PREDICTION IN RANDOM COEFFICIENT REGRESSION[1]

By

Rudolf Beran

Technical Report No. 299
April 1991

[1]This research was supported in part by NSF Grant DMS 9001710

Department of Statistics
University of California
Berkeley, California

PREDICTION IN RANDOM COEFFICIENT REGRESSION[1]

Rudolf Beran
University of California, Berkeley

ABSTRACT

Random coefficient regression and autoregressive models are important in diverse applications such as the classical statistical analysis of random and mixed effects models, the modelling of certain econometric and biological time series, and as a means for image compression. This paper develops nonparametric prediction intervals for a random coefficient regression model. The construction of these intervals requires a consistent estimate for the joint distribution of the model's random coefficients. Two such consistent estimates — a new one using minimum distance ideas and an earlier one based on estimated moments (Beran and Hall 1990) — are discussed.

AMS (1985) Subject classifications. Primary 62G05; secondary 62M20

Key words and phrases. Prediction interval, random coefficient, identifiable, consistent estimate, linear regression, bootstrap.

[1] This research was supported in part by NSF Grant DMS 9001710.

1. Introduction. Suppose that we observe n paired observations $\{(Y_i, X_i): 1 \le i \le n\}$ from the model

$$Y_i = A_i + B_i X_i \qquad i \ge 1 \tag{1.1}$$

where the $\{(A_i, B_i, X_i)\}$ are iid trivariate random vectors and (A_i, B_i) is independent of X_i for each i. Model (1.1) is called a *random coefficient* linear regression model. In a variant of this model, the $\{X_i\}$ are constants whose value may also depend on n. If we write $A_i = a + a_i$, $B_i = b + b_i$ with $a = EA_i$ and $b = EB_i$, then (1.1) can be put in the form

$$Y_i = (a + bX_i) + (a_i + b_i X_i). \tag{1.2}$$

Model (1.1) or (1.2) includes several important special cases:

(a) Ordinary linear regression. In (1.2), set $b_i = 0$ w.p.1.

(b) Structured heteroscedastic linear regression. This is (1.2) when b_i is not a degenerate random variable.

(c) Location-scale mixture model. In (1.1), the $\{X_i\}$ are not observed but their distribution is assumed known.

A variant of (1.1) is the model

$$Y_{ij} = A_{ij} + B_{ij} X_i, \quad 1 \le i \le I, \quad 1 \le j \le J \tag{1.3}$$

where the random vectors $\{\{A_{ij}, B_{ij}: 1 \le j \le J\}, X_i\}$ are iid and the random vector $\{A_{ij}, B_{ij}: 1 \le j \le J\}$ is independent of X_i for each i. Model (1.3) includes random effects and mixed effects models for the one-way layout. For instance, setting $B_{ij} = 1$ w.p.1 and requiring $EA_{ij} = 0$ yields a classical random effects model (Scheffé 1959, Chapter 7).

Random coefficient regression models, their autoregressive analogs, and models combining both features have been used to analyze certain econometric and biological time series. Good surveys of this work are Nicholls and Quinn (1982) and Nicholls and Pagan (1985). A very different application of bivariate random coefficient autoregressive models to image compression is given by Barnsley and Elton (1988). Nevertheless, many basic statistical problems associated with random coefficient models remain unsolved.

This paper seeks to construct a nonparametric prediction interval for Y_{n+1} under model (1.1), given the learning sample $S_n = \{(Y_i, X_i): 1 \le i \le n\}$ and the condition that $X_{n+1} = x$. As discussed in section 2, the proposed prediction interval requires a consistent estimate of F_{AB}, the distribution of the random coefficients (A_i, B_i). Section 3 presents two such consistent estimates — one using minimum distance ideas and the other based on estimated moments. Proofs for the minimum distance estimate are gathered in section 4. The estimated moment approach has been treated previously by Beran and Hall (1990) in the case where A_i and B_i are assumed independent.

2. Prediction intervals. We make the assumptions on model (1.1) that are stated in the first paragraph of the Introduction. The distributions F_{AB} of (A_i, B_i) and F_X of X_i are unknown. The problem is to construct a good prediction interval for Y_{n+1} given that $X_{n+1} = x$, on the basis of the learning sample $S_n = \{(Y_i, X_i) : 1 \le i \le n\}$.

For every real x, let $A_x(\cdot, F_{AB})$ denote the cdf of $A_i + B_i x$. Suppose $\hat{F}_{AB,n}$ is a consistent estimate of F_{AB} in the sense of weak convergence. Two methods for constructing $\hat{F}_{AB,n}$ are the topic of section 3. For every α in (0, 1), define upper and lower critical values from the quantiles of the estimated cdf $A(\cdot, \hat{F}_{AB,n})$:

$$\begin{aligned} \hat{c}_{x,n} &= A_x^{-1}[(1-\alpha)/2, \hat{F}_{AB,n}] \\ \hat{d}_{x,n} &= A_x^{-1}[(1+\alpha)/2, \hat{F}_{AB,n}]. \end{aligned} \tag{2.1}$$

The corresponding prediction interval for Y_{n+1} given $X_{n+1} = x$ is then

$$D_{x,n} = \{y : \hat{c}_{x,n} \le y \le \hat{d}_{x,n}\}. \tag{2.2}$$

Clearly, $D_{x,n}$ is a function of x and of the learning sample S_n.

The *conditional coverage probability* of $D_{x,n}$ for Y_{n+1}, given S_n and $X_{n+1} = x$, is

$$CP(D_{x,n} | x, S_n, F_{AB}) = A_x(\hat{d}_{x,n}, F_{AB}) - A_x(\hat{c}_{x,n}, F_{AB}). \tag{2.3}$$

The *coverage probability* of $D_{x,n}$ for Y_{n+1}, given $X_{n+1} = x$, is then

$$\begin{aligned} CP(D_{x,n} | x, F_{AB}) &= P[Y_{n+1} \in D_{x,n} | X_{n+1} = x, F_{AB}] \\ &= E\, CP(D_{x,n} | x, S_n, F_{AB}), \end{aligned} \tag{2.4}$$

the expectation being taken over the distribution of learning sample S_n.

PROPOSITION 1. Suppose the cdf $A_x(t, F_{AB})$ is continuous in t and $\hat{F}_{AB,n}$ converges weakly to F_{AB} in probability. Then, as n increases,

$$CP(D_{x,n} | x, S_n, F_{AB}) \xrightarrow{P} \alpha \tag{2.5}$$

and

$$CP(D_{x,n} | x, F_{AB}) \to \alpha \tag{2.6}$$

for every support point x of the distribution of X_i.

Thus, prediction interval $D_{x,n}$ has asymptotic coverage probability α for Y_{n+1}, given $X_{n+1} = x$. Section 4 contains a proof of Proposition 1. The following remarks supplement the Proposition.

(a) It follows from the proof of Proposition 1 that $P[Y_{n+1} > \hat{d}_{x,n} | X_{n+1} = x, S_n, F_{AB}]$ and $P[Y_{n+1} < \hat{c}_{x,n} | X_{n+1} = x, S_n, F_{AB}]$ both converge in probability to $(1-\alpha)/2$ as n increases. In an obvious sense, prediction interval $D_{x,n}$ is thus *probability-centered* for Y_{n+1} when n is large.

(b) A sufficient condition for the continuity in t of $A_x(t, F_{AB})$ is that F_{AB} be absolutely continuous with respect to Lebesgue measure in R^2.

(c) The standard linear regression model arises from (1.1) when $B_i = b$ w.p.1 and A_i has mean a and variance σ^2. The constants (a, b, σ^2) are unknown. In this special case, the distributions of A_i and of B_i are trivially independent. Let $\hat{b}_n$ denote the least squares estimate of b. Define $\hat{F}_{A,n}$ to be the empirical distribution of the residuals $\{Y_i - \hat{b}_n X_i : 1 \le i \le n\}$. The estimate $\hat{F}_{AB,n}$ is then defined as the product probability formed from $\hat{F}_{A,n}$ and the point mass at $\hat{b}_n$. With probability one, $\hat{F}_{AB,n}$ converges weakly to F_{AB}. This assertion may be checked by using the bounded Lipschitz metric for weak convergence.

Proposition 1 applies and the endpoints of prediction interval $D_{x,n}$ are just

$$\begin{aligned} \hat{c}_{x,n} &= \hat{b}_n x + \hat{F}_{A,n}^{-1}[(1-\alpha)/2] \\ \hat{d}_{x,n} &= \hat{b}_n x + \hat{F}_{A,n}^{-1}[(1+\alpha)/2]. \end{aligned} \tag{2.7}$$

$D_{x,n}$ can be viewed as the intersection of two one-sided prediction intervals for Y_{n+1}, each having asymptotic coverage probability $(1+\alpha)/2$. For a different analysis of these one-sided intervals in a standard linear regression model, see Beran (1990).

(d) A bootstrap algorithm based on a random sample from $\hat{F}_{AB,n}$ is a convenient way to approximate the cdf $A_x(\cdot, \hat{F}_{AB,n})$ and so the endpoints of prediction interval $D_{x,n}$. Draw q bootstrap variables $\{(A_k^*, B_k^*) : 1 \le k \le q\}$ from the estimated distribution $\hat{F}_{AB,n}$. The empirical cdf of the values $\{A_k^* + B_k^* x : 1 \le k \le q\}$ approximates $A_x(\cdot, \hat{F}_{AB,n})$ for large q.

(e) The cdf $A_x(\cdot, F_{AB})$ is the conditional cdf of Y_{n+1} given $X_{n+1} = x$. Suppose $\hat{A}_{x,n}^{-1}(\beta)$ is any consistent nonparametric estimate of $A_x^{-1}(\beta, F_{AB})$; Stone (1977) gives several constructions for $\hat{A}_{x,n}^{-1}(\beta)$. The prediction interval

$$\tilde{D}_{x,n} = \{y : \hat{A}_{x,n}^{-1}[(1-\alpha)/2] \le y \le \hat{A}_{x,n}^{-1}[(1+\alpha)/2]\} \tag{2.8}$$

also satisfies the conclusions of Proposition 1. Because the construction of $\tilde{D}_{x,n}$ does not fully use the structure of model (1.1), it is likely that $D_{x,n}$ is a more efficient prediction interval under model (1.1). The question needs more work.

3. Consistent estimation of F_{AB}. Both the minimum distance and moment-based estimates for F_{AB} that are described in this section require identifiability of F_{AB}. To discuss the latter point, let F_X denote the distribution of X_i and let $P_{YX} = P(F_{AB}, F_X)$ denote the distribution of (Y_i, X_i). Here P is the function of F_{AB} and F_X that is determined by model (1.1). The following result gives sufficient conditions for identifiability of F_{AB} from knowledge of P_{YX}. Proofs for both Proposition in this section are deferred to section 4.

PROPOSITION 2. Suppose that the support of F_X contains an open interval, that all moments of F_{AB} exist and that F_{AB} is uniquely determined by its moments. If $\tilde{F}_{AB}$ is a distribution on R^2 such that

$$P(\tilde{F}_{AB}, F_X) = P(F_{AB}, F_X) \tag{3.1}$$

then

$$\tilde{F}_{AB} = F_{AB}. \tag{3.2}$$

In our application of Proposition 2, we will assume that the supports of F_{AB} and $\tilde{F}_{AB}$ both lie within a fixed compact subset K of R^2. Then, the moment assumptions for the Proposition are satisfied automatically.

3.1. Minimum distance estimates. Suppose d metrizes weak convergence of probabilities on R^2. A numerically tractable metric d is generated by the $L^2(\mu)$ norm on bivariate characteristic functions, where μ is a probability measure on R^2 with strictly positive Lebesgue density. For the theory in this section, the choice of weak convergence metric does not matter.

Assume that the support of F_{AB} lies within a fixed compact subset K of R^2. Let $C_m = C_m(K)$ denote the set of all distributions in R^2 supported on at most m points in K, the mass of each support point being some multiple of 1/m. For example, C_3 contains all distributions that give probability 1/3 to three distinct points in K, plus all distributions that give probability 2/3 to one point in K and probability 1/3 to another point in K, plus all point mass distributions on K. Thus, C_m contains the unobservable empirical cdf of the random vector $\{(A_i, B_i): 1 \le i \le m\}$.

Let $\hat{P}_n$ denote the empirical cdf of the learning sample $S_n = \{(Y_i, X_i): 1 \le i \le n\}$ and let $\hat{F}_{X,n}$ denote the empirical cdf of the $\{X_i: 1 \le i \le n\}$. Define a minimum distance estimate $\hat{F}_{AB,n}$ of F_{AB} through the requirement

$$d[\hat{P}_n, P(\hat{F}_{AB,n}, F_{X,n})] = \inf_{G \in C_m} d[\hat{P}_n, P(G, \hat{F}_{X,n})] + \varepsilon_n \tag{3.3}$$

where $m = m_n$ depends on n and ε_n is a very small constant. Computational feasibility is the main reason for restricting the minimization on the right side of (3.3) to discrete distributions G in C_m. Consistency of $\hat{F}_{AB,n}$ is the subject of the following result.

PROPOSITION 3. Suppose that the support of F_X contains an open interval, that the support of F_{AB} lies within the compact K, and that $m_n \to \infty$, $\varepsilon_n \to 0$ as n increases. Then with probability one,

$$d(\hat{F}_{AB,n}, F_{AB}) \to 0 \tag{3.4}$$

as n increases.

Some remarks related to this Proposition:

(a) Computing $\hat{F}_{AB,n}$ is a question of minimizing a function of m_n variables — the variables being m_n candidate support points for $\hat{F}_{AB,n}$ with ties permitted. Rate-of-convergence arguments in (4.9) and (4.10) indicate that $m_n = n$ is best. However, computational constraints may force a smaller choice of m_n.

(b) Rates-of-convergence for $\hat{F}_{AB,n}$ and effective numerical algorithms for $\hat{F}_{AB,n}$ are currently under investigation by P.W. Millar and this author.

(c) Let

$$\mu_{j,k} = E(A_1^j B_1^k). \tag{3.5}$$

As will be discussed in subsection 3.2, the moments $\{\mu_{j,k}\}$ of F_{AB} of given order can be estimated consistently by a least squares algorithm. The estimated means and standard deviations for A_1 and B_1 provide guidance for the practical choice of the compact set K in the calculation of $\hat{F}_{AB,n}$.

(d) Several theoretical variants of $\hat{F}_{AB,n}$ preserve the consistency property (3.4). For instance, the minimization in definition (3.3) could be over all distributions G whose support lies in K. Or F_{AB} and F_X could be estimated simultaneously by the minimum distance criterion. However, these alternative estimates are much harder to calculate.

3.2. Moment-based estimates. The moments $\{\mu_{j,k}\}$ of F_{AB} all exist under the compact support assumption. From model (1.1),

$$E(Y_i^r) = \sum_{k=0}^{r} \binom{r}{k} \mu_{r-k,k} E(X_i^k). \tag{3.6}$$

Define the least squares estimates $\{\hat{\mu}_{r-k,k} : 0 \le k \le r\}$ to be the values that minimize

$$\sum_{i=1}^{n} [Y_i^r - \sum_{k=0}^{r} \binom{r}{k} \mu_{r-k,k} X_i^k]^2 \tag{3.7}$$

over all real $\{\mu_{r-k,k} : 0 \le k \le r\}$. Under moderate assumptions on F_X, the estimated moments $\{\hat{\mu}_{r-k,k} : 0 \le k \le r\}$ converge with probability one to the moments $\{\mu_{r-k,k} : 0 \le k \le r\}$.

This convergence suggests a moment-based procedure for estimating F_{AB}:

(i) Estimate the moments $\{\mu_{jk} : 0 \le j,k \le r_n\}$ by carrying out the least squares method above for $0 \le r \le r_n$.

(ii) Construct a distribution $\hat{F}_{AB,n}$ whose moments up to order r_n approximate the $\{\hat{\mu}_{jk} : 0 \le j,k \le n\}$ with increasing accuracy as n increases. One approach is through numerical conversion of an estimated Taylor expansion for the characteristic function of F_{AB}. Other methods are bivariate extensions of Hausdorff's univariate distribution constructions; see Shohat and Tamarkin (1943, p.90ff) for the latter.

We observe, without proof, that $\hat{F}_{AB,n}$ converges weakly to F_{AB} with probability one, provided $r_n \to \infty$ much more slowly than n.

When A_i and B_i are independent, $\mu_{j,k} = \alpha_j \beta_k$ with $\alpha_j = E(A_1^j)$ and $\beta_k = E(B_1^k)$. In this case, a simpler recursive scheme yields consistent moment estimates, as follows. When $r = 1$, minimization of (3.7) yields least squares estimates $\hat{\alpha}_1$, $\hat{\beta}_1$. Given the estimates $\{\hat{\alpha}_j, \hat{\beta}_j : 1 \le j \le r-1\}$, define the estimates $\hat{\alpha}_r$, $\hat{\beta}_r$ to be the values that minimize

$$\sum_{i=1}^{n} [\, Y_i^r - \sum_{k=1}^{r-1} \binom{r}{k} \hat{\alpha}_{r-k} \hat{\beta}_k - \alpha_r - \beta_r X_i^r \,]^2 \tag{3.8}$$

over all real α_r, β_r. A reconstruction algorithm of the type described in the preceding paragraph then yields marginal distribution estimates $\hat{F}_{A,n}$ and $\hat{F}_{B,n}$ that are based on the estimated moments $\{\hat{\alpha}_j : 1 \le j \le r_n\}$ and $\{\hat{\beta}_j : 1 \le j \le r_n\}$ respectively.

Beran and Hall (1990) prove that $\hat{F}_{A,n}$, $\hat{F}_{B,n}$ convey weakly w.p.1 to F_A, F_B respectively, provided r_n equals the integer part of $\varepsilon (\log n)^{1/2}$ for sufficiently small positive ε. A numerical example in their paper illustrates the feasibility of the moment method while noting possible difficulties with roundoff errors when r_n exceeds 30.

4. **Proofs**. This section gives proofs for Propositions 1 to 3.

PROOF OF PROPOSITION 1. Suppose $\{F_{AB,n}\}$ is any sequence of distributions converging weakly to F_{AB}. From the definition of cdf $A_x(\cdot, F_{AB})$, it follows that $A_x(\cdot, F_{AB,n})$ converges weakly to $A_x(\cdot, F_{AB})$. By the continuity of $A_x(t, F_{AB})$ in t, Polya's theorem, and the assumed consistency of $\{\hat{F}_{AB,n}\}$

$$\sup_t |A_x(t, \hat{F}_{AB,n}) - A_x(t, F_{AB})| \overset{P}{\rightarrow} 0. \tag{4.1}$$

In view of (4.1) and (2.1),

$$\begin{aligned} A_x(\hat{d}_{x,n}, F_{AB}) &= A_x(\hat{d}_{x,n}, \hat{F}_{AB,n}) + o_p(1) \\ &= (1+\alpha)/2 + o_p(1). \end{aligned} \tag{4.2}$$

Similarly,

$$A_x(\hat{c}_{x,n}, F_{AB}) = (1-\alpha)/2 + o_p(1). \tag{4.3}$$

The Proposition follows from (2.3), (4.2) and (4.3).

PROOF OF PROPOSITION 2. Suppose x lies in the open interval that is contained in the support of F_X. Let $\psi_x(t)$ and $\tilde{\psi}_x(t)$ denote the characteristic functions of $A_x(\cdot, F_{AB})$ and $A_x(\cdot, \tilde{F}_{AB})$ respectively. In view of (3.1)

$$\tilde{\psi}_x(t) = \psi_x(t), \quad -\infty < t < \infty. \tag{4.4}$$

Let $\phi(t,u)$ and $\tilde{\phi}(t,u)$ denote the characteristic functions of F_{AB} and $\tilde{F}_{AB}$ respectively. From (4.4) and model (1.1),

$$\tilde{\phi}(t, tx) = \phi(t, tx), \quad -\infty < t < \infty. \tag{4.5}$$

It remains to show that the moments of $\tilde{F}_{AB}$ equal those of F_{AB}.

Define the differential operators

$$\begin{aligned} D_{1,0} &= \frac{\partial}{\partial t} - \frac{x}{t}\frac{\partial}{\partial x} \\ D_{0,1} &= \frac{1}{t}\frac{\partial}{\partial x} \end{aligned} \tag{4.6}$$

and write $\phi^{(i,j)}(t,u) = \partial^{i+j}\phi(t,u)/\partial t^i \partial u^j$. For every t and every integer pair (i, j),

$$\phi^{(i,j)}(t, xt) = D_{1,0}^i D_{0,1}^j \phi(t, tx) \tag{4.7}$$

where the superscripts on the right side represent iteration of the respective operators. Setting $t = 0$ in (4.7) and in its analog for $\tilde{\phi}$ establishes the desired equality of moments.

PROOF OF PROPOSITION 3. For notational simplicity, the subscript n is omitted selectively during this argument. The first step is to show that

$$d[P(\hat{F}_{AB}, \hat{F}_X), P(F_{AB}, F_X)] \rightarrow 0 \quad \text{w.p.1.} \tag{4.8}$$

Indeed, by the triangle inequality and definition (3.3) of $\hat{F}_{AB}$,

$$d[P(\hat{F}_{AB}, \hat{F}_X), P(F_{AB}, F_X)]$$

$$\leq d[\hat{P}, P(\hat{F}_{AB}, \hat{F}_X)] + d[\hat{P}, P(F_{AB}, F_X)] \qquad (4.9)$$
$$\leq \inf_{G \in C_m} d[\hat{P}, P(G, \hat{F}_X)] + d[\hat{P}, P(F_{AB}, F_X)] + \varepsilon_n .$$

Let $\hat{\hat{F}}_{AB}$ denote the unobservable empirical cdf of the $\{(A_i, B_i) : 1 \leq i \leq m\}$. The last line in (4.9) is

$$\begin{aligned} &\leq d[\hat{P}, P(\hat{\hat{F}}_{AB}, \hat{F}_X)] + d[\hat{P}, P(F_{AB}, F_X)] + \varepsilon_n \\ &\leq 2d[\hat{P}, P(F_{AB}, F_X)] + d[P(\hat{\hat{F}}_{AB}, \hat{F}_X), P(F_{AB}, F_X)] + \varepsilon_n . \end{aligned} \qquad (4.10)$$

In this upper bound, the first term tends to zero w.p.1 by Glivenko-Cantelli applied to $\hat{P}_n$. The second term tends to zero w.p.1 by the weak continuity of $P(\cdot,\cdot)$ in its two arguments and by Glivenko-Cantelli applied to $\hat{\hat{F}}_{AB}$ and to $\hat{F}_X$. Hence (4.8) holds.

Suppose next that $\{H_n\}$ are distributions supported within K and $\{F_n\}$ are distributions on the real line such that

$$d[P(H_n, F_n), P(F_{AB}, F_X)] \rightarrow 0 \qquad (4.11)$$

and $F_n \Rightarrow F_X$. Then $H_n \Rightarrow F_{AB}$. Indeed, suppose not. Since the distributions $\{H_n\}$ are tight, by going to a subsequence, assume without loss of generality that $H_n \Rightarrow \tilde{F}_{AB} \neq F_{AB}$. Then, by the weak continuity of $P(\cdot,\cdot)$,

$$d[P(H_n, F_n), P(\tilde{F}_{AB}, F_X)] \rightarrow 0. \qquad (4.12)$$

Convergences (4.11) and (4.12) imply that $P(\tilde{F}_{AB}, F_X) = P(F_{AB}, F_X)$. Hence, by Proposition 1, $\tilde{F}_{AB} = F_{AB}$. This contradiction establishes that $H_n \Rightarrow F_{AB}$.

Setting $H_n = \hat{F}_{AB,n}$, $F_n = \hat{F}_{X,n}$ and using (4.8) now proves (3.4).

REFERENCES

BARNSLEY, M.F. and ELTON, J.H. (1988). A new class of Markov processes for image encoding. *Adv. Appl. Prob.* **20**, 14-32.

BERAN, R. (1990). Calibrating prediction regions. *J. Amer. Statist. Assoc.* **85**, 715-723.

BERAN, R. and HALL, P. (1990). Estimating coefficient distributions in random coefficient regressions. Unpublished paper.

NICHOLLS, D.F. and PAGAN, A.R. (1985). Varying coefficient regression. In: *Handbook of Statistics*, Vol. 5 (E.J. Hannan, P.R. Krishnaich, M.M. Rao, eds.) North-Holland, Amsterdam.

NICHOLLS, D.F. and QUINN, B.G. (1982). *Random Coefficient Autoregressive Models: An Introduction.* Springer-Verlag, New York.

SCHEFFÉ, H. (1959). *The Analysis of Variance*. Wiley, New York.

SHOHAT, J.A. and TAMARKIN, J.D. (1943). *The Problem of Moments*. American Mathematical Society, New York.

STONE, C.J. (1977). Consistent nonparametric regression. *Ann. Statist.* **5**, 595-645.