must be between 1 and q-1 where q is rank of $\underset{\sim}{R}$. If $\underset{\sim}{\beta}=\underset{\sim}{b}$, then M may have rank zero. If $\underset{\sim}{R}\beta = \underset{\sim}{r}$, then $\underset{\sim}{M}$ may have rank q, in which case it might as well be $\underset{\sim}{M}=\underset{\sim}{I}_q$. The following theorem describes the matrix $\underset{\sim}{M}$ with rank one necessary to produce $\underset{\sim}{\beta}$.

<u>Theorem 4</u> The estimate (15) is least squares subject to the constraint $\underset{\sim}{M}\underset{\sim}{R}\beta = \underset{\sim}{M}\underset{\sim}{r}$, where

$$\underset{\sim}{M} = (\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma})'\,\underset{\sim}{A}^{-1}, \qquad \text{and}$$

$$\underset{\sim}{\gamma} = (\underset{\sim}{R}\underset{\sim}{b}+\underset{\sim}{r})/2 \pm \underset{\sim}{R}\,\underset{\sim}{V}\,\underset{\sim}{\psi}\sqrt{c/m}.$$

The variance of least-squares subject to this restriction is

$$V(\underset{\sim}{\hat\beta}) = \underset{\sim}{V} - \underset{\sim}{V}\,\underset{\sim}{R}'\,\underset{\sim}{A}^{-1}(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma})(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma})'\,\underset{\sim}{A}^{-1}\underset{\sim}{R}\,\underset{\sim}{V} / \ (\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma})'\,\underset{\sim}{A}^{-1}\ (\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma})$$

if $(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma}) \neq \underset{\sim}{0}$.

<u>Proof:</u> This follows directly from (11) and (12) using the constraints indicated:

$$\underset{\sim}{\hat\beta} = \underset{\sim}{b} - \underset{\sim}{V}\,\underset{\sim}{R}'\,\underset{\sim}{A}^{-1}(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma})((\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma})'\,\underset{\sim}{A}^{-1}\,\underset{\sim}{R}\,\underset{\sim}{V}\,\underset{\sim}{R}'\,\underset{\sim}{A}^{-1}(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma}))^{-1}$$
$$(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma})'\,\underset{\sim}{A}^{-1}(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})$$

but $(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma}) = (\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})/2 \pm \underset{\sim}{R}\,\underset{\sim}{V}\,\underset{\sim}{\psi}\sqrt{c/m}$, and

$$\begin{aligned}
(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma})'\,\underset{\sim}{A}^{-1}(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma}) &= [(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r}) \pm 2\underset{\sim}{R}\,\underset{\sim}{V}\,\underset{\sim}{\psi}\sqrt{c/m}]'\,\underset{\sim}{A}^{-1} \\
&\qquad [(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r}) \pm 2\underset{\sim}{R}\,\underset{\sim}{V}\,\underset{\sim}{\psi}\sqrt{c/m}]/4 \\
&= (\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})'\,\underset{\sim}{A}^{-1}(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})/4 + \underset{\sim}{\psi}'\underset{\sim}{V}\,\underset{\sim}{R}'\,\underset{\sim}{A}^{-1}\,\underset{\sim}{R}\,\underset{\sim}{V}\,\underset{\sim}{\psi}\ c/m \\
&\qquad (\pm)\ (\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})'\,\underset{\sim}{A}^{-1}\underset{\sim}{R}\,\underset{\sim}{V}\,\underset{\sim}{\psi}\sqrt{c/m} \\
&= (\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})'\,\underset{\sim}{A}^{-1}\,(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})/2 \pm (\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})'\,\underset{\sim}{A}^{-1}\underset{\sim}{R}\,\underset{\sim}{V}\,\underset{\sim}{\psi}\sqrt{c/m} \\
&= (\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})'\,\underset{\sim}{A}^{-1}(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma}).
\end{aligned}$$

Thus
$$\begin{aligned}
\underset{\sim}{\hat\beta} &= \underset{\sim}{b} - \underset{\sim}{V}\,\underset{\sim}{R}'\,\underset{\sim}{A}^{-1}(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma}) \\
&= \underset{\sim}{b} - \underset{\sim}{V}\,\underset{\sim}{R}'\,\underset{\sim}{A}^{-1}\,[(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})/2 \pm \underset{\sim}{R}\,\underset{\sim}{V}\,\underset{\sim}{\psi}\sqrt{c/m}]
\end{aligned}$$

The following theorem implies a matrix $\underset{\sim}{M}$ with rank q-1 necessary to produce $\underset{\sim}{\hat\beta}$.

<u>Theorem 5</u> The estimate (15) is least-squares subject to the restriction that $(\underset{\sim}{R}\beta-\underset{\sim}{r})$ is proportional to $(\underset{\sim}{\gamma}-\underset{\sim}{r})$, $(\underset{\sim}{R}\beta-\underset{\sim}{r}) = (\underset{\sim}{\gamma}-\underset{\sim}{r})\Theta$ where Θ is a scalar parameter. The sampling variance of $\hat{\underset{\sim}{\beta}}$ given this restriction is

$$V(\hat{\underset{\sim}{\beta}}) = \underset{\sim}{V} - \underset{\sim}{V}\,\underset{\sim}{R}'\,\underset{\sim}{A}^{-1}\,[\underset{\sim}{A} - (\underset{\sim}{\gamma}-\underset{\sim}{r})V(\hat{\Theta})(\underset{\sim}{\gamma}-\underset{\sim}{r})']\underset{\sim}{A}^{-1}\,\underset{\sim}{R}\,\underset{\sim}{V}$$

where

$$V(\hat{\Theta}) = [(\underset{\sim}{\gamma}-\underset{\sim}{r})'\,\underset{\sim}{A}^{-1}(\underset{\sim}{\gamma}-\underset{\sim}{r})]^{-1}$$

<u>Proof.</u> Given Θ the estimator is

$$\hat{\underset{\sim}{\beta}}(\Theta) = \underset{\sim}{b} - \underset{\sim}{V}\,\underset{\sim}{R}'\,\underset{\sim}{A}^{-1}(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r}-(\underset{\sim}{\gamma}-\underset{\sim}{r})\Theta).$$

The least squares estimate of Θ can be found by minimizing the quadratic form

$$(\hat{\underset{\sim}{\beta}}-\underset{\sim}{b})'\,\underset{\sim}{V}^{-1}(\hat{\underset{\sim}{\beta}}-\underset{\sim}{b})$$
$$= (\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r}-(\underset{\sim}{\gamma}-\underset{\sim}{r})\Theta)'\,\underset{\sim}{A}^{-1}\,\underset{\sim}{R}\,\underset{\sim}{V}\,\underset{\sim}{R}'\,\underset{\sim}{A}^{-1}(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r}-(\underset{\sim}{\gamma}-\underset{\sim}{r})\Theta).$$

Setting the derivative of this with respect to Θ to zero yields

$$\Theta = (\underset{\sim}{\gamma}-\underset{\sim}{r})'\,\underset{\sim}{A}^{-1}(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{r})/(\underset{\sim}{\gamma}-\underset{\sim}{r})'\,\underset{\sim}{A}^{-1}(\underset{\sim}{\gamma}-\underset{\sim}{r}).$$

But because γ lies on the ellipsoid (16), $\Theta = 1$. Thus $\hat{\underset{\sim}{\beta}} = \underset{\sim}{b} - \underset{\sim}{V}\,\underset{\sim}{R}'\,\underset{\sim}{A}^{-1}(\underset{\sim}{R}\underset{\sim}{b}-\underset{\sim}{\gamma})$, which is least squares given $\underset{\sim}{R}\beta = \underset{\sim}{\gamma}$.

The variance of $\hat{\underset{\sim}{\beta}}$ given Θ is given by Equation (11). Since $\hat{\underset{\sim}{\beta}}$ given Θ is a linear function of Θ, we must add to this variance the matrix $\underset{\sim}{V}\,\underset{\sim}{R}'\,\underset{\sim}{A}^{-1}(\underset{\sim}{\gamma}-\underset{\sim}{r})\,V(\hat{\Theta})(\underset{\sim}{\gamma}-\underset{\sim}{r})'\,\underset{\sim}{A}^{-1}\,\underset{\sim}{R}\,\underset{\sim}{V}$ where $V(\hat{\Theta}) = (\underset{\sim}{\gamma}-\underset{\sim}{r})'\,\underset{\sim}{A}^{-1}\,\underset{\sim}{R}\,\mathrm{Var}(\underset{\sim}{b})\,\underset{\sim}{R}'\,\underset{\sim}{A}^{-1}(\underset{\sim}{\gamma}-\underset{\sim}{r})/((\underset{\sim}{\gamma}-\underset{\sim}{r})'\,\underset{\sim}{A}^{-1}(\underset{\sim}{\gamma}-\underset{\sim}{r}))^2 = [(\underset{\sim}{\gamma}-\underset{\sim}{r})'\,\underset{\sim}{A}^{-1}(\underset{\sim}{\gamma}-\underset{\sim}{r})]^{-1}$.

<u>Standard Deviations</u>

The program reports for each constrained regression the maximum and the minimum standard errors consistent with the given constrained regression. (A fixed estimate of σ^2, s^2, is used in all cases).

<u>Case 1</u> $\hat{\beta} = b$. The maximal variances make use of $\underset{\sim}{M} = \underset{\sim}{0}$, $Var(\underset{\sim}{\hat{\beta}}) = \underset{\sim}{V}$.

The minimal is given by theorem (5).

<u>Case 2</u> $R\hat{\beta} = r$. The minimal variance is given by formula (12) and the maximal by Theorem 4.

<u>Case 3</u> Otherwise, the maximal variances are given by theorem (4) and the minimal by theorem (5).

The reader may verify that $Var(\underset{\sim}{\psi}'\underset{\sim}{\beta})$ is unique.

VI. Extensions

The following additional material is in the planning and/or programming stage.

Posterior Modes. The location of modes along the contract curve given the Student prior (2) and a diffuse prior for σ^{-2}, $f(\sigma^{-2}) \alpha \sigma^{-2}$.

Measures of Dispersion. Exact or approximate posterior confidence intervals.

Local Sensitivity Analysis. The derivatives of features of the posterior distribution with respect to parameters of the prior distribution.

Simplification Analysis. Given the posterior distributions, simplify the model for forecasting and/or control.

Proxy Variable Analysis. Estimates and measures of dispersion when variables are measured with error, and when some prior information is available.

Contract Curves with other Prior Metrics

VII. Examples

(a) Two-dimensional example.

INPUTS $X'X = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ $X'Y = \begin{bmatrix} 9 \\ 6 \end{bmatrix}$ $s^2 = 3$

$T = 12$

$N^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $b^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

OUTPUTS (illustrated in Figure 1)

1. **Rotation Invariant Average Regressions**

(J = Number of Constraints)

ESS	J	1	2
72	2	.0	.0
34.9	1	2.25	1.5
30	0	4.0	1.0

Discussion. The first ideal point is the prior location (0,0), the third is the data location $(X'X)^{-1}X'Y = (4,1)$ and the second is a special weighted average of least squares with the first variable omitted (0,3) and least-squares with the second variable omitted (4.5,0). The triangle formed by these three points contains the contract curve. Note that only the first and last ideal points are actually on the curve.

2. Points on the Contract Curve

The points on the contract curve in steps of equal data likelihood are reported in Table 1. A description of the information provided may be found in Section IVb of this write-up.

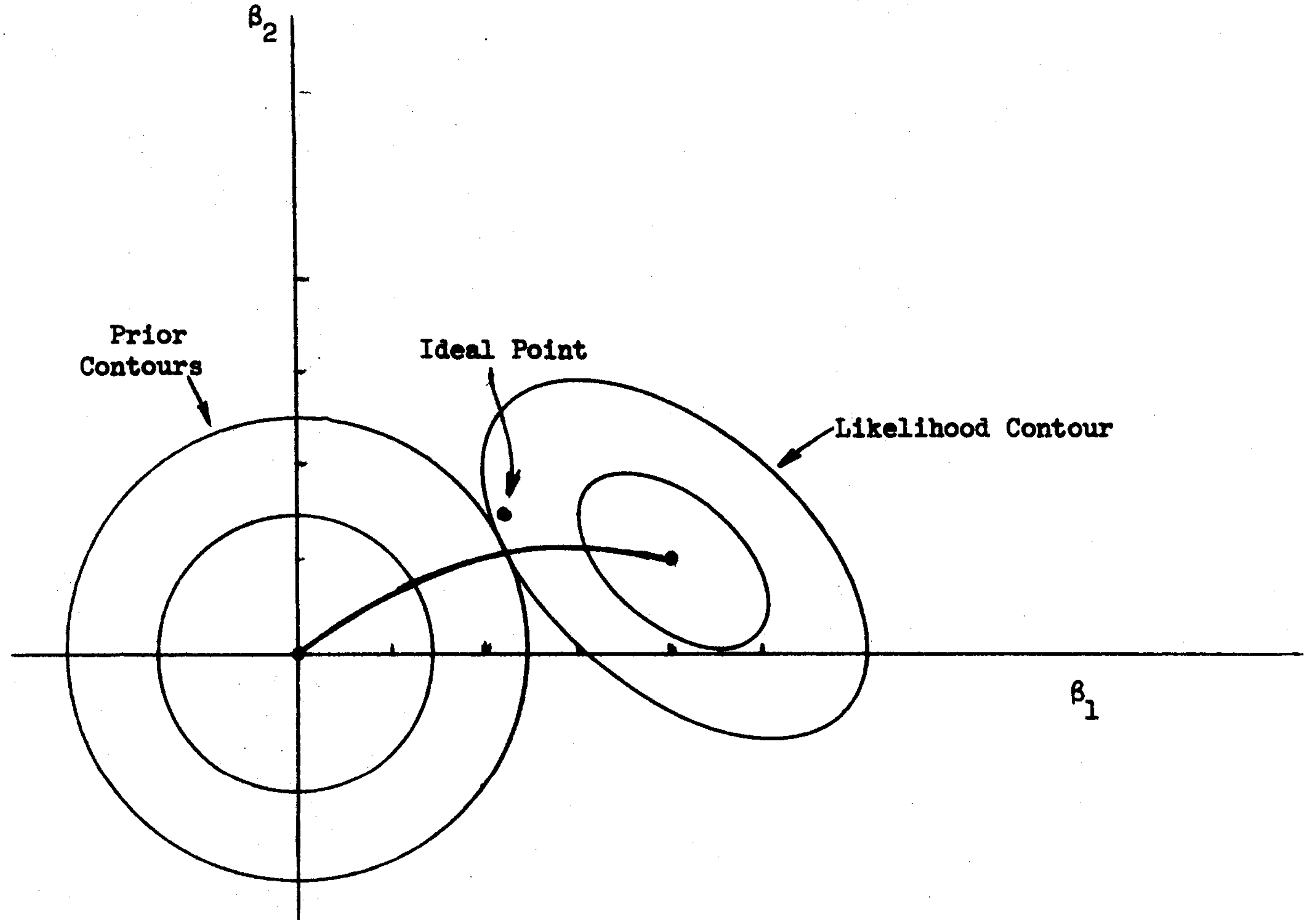

Figure 1   Contract Curve and Ideal Points

Table 1

A Trace of the Contract Curve

** AVERAGE REGRESSIONS **

SELECTED POINTS ON THE INFORMATION CONTRACT CURVE

CONTRACT CURVE IN STEPS OF EQUAL DATA LIKELIHOOD

	DAT LIK	PRI LIK	DAT DIST	PRI DIST	EUC DIST TO DATA	EUC DIST TO PRIOR	RHO	SIGMAT	FDATA	CHIPRIORD
1	0.0009	1.0000	1.0000	0.0	1.0000E+00	0.0		0.0	7.0000E+00	0.0
2	0.1008	0.2422	0.5725	0.4385	6.1433E-01	4.[illegible]846E-01	1.1846D+00	9.1880E-01	2.2944E+00	2.8363E+00
3	0.2007	0.1192	0.4790	0.5002	5.2609E-01	5.0021E-01	7.9875D-01	1.1189E+00	1.6058E+00	4.2535E+00
4	0.3006	0.0667	0.4144	0.5543	4.6350E-01	5.6432E-01	6.0591D-01	1.2847E+00	1.2018E+00	5.4137E+00
5	0.4005	0.0392	0.3615	0.6173	4.1101E-01	6.1730E-01	4.7839D-01	1.4458E+00	9.1493E-01	6.4780E+00
6	0.5005	0.0233	0.3145	0.6653	3.6305E-01	6.6497E-01	3.8236D-01	1.6172E+00	6.9224E-01	7.5175E+00
7	0.6004	0.0137	0.2700	0.7105	3.1650E-01	7.1054E-01	3.0396D-01	1.8138E+00	5.1023E-01	8.5827E+00
8	0.7003	0.0077	0.2256	0.7565	2.6871E-01	7.5649E-01	2.3580D-01	2.0594E+00	3.5628E-01	9.7286E+00
9	0.8002	0.0040	0.1785	0.8039	2.1629E-01	8.0593E-01	1.7271D-01	2.4063E+00	2.2292E-01	1.1042E+01
10	0.9001	0.0017	0.1226	0.8654	1.5179E-01	8.6538E-01	1.0860D-01	3.0345E+00	1.0526E-01	1.2731E+01
11	1.0000	0.0002	0.0	1.0000	0.0	1.0000E+00	0.0		0.0	1.7000E+01

THE CALCULATED POINTS FOLLOW:

	1	2
1	0.0	0.0
2	1.4738E+00	8.1499E-01
3	1.8315E+00	9.4819E-01
4	2.0891E+00	1.0244E+00
5	2.3770E+00	1.0751E+00
6	2.5071E+00	1.1099E+00
7	2.7018E+00	1.1327E+00
8	2.9015E+00	1.1445E+00
9	3.1199E+00	1.1438E+00
10	3.3865E+00	1.1237E+00
11	4.0000E+00	1.0000E+00

3. Bounds

The bounds over constrained least squares points are illustrated in Figure 2. The ellipse of constrained least squares points is the locus of tangencies between rays out of the origin and likelihood ellipses. This ellipse is described in terms of the points labelled A through K. These points and their associated standard errors are reported in Table 2.

Discussion. Point A is chosen to maximize β_1, B to minimize β_1, C to maximize β_2 and D to minimize β_2. Point E maximizes $\beta_1 + \beta_2$; point F minimizes $\beta_1 + \beta_2$. Point G maximizes β_1 given constraints that do not involve β_1 (i.e., only $\beta_2 = 0$), and point H similarly minimizes β_1. Points I and G maximize and minimize β_2 given constraints that do not involve β_2. Point J and K indicate the extreme values of $3\beta_1 + 2\beta_2$. For these last two points, two different standard errors are reported, since they can each be generated with two different constraints.

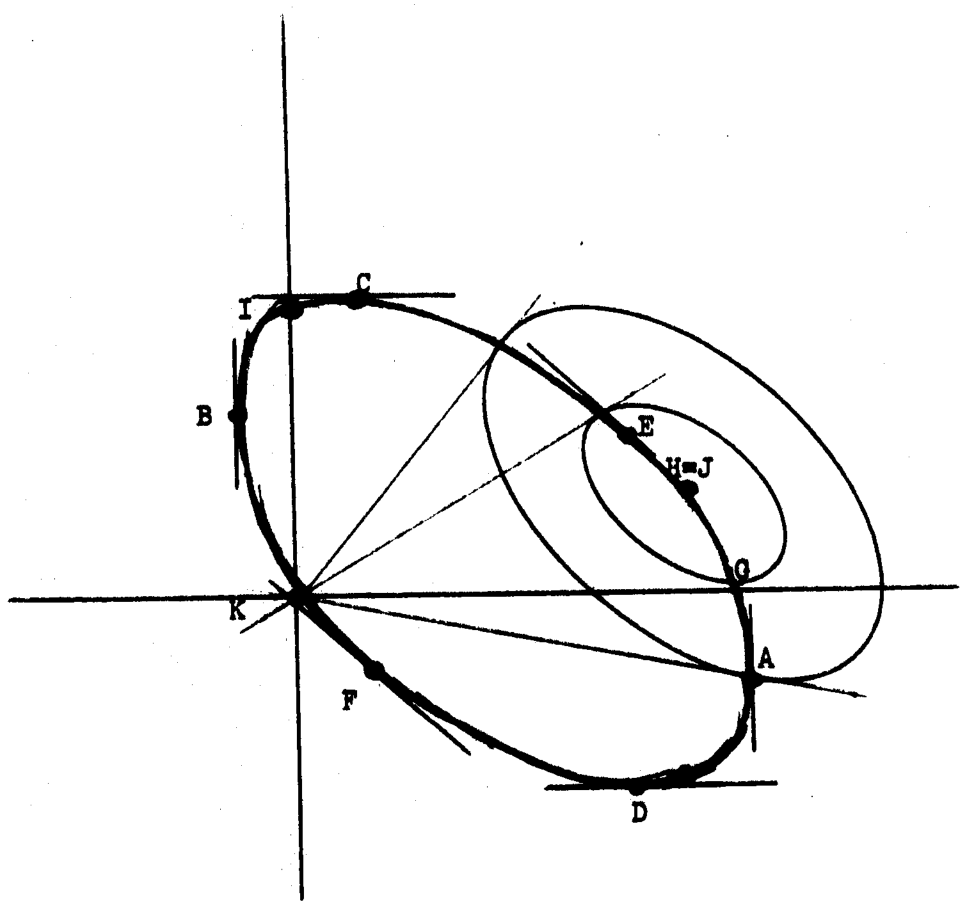

Figure 2 Set of Constrained Regressions

Table 2

Points on the Feasible Ellipsoid

	Point		Standard errors[1]	
A	4.65	-.82	1.32	.24
B	-.65	1.82	.49	1.39
C	.68	3.15	.23	1.09
D	3.32	-2.15	1.39	.90
E	3.32	1.82	.90	.49
F	.68	-.82	1.1	1.33
G	4.5	0	1.22	0
H	4.0	1.0	1.41	1.41
I	0	3.0	0	1.22
J	4.0	1.0	1.41(1.01)	1.41(.27)
K	0	0	.93(0)	1.39(0)

[1]These are standard errors of the coefficients given the constraint(s) necessary to produce the estimate. When two figures are given two constraints will produce the same estimate.

(b) Doubtful Variables

It is very common to have a model with a few explanatory variables which are known to belong in the equation, and a longer list of "doubtful" explanatory variables. The first set of variables is likely to be the focus of the analysis, and the second set is used to "control" for other influences. If the list of doubtful variables is long, estimation with all the doubtful variables included in the equation will produce large standard errors on the coefficients of the "focus" variables. In this situation, it is typical to try different subsets of the doubtful variables, and it is hoped that the coefficients of the focus variables will not change much as the list of doubtful variables is changed. But this search is both haphazard and non-exhaustive. Furthermore, if the coefficients of the focus variables change very much, this *ad hoc* search does not suggest how to average the many computed estimates into a single number.

SEARCH is ideally suited to deal with this problem. The interesting bounds that the program can report are the extreme estimates of the focus coefficients with ideally chosen doubtful variables included in the equation. There is no way of "fiddling" with the doubtful variables to get an estimate outside the reported range. The points on the contract curve reported by the program are mixtures of the 2^q regressions that could be computed using subsets of the q doubtful variables. Thus the program both searches exhaustively the set of possible regressions and also suggests weighted averages of the regressions, the latter being important when the bounds are wide.

The following example has eight "doubtful" regional dummy variables. The dependent variable is the wage rate and the focus variables are education of the wage earner, his age and the square of his age. A dummy variable for a region is necessary if the labor market in the given region is "separated" from the markets in other regions. To say that the dummy variables are doubtful is to say that in the absence of evidence to the contrary, we should view the labor market as a national market.

The estimated model with all the dummy variables included is (standard errors in parentheses)

$$\begin{aligned}
W = {} & \underset{(.34)}{.041}\, D_1 + \underset{(.32)}{.098}\, D_2 + \underset{(.46)}{.051}\, D_3 - \underset{(.34)}{.019}\, D_4 \\
& + \underset{(.46)}{.004}\, D_5 - \underset{(.43)}{.178}\, D_6 + \underset{(.50)}{.086}\, D_7 + \underset{(.35)}{.060}\, D_8 \\
& + \underset{(.030)}{.05}\, \mathrm{EDUC} + \underset{(.047)}{.137}\, \mathrm{AGE} - \underset{(.0006)}{.0015}\, (\mathrm{AGE})^2 + \underset{(.96)}{5.737}
\end{aligned}$$

where D_1 = Mid-Atlantic

D_2 = E. No. Central

D_3 = W. No. Central

D_4 = S. Atlantic

D_5 = E. So. Central

D_6 = W. So. Central

D_7 = Mountain

D_8 = Pacific

(New England omitted)

The bounds for the coefficients of the three focus variables are reported in the table below. The numbers in parentheses are the standard errors of these coefficients if the model that implied the estimate could be taken as given. (Remember that these bounds include regressions subject to constraints such as $\beta_1 = \beta_2$, which says the Mid-Atlantic and E. No. Central regions can be aggregated. They also include constraints of the form $\beta_i = 0$.)

EDUC	AGE	$(AGE)^2$
.0577 (.0177)	.139 (.029)	-.00147 (.00035)
.0446 (.0178)	.131 (.029)	-.00155 (.00035)

Each of these coefficients is quite insensitive to the choice of regional dummy variables.

Choice of points within these (narrow) bounds requires a more completely specified prior. Suppose that the coefficients of the doubtful variables are thought to be small in the sense that $\Sigma_{i=1}^{8} \beta_i^2$ is likely to be small. This prior "metric" implies the contract curve ~~in~~completely reported in Table 3. On this contract curve the extremes of all coefficients occur at the end points. One end point is least squares with all the dummies included; the other is least squares with all the dummies excluded. The extremes for the focus variables are:

EDUC	AGE	$(AGE)^2$
.0521	.1332	-.001489
.0502	.1336	-.001535

These bounds are almost points and it hardly seems necessary to select a particular point on the contract curve. But notice from Table 3 that the equation with the dummies omitted has a low likelihood ratio (equivalently a large F) and the data have a distinct preference for an estimate close to the unconstrained least squares points.

To conclude, for this particular problem, the ambiguity in the specification does not translate into substantial ambiguity in the focus coefficients. The specification error implies for example an interval of estimates for the education coefficient from .0446 to .0577. But the sampling standard error of this coefficient in the unconstrained model is .03, which is large compared to the specification range .0577 - .0446 = .0131. To put it briefly, the sampling error is more important than the specification error.

Table 3 Points on Contract Curve

Likelihood ratio	EDUC	AGE	$(AGE)^2$
.14	.0521	1.33	.00148
.31	.0517	1.34	.00150
.48	.0514	1.34	.00150
.66	.0511	1.35	.00151
.83	.0507	1.35	.00152
1.0	.0502	1.37	.00153

(c) <u>Distributed Lags</u>

Forty-seven quarterly observations, adjusted for auto-correlation .96, were used to estimate the following distributed lag process (std. errors in parentheses):

$$\begin{array}{lllllllllll} M_t = & .13\ Y_t & + & 1.96\ Y_{t-1} & - & .91\ Y_{t-2} & + & .55\ Y_{t-3} & - & .32\ Y_{t-4} & \\ & (.43) & & (.49) & & (.49) & & (.51) & & (.41) & \\ & -.42\ P_t & - & .52\ P_{t-1} & + & .33\ P_{t-2} & - & .72\ P_{t-3} & + & .23\ P_{t-4} & - .15 \\ & (.54) & & (.51) & & (.42) & & (.53) & & (.52) & \end{array}$$

where M is the logarithm of the quarterly flow of imports divided by a price index of imports, Y_t is the logarithm of GNP divided by the GNP price deflator, P_t is the logarithm of import prices divided by the GNP deflator.

This equation would be regarded as unlikely because of the peculiar changes in sign of each of the sets of the distributed lag coefficients. The priors now to be constructed are intended to smooth these coefficients. Four different priors are considered.

<u>Prior 1</u> <u>"Ridge" Regression,</u> $\underset{\sim}{N}^* = \underset{\sim}{I}$, $\underset{\sim}{b}^* = \underset{\sim}{0}$.

A spherical prior on the (slope) coefficients located at the origin is the first prior. This unlikely prior does imply some interesting bounds to be discussed below.

<u>Prior 2</u> <u>Small Differences</u>

$$\underset{\sim}{R} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \end{bmatrix}$$

$$\underset{\sim}{r}' = [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]$$

$$\underset{\sim}{V} = \underset{\sim}{I}$$

This prior reflects the fact that the first five coefficients are likely to be similar, $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5$, and the next five coefficients are likely to be similar, $\beta_6 = \beta_7 = \beta_8 = \beta_9 = \beta_{10}$. (This is one of Shiller's (1973) proposals.)

Prior 3

$$\underset{\sim}{R} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\underset{\sim}{r}' = [\,0 \quad 0 \quad 0 \quad 0\,]$$

$$\underset{\sim}{V} = \underset{\sim}{I}$$

This prior will smooth only the income coefficient pattern.

Prior 4

$$\underset{\sim}{R} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \end{bmatrix}$$

$$\underset{\sim}{r}' = [\,0 \quad 0 \quad 0 \quad 0\,]$$

$$\underset{\sim}{V} = \underset{\sim}{I}$$

This prior will smooth only the price coefficients.

Discussion of Ideal Points

The "ideal" points for each of the priors are given in Table 4. For each prior the first ideal point is constrained least squares given all the constraints implicit in the prior. The last ideal point is constrained least squares given none of the constraints, that is, it is just the unconstrained least squares estimate. The intermediate points are weighted averages of constrained least squares points. The next to last point is a weighted average of all regressions that involve one constraint. The next point uses constraints two at a time...

Any point on the contract curve is a weighted average of these ideal points. The traces of the contract curve risk missing important features of highly variable curves, but the ideal points cannot. Unfortunately, as will be seen below, the ideal points may be rather far from the curve.

Discussion of Points on the Contract Curve

Points on the contract curve are reported in Tables 5.1-5.4. These priors are intended to smooth the pattern of coefficients. In Table 5.1 the first point at which the signs of the income elasticities are equal has relative likelihood .9962 and is indicated by an arrow. A point with somewhat lower relative likelihood but with a smoother pattern of coefficients is also indicated by an arrow. (The third and fourth ideal points are actually not too different from these points.)

Points with comparable relative likelihood are "arrowed" for each of the next two priors reported in Table 5.2 and Table 5.3. Recall that prior 1 asserts that the coefficients are small but says nothing about the smoothness of the lag pattern, whereas the other priors indicate nothing about the size of the coefficients but instead are intended to smooth the coefficients. It is then surprising that the first prior seems to smooth the pattern of coefficients almost as well as the other priors. In Table 4.2 it is not possible to pick a point with higher relative likelihood that does not have a sign change in the income coefficients. The third

Table 4

Ideal Points (Rotation Invariant Average Regressions)

	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}
	0	0	0	0	0	0	0	0	0	0	-.04
	.14	.21	.08	.07	.05	-.05	-.01	-.03	.06	-.03	-.08
	.25	.39	.09	.08	.05	-.12	-.02	-.06	-.12	-.05	-.10
	.31	.56	.06	.08	.04	-.21	-.03	-.07	-.19	-.07	-.12
	.35	.72	.01	.07	.03	-.31	-.04	-.07	-.25	-.07	-.13
Prior 1	.37	.89	-.07	.07	.01	-.41	-.05	-.06	-.30	-.06	-.13
	.37	1.06	-.17	.09	-.02	-.50	-.08	-.03	-.35	-.04	-.14
	.35	1.24	-.29	.12	-.05	-.57	-.11	.01	-.40	-.01	-.15
	.32	1.42	-.43	.19	-.10	-.61	-.18	.08	-.46	.04	-.15
	.25	1.65	-.62	.32	-.17	-.58	-.30	.17	-.55	.11	-.15
	.13	1.96	-.91	.56	-.33	-.42	-.53	.33	-.72	.23	-.15
	.28	.28	.28	.28	.28	-.31	-.31	-.31	-.31	-.31	-.15
	.50	.45	.27	.14	.07	-.28	-.25	-.26	-.29	-.30	-.15
	.61	.59	.23	.55	-.02	-.41	-.24	-.19	-.23	-.19	-.16
	.65	.72	.16	-.00	-.05	-.55	-.22	-.12	-.21	-.12	-.16
Prior 2	.65	.86	.06	-.04	-.04	-.68	-.18	-.06	-.22	-.07	-.16
	.60	1.03	-.06	-.04	-.03	-.77	-.14	-.02	-.27	-.03	-.16
	.50	1.23	-.21	.01	-.04	-.80	-.12	.01	-.34	.01	-.16
	.36	1.50	-.45	.15	-.10	-.74	-.18	.09	-.44	.06	-.16
	.13	1.96	-.91	.56	-.33	-.42	-.53	.33	-.72	.23	-.15
	.29	.29	.29	.29	.29	-.93	.33	-.19	-.74	-.38	-.16
	.69	.59	.27	.04	-.09	-1.05	.15	-.02	-.52	.02	-.16
Prior 3	.73	.79	.11	-.05	-.06	-1.04	.08	.03	-.44	.08	-.16
	.53	1.2	-.15	-.06	-.02	-.88	-.07	.07	-.45	.10	-.16
	.13	1.96	-.91	.56	-.33	-.42	-.52	.33	-.72	.23	-.15
	.05	1.86	-.76	.33	-.08	-.24	-.24	-.24	-.24	-.24	-.15
	.12	1.85	-.75	.31	-.11	-.38	-.29	-.17	-.19	-.13	-.15
Prior 4	.17	1.83	-.76	.32	-.14	-.52	-.30	-.03	-.24	-.09	-.15
	.18	1.84	-.78	.38	-.20	-.51	-.30	.10	-.43	.04	-.15
	.13	1.96	-.91	.56	-.33	-.42	-.53	.33	-.72	.23	-.15

prior, which smooths only the income coefficients does do relatively better in that sense. Notice that prior 4 does not smooth the income coefficients, as is to be expected.

Perusal of the contract curves suggests that the response to the income stimulus is more rapid than the response to the price stimulus, and probably "smoother" as well. (Smoothness refers to the shape of the distribution of coefficients).

<u>Discussion of bounds</u>

The bounds for the coefficients are reported in Table 6. The type I bounds (potentially) involve constraints on the bounded coefficient; type II bounds force the prior to be diffuse on the coefficient in question.

Consider the bounds for β_1. The first bound [1.42, -1.29], indicates the potential set of estimates for β_1 given any homogeneous constraints on the first ten coefficients. The second bound, [1.3, -.19], makes use of constraints on coefficients other than β_1. The type I prior 2 bound [1.15, -.74] uses linear combinations of the two sets of constraints $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5$, $\beta_6 = \beta_7 = \beta_8 = \beta_9 = \beta_{10}$. The type II bound [.98, -.16] cannot use constraints involving β_1. The prior 3 bounds [1.04, -.62] and [.82, .12] use linear combinations of the constraints $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5$. The prior 4 bounds use the constraints $\beta_6 = \ldots = \beta_{10}$.

The prior one bounds for β_1 are quite wide. But they also potentially involve highly unlikely restrictions. The prior 2 bounds make use of a more reasonable family of models and they are somewhat narrower. The family of models is still unreasonably large since it includes constraints of the form $\beta_1 - \beta_2 = \beta_6 - \beta_7$ (a linear combination of two constraints). Ideally we could compute bounds that make use of constraints of the form $\underset{\sim}{M}\underset{\sim}{R}\underset{\sim}{\beta} = \underset{\sim}{M}\underset{\sim}{r}$ where R

$\sigma_1^2 N^*$

SELECTED POINTS ON THE INFORMATION CONTRACT CURVE

CONTRACT CURVE IN STEPS OF EQUAL DATA LIKELIHOOD

DAT LIK	PRI LIK	DAT DIST	PRI DIST	EUC DIST TO DATA	EUC DIST TO PRIOR	RHO	SIGMA1	FDATA	CHIPR
0.9873	1.0000	1.0000	0.0	1.0000E 00	0.0		0.0	1.7836E-03	0.0
0.9885	0.9971	0.9484	0.0306	9.8177E-01	3.0855E-02	2.1223D-01	3.0790E 00	1.6042E-03	5.289
0.9898	0.9874	0.8939	0.0640	9.6126E-01	6.4515E-02	9.2255D-02	4.5370E 00	1.4251E-03	2.314
0.9911	0.9688	0.8359	0.1009	9.3769E-01	1.0171E-01	5.2409D-02	5.8378E 00	1.2461E-03	5.758
0.9924	0.9388	0.7736	0.1424	9.0984E-01	1.4349E-01	3.2652D-02	7.1588E 00	1.0674E-03	1.147
0.9921	0.9452	0.7850	0.1347	9.1518E-01	1.3566E-01	3.5393D-02	6.9170E 00	1.0992E-03	1.025
0.9949	0.8311	0.6348	0.2438	8.3525E-01	2.4521E-01	1.3724D-02	1.0291E 01	7.1874E-04	3.362
0.9962	0.7328	0.5465	0.3161	7.7634E-01	3.1752E-01	8.2996D-03	1.2688E 01	5.3269E-04	5.651
0.9975	0.5969	0.4460	0.4073	6.9567E-01	4.0866E-01	4.6905D-03	1.6153E 01	3.5481E-04	9.383
0.9988	0.3853	0.3065	0.5537	5.5029E-01	5.5479E-01	1.9836D-03	2.3610E 01	1.6756E-04	1.734
1.0000	0.0446	0.0	1.0000	0.0	1.0000E 00	0.0		0.0	5.656

THE CALCULATED POINTS FOLLOW:

0.0000	0.0000	0.0000	-0.0000	-0.0000	0.0000	0.0000	-0.0000	-0.0000	-0.0000	-0.0
0.0375	0.0552	0.0196	0.0170	0.0119	-0.0136	-0.0021	-0.0078	-0.0155	-0.0077	-0.0
0.0779	0.1171	0.0375	0.0329	0.0227	-0.0310	-0.0048	-0.0165	-0.0340	-0.0162	-0.0
0.1211	0.1877	0.0524	0.0472	0.0316	-0.0536	-0.0083	-0.0262	-0.0563	-0.0256	-0.0
0.1672	0.2697	0.0622	0.0591	0.0380	-0.0836	-0.0127	-0.0367	-0.0836	-0.0356	-0.0
0.1588	0.2541	0.0610	0.0572	0.0371	-0.0776	-0.0118	-0.0348	-0.0783	-0.0338	-0.0
0.2619	0.4786	0.0519	0.0728	0.0390	-0.1771	-0.0256	-0.0558	-0.1568	-0.0545	-0.1
0.3093	0.6325	0.0151	0.0758	0.0301	-0.2571	-0.0369	-0.0595	-0.2109	-0.0600	-0.1
0.3432	0.8287	-0.0617	0.0827	0.0109	-0.3627	-0.0560	-0.0489	-0.2770	-0.0545	-0.1
0.3451	1.1425	-0.2389	0.1264	-0.0383	-0.5041	-0.1109	0.0048	-0.3768	-0.0161	-0.1
0.1326	1.9580	-0.9063	0.5564	-0.3296	-0.4197	-0.5276	0.3280	-0.7179	0.2308	-0.1

Prior 1

Table 5-1

SELECTED POINTS ON THE INFORMATION CONTRACT CURVE

CONTRACT CURVE IN STEPS OF EQUAL DATA LIKELIHOOD

DAT LIK	PRI LIK	DAT DIST	PRI DIST	EUC DIST TO DATA	EUC DIST TO PRIOR	RHO	SIGMA1	FDATA	CH
0.9936	1.0000	1.0000	0.0	1.0000E 00	0.0		0.0	8.9506E-04	0.[illegible]
0.9942	0.9975	0.9493	0.0170	9.8139E-01	4.6053E-02	1.1784D-01	3.5783E 00	8.0656E-04	4.[illegible]
0.9949	0.9883	0.8962	0.0370	9.6025E-01	9.7596E-02	4.5160D-02	5.6733E 00	7.1894E-04	2.1
0.9956	0.9621	0.8281	0.0671	9.2952E-01	1.6781E-01	1.8579D-02	8.5492E 00	6.1374E-04	7.0
0.9962	0.9246	0.7741	0.0955	9.0219E-01	2.2376E-01	1.0667D-02	1.1201E 01	5.3635E-04	1.4
0.9968	0.8499	0.7065	0.1375	8.6492E-01	2.9005E-01	5.6949D-03	1.5003E 01	4.4671E-04	2.9
0.9974	0.7253	0.6303	0.1933	8.1871E-01	3.5787E-01	3.0893D-03	1.9908E 01	3.5559E-04	5.8
0.9965	0.8863	0.7356	0.1185	8.8139E-01	2.6214E-01	7.3801D-03	1.3300E 01	4.8439E-04	2.1
0.9991	0.1727	0.3700	0.4519	5.8290E-01	5.7214E-01	5.0352D-04	4.6255E 01	1.2255E-04	3.1
0.9965	0.8934	0.7420	0.1145	8.8490E-01	2.5590E-01	7.8275D-03	1.2941E 01	4.9282E-04	2.0
1.0000	0.0002	0.0	1.0000	0.0	1.0000E 00	0.0		0.0	1.5

THE CALCULATED POINTS FOLLOW:

0.2884	0.2884	0.2884	0.2884	0.2884	-0.3077	-0.3077	-0.3077	-0.3077	-0.3077	-0
0.3487	0.3366	0.2827	0.2466	0.2283	-0.3043	-0.2907	-0.2926	-0.3036	-0.3040	-0
0.4147	0.3937	0.2736	0.1996	0.1636	-0.3114	-0.2740	-0.2727	-0.2948	-0.2919	-0
0.4979	0.4785	0.2538	0.1368	0.0843	-0.3458	-0.2554	-0.2391	-0.2756	-0.2554	-0
0.5538	0.5533	0.2285	0.0898	0.0334	-0.3965	-0.2428	-0.2060	-0.2570	-0.2212	-0
0.6009	0.6527	0.1831	0.0409	-0.0080	-0.4794	-0.2263	-0.1608	-0.2387	-0.1684	-0
0.6208	0.7703	0.1152	0.0033	-0.0295	-0.5766	-0.2033	-0.1114	-0.2345	-0.1151	-0
0.5840	0.6092	0.2045	0.0603	0.0068	-0.4421	-0.2338	-0.1805	-0.2452	-0.1910	-0
0.4881	1.2435	-0.2366	0.0537	-0.0678	-0.7409	-0.1665	0.0294	-0.3531	0.0169	-0
0.5796	0.5998	0.2088	0.0649	0.0106	-0.4342	-0.2353	-0.1848	-0.2470	-0.1961	-0
0.1326	1.9580	-0.9063	0.5564	-0.3296	-0.4197	-0.5276	0.3280	-0.7179	0.2308	-0

Prior 2

Table 5-2

SELECTED POINTS ON THE INFORMATION CONTRACT CURVE

CONTRACT CURVE IN STEPS OF EQUAL DATA LIKELIHOOD

CAT LIK	PRI LIK	DAT DIST	PRI DIST	EUC DIST TO DATA	EUC DIST TO PRIOR	RHO	SIGMA1	FDATA	CHI
0.9950	1.0000	1.0000	0.0	1.0000E 00	0.0		0.0	6.9116E-04	0.0
0.9955	0.9981	0.9486	0.0163	9.7662E-01	4.9391E-02	1.2222D-01	3.3751E 00	6.2190E-04	3.48
0.9960	0.9913	0.8942	0.0348	9.5197E-01	1.0475E-01	4.9347D-02	5.2278E 00	5.5266E-04	1.59
0.9965	0.9769	0.8364	0.0568	9.2524E-01	1.6819E-01	2.4848D-02	7.2470E 00	4.8346E-04	4.24
0.9970	0.9495	0.7742	0.0847	8.9434E-01	2.4200E-01	1.2567D-02	1.0019E 01	4.1429E-04	9.42
0.9975	0.8948	0.7067	0.1240	8.5306E-01	3.2244E-01	5.7450D-03	1.4559E 01	3.4516E-04	2.02
0.9980	0.7823	0.6319	0.1843	7.9101E-01	3.9210E-01	2.5403D-03	2.1499E 01	2.7601E-04	4.46
0.9985	0.5834	0.5460	0.2730	7.0240E-01	4.5467E-01	1.2238D-03	3.0387E 01	2.0608E-04	9.79
0.9986	0.5260	0.5240	0.2981	6.7768E-01	4.7062E-01	1.0432D-03	3.2763E 01	1.8980E-04	1.16
0.9996	0.0702	0.2824	0.6061	3.7770E-01	6.8274E-01	2.3649D-04	6.6148E 01	5.5136E-05	4.82
1.0000	0.0007	0.0	1.0000	0.0	1.0000E 00	0.0		0.0	1.31

THE CALCULATED POINTS FOLLOW:

0.2889	0.2889	0.2889	0.2889	0.2889	-0.9334	0.3315	-0.1859	-0.7381	-0.3799	-0.
0.3463	0.3328	0.2853	0.2518	0.2351	-0.9506	0.3056	-0.1612	-0.7062	-0.3220	-0.
0.4101	0.3837	0.2794	0.2098	0.1762	-0.9692	0.2760	-0.1334	-0.6701	-0.2574	-0.
0.4821	0.4456	0.2689	0.1611	0.1110	-0.9894	0.2411	-0.1013	-0.6283	-0.1840	-0.
0.5620	0.5263	0.2476	0.1032	0.0409	-1.0096	0.1979	-0.0633	-0.5787	-0.1004	-0.
0.6352	0.6392	0.1998	0.0397	-0.0213	-1.0208	0.1432	-0.0207	-0.5244	-0.0147	-0.
0.6576	0.7878	0.1105	-0.0042	-0.0510	-1.0024	0.0789	0.0182	-0.4857	0.0461	-0.
0.6175	0.9629	-0.0143	-0.0042	-0.0626	-0.9482	0.0042	0.0541	-0.4771	0.0808	-0.
0.6015	1.0072	-0.0479	0.0029	-0.0660	-0.9310	-0.0154	0.0631	-0.4792	0.0873	-0.
0.3876	1.4670	-0.4358	0.1984	-0.1533	-0.7065	-0.2445	0.1748	-0.5627	0.1503	-0.
0.1326	1.9580	-0.9063	0.5564	-0.3296	-0.4197	-0.5276	0.3280	-0.7179	0.2308	-0.

Prior 3

Table 5-3

SELECTED POINTS ON THE INFORMATION CONTRACT CURVE

CONTRACT CURVE IN STEPS OF EQUAL DATA LIKELIHOOD

CAT LIK	PRI LIK	DAT DIST	PRI DIST	EUC DIST TO DATA	EUC DIST TO PRIOR	RHO	SIGMA1	FDATA	CHI
0.9990	1.0000	1.0000	0.0	1.0000E 00	0.0		0.0	1.3330E-04	0.0
0.9991	0.9985	0.9487	0.0329	9.6553E-01	6.8288E-02	3.0384D-02	5.9500E 00	1.1997E-04	2.69
0.9992	0.9933	0.8944	0.0703	9.2999E-01	1.3557E-01	1.2467D-02	9.2524E 00	1.0663E-04	1.22
0.9993	0.9827	0.8366	0.1131	8.9204E-01	2.0094E-01	6.7414D-03	1.2533E 01	9.3299E-05	3.18
0.9994	0.9645	0.7745	0.1625	8.4963E-01	2.6461E-01	4.0471D-03	1.6111E 01	7.9968E-05	6.56
0.9995	0.9359	0.7070	0.2201	7.9989E-01	3.2796E-01	2.5461D-03	2.0230E 01	6.6636E-05	1.20
0.9996	0.8924	0.6323	0.2884	7.3891E-01	3.9354E-01	1.6251D-03	2.5220E 01	5.3284E-05	2.06
0.9997	0.8275	0.5466	0.3720	6.6056E-01	4.6603E-01	1.0202D-03	3.1699E 01	3.9825E-05	3.44
0.9997	0.8594	0.5861	0.3327	6.9788E-01	4.3270E-01	1.2589D-03	2.8589E 01	4.5796E-05	2.75
0.9999	0.5601	0.2898	0.6508	3.7631E-01	6.9443E-01	2.7008D-04	6.1063E 01	1.1198E-05	1.05
1.0000	0.2544	0.0	1.0000	0.0	1.0000E 00	0.0		0.0	2.48

THE CALCULATED POINTS FOLLOW:

0.0489	1.8595	-0.7630	0.3292	-0.0810	-0.2426	-0.2426	-0.2426	-0.2426	-0.2426	-0.
0.0708	1.8574	-0.7604	0.3229	-0.0899	-0.2870	-0.2561	-0.2175	-0.2278	-0.2076	-0.
0.0922	1.8544	-0.7586	0.3186	-0.0992	-0.3324	-0.2677	-0.1878	-0.2196	-0.1756	-0.
0.1120	1.8509	-0.7578	0.3174	-0.1091	-0.3773	-0.2774	-0.1539	-0.2201	-0.1467	-0.
0.1297	1.8475	-0.7583	0.3197	-0.1201	-0.4197	-0.2854	-0.1161	-0.2309	-0.1198	-0.
0.1446	1.8449	-0.7608	0.3265	-0.1328	-0.4577	-0.2929	-0.0746	-0.2533	-0.0925	-0.
0.1563	1.8445	-0.7659	0.3387	-0.1484	-0.4893	-0.3020	-0.0293	-0.2884	-0.0623	-0.
0.1641	1.8480	-0.7751	0.3579	-0.1683	-0.5119	-0.3157	0.0213	-0.3381	-0.0259	-0.
0.1612	1.8458	-0.7704	0.3484	-0.1588	-0.5032	-0.3088	-0.0019	-0.3141	-0.0430	-0.
0.1615	1.8843	-0.8240	0.4401	-0.2392	-0.5067	-0.3907	0.1675	-0.5144	0.0934	-0.
0.1326	1.9580	-0.9063	0.5564	-0.3296	-0.4197	-0.5276	0.3280	-0.7179	0.2308	-0.

Prior 4

Table 5-4

Table 6

Bounds for Coefficients

		β_1	β_2	β_3	β_4	β_5	$\Sigma_{i=1}^{5}\beta_i$	β_6	β_7	β_8	β_9	β_{10}	$\Sigma_{i=6}^{10}\beta_i$
PRIOR 1	Type I	1.42	2.55	1.13	1.89	1.14	1.92	1.50	1.35	1.51	1.32	1.77	1.92
		-1.29	-.59	-2.0	-1.33	-1.47	-.51	-1.92	-1.88	-1.19	-2.04	-1.55	-3.03
	Type II	1.42	2.39	1.08	1.82	.96	not	.37	.78	.81	0	.62	not
		-.19	1.10	-1.38	-.76	-.93	calc.	-1.62	-1.41	-.80	-1.75	-.97	calc.
PRIOR 2	Type I	1.15	2.22	.80	1.55	.89	1.65	.81	.69	.92	.63	1.10	-.49
		-.74	.02	-1.42	-.71	-.93	1.2	-1.53	-1.52	-.90	-1.66	-1.17	-2.15
	Type II	.98	2.21	0	.81	.27	same	.05	.52	.64	-.21	.50	same
		-.16	1.41	-1.26	-.61	-.81		-1.30	-.71	-.37	-1.33	-.65	
PRIOR 3	Type I	1.04	2.09	.66	1.41	.78	1.61	-.13	.44	.45	-.25	.37	-1.00
		-.62	.156	-1.28	-.57	-.82	1.24	-1.22	-.63	-.31	-.33	-.14	-2.00
	Type II	.82	2.1	-.06	.57	.08	same			same as above			same
		.12	1.5	-1.11	-.48	-.68							
PRIOR 4	Type I	.27	2.12	-.62	.68	-.03	1.46	.12	.04	.39	-.04	.43	-.87
		-.08	1.70	-1.05	.21	-.38	1.35	-.78	-.81	-.31	-.92	-.44	-1.45
	Type II			same as above			same	-.40	-.26	.37	-.57	.26	same
								-.72	-.63	.12	-.80	.007	

and r are given and where $\underset{\sim}{M}$ is restricted to be a block diagonal matrix, thereby disallowing constraints such as $\beta_1 - \beta_2 = \beta_6 - \beta_7$. The resulting bounds are much more difficult to compute and I have opted instead to use priors 3 and 4. But the family of constraints underlying prior 3 is still too large since it includes constraints such as $\beta_1 - \beta_2 = -(\beta_2 - \beta_3)$, which implies a very rugged, not a smooth, coefficient pattern. Thus, ideally, in this case, $\underset{\sim}{M}$ would be restricted to be a positive matrix; but again the resulting bounds are difficult to compute.

Incidentally, polynomial constraints of all orders are implicit in these bounds. The constraint that the coefficients lie on a polynomial of order zero is the constraint $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5$. The constraint that they lie on a line is $\beta_1 - \beta_2 = \beta_2 - \beta_3 = \beta_3 - \beta_4 = \beta_4 - \beta_5$, or a quadratic $(\beta_1 - \beta_2) - (\beta_2 - \beta_3) = (\beta_2 - \beta_3) - (\beta_3 - \beta_4) = (\beta_3 - \beta_4) - (\beta_4 - \beta_5)$. But in forming the bounds reported in Table 6 all of these constraints are feasible. Thus the method applies when the order of the polynomial constraint is unclear.

Although the bounds for individual coefficients are generally wide, the bound for the sum of the coefficients is relatively small. The prior one bound for the long run income elasticity is wide [1.92, -.51]. But the smoothness priors imply much narrower bounds [1.65, 1.2] and [1.61, 1.24]. Notice that the "mispecification uncertainty" has a relatively greater effect on the long run price elasticity in the sense that the bounds are relatively large.

REFERENCES

Chamberlain, G. and E. Leamer (1975), "Matrix weighted averages and posterior bounds," J. Roy. Statis. Soc. B 37, pp. 73-84.

Dickey, J.M. (1975), "Bayesian alternatives to the F-test and least-squares estimates in the normal linear model," in S.E. Fienberg and A. Zellner (eds) Bayesian Studies in Econometrics and Statistics, North Holland, Amsterdam.

Hoerl, A.E. and R.W. Kennard (1970), "Ridge regression: biased estimation for nonorthogonal problems," Technometrics 12, pp. 55-67.

Leamer, E.E. and G. Chamberlain (1975), "A Bayesian interpretation of pretesting," J. Roy. Statis. Soc., pp. 85-94.

Shiller, R. (1973), "A distributed lag estimator derived from smoothness priors," Econometrica 41, pp. 775-788.