

# Fitting and interpreting correlated random coefficient (CRC) models using Stata

Oscar Barriga Cabanillas  
University of California  
Davis, CA  
obarriga@ucdavis.edu

Jeffrey D. Michler  
University of Illinois  
Urbana, IL  
jmichler@illinois.edu

Aleksandr Michuda  
University of California  
Davis, CA  
amichuda@ucdavis.edu

Emilia Tjernström  
University of Wisconsin  
Madison, WI  
tjernstroem@wisc.edu

**Abstract.** This article introduces the user-written command `randcoef`, which fits the correlated random effects (CRE) and correlated random coefficient (CRC) models discussed in Suri (2011). While the approach has been around for a decade, its use has been limited by the computationally intensive nature of the estimation procedure that relies on using the optimal minimum distance (OMD) estimator. The command can accommodate up to five rounds of panel data and offers several options, including alternative weight matrices for estimation and the inclusion of additional endogenous regressors. We also present post-estimation analysis using sample data in order to facilitate understanding and interpretation of results.

**Keywords:** `randcoef`, correlated random effects, correlated random coefficients, technology adoption, heterogeneity

## 1 Introduction

Random coefficient models have frequently been used to help explain heterogeneity in returns to human capital (Heckman and Vytlacil 1998) or the adoption of a new technology (Suri 2011). At its most basic, the random coefficient model can be written as:

$$y_i = \alpha_{0i} + \alpha_{1i}h_i. \quad (1)$$

The outcome variable (earnings in Heckman and Vytlacil (1998), maize yields in Suri (2011)), is a function of the rate of return,  $\alpha_{1i}$ , to a choice variable  $h_i$  (for example, investment in human capital or technology adoption). The rate of return varies by individual, as does the intercept term. Assume  $\alpha_{0i} = \bar{\alpha}_0 + \epsilon_{0i}$  and  $\alpha_{1i} = \bar{\alpha}_1 + \epsilon_{1i}$ , where  $\epsilon_{0i}$  and  $\epsilon_{1i}$  are zero in expectation. We can then rewrite equation (1) as:

$$y_i = \bar{\alpha}_0 + \bar{\alpha}_1h_i + (\epsilon_{0i} + \epsilon_{1i}h_i). \quad (2)$$

The difficulty in identifying the model arises when  $\alpha_{1i}$ , the rate of return for the individ-

ual, is correlated with the individual's choice of  $h_i$ . In the case of schooling, this would be the case if an individual's return to education influences her choice to invest in education. In the case of technology adoption, the classic example is when an individual's return to adoption influences his choice to adopt the new technology.

The literature has produced several alternative approaches to identifying the correlated random coefficient (CRC) model in equation (2). Both Heckman and Vytlačil (1998) and Wooldridge (2003) use instrumental variables. More recently, Suri (2011) developed a generalization of the Chamberlain (1984) fixed effects method and applied it to the CRC model. In this article, we present the `randcoef` command, which enables a user to fit Suri's CRC model. The approach, which is structural in nature, uses a set of reduced form parameters to recover the structural parameters of interest via the optimal minimum distance estimator (OMD). The command allows for the use of equally weighted minimum distance (EWMD) or diagonally weighted minimum distance (DWMD) instead of the inverse of the variance-covariance matrix of the reduced form estimates. Additionally, it allows for the inclusion of exogenous covariates as controls and can also accommodate the inclusion of an additional endogenous covariate.<sup>1</sup>

The structural approach to identifying the CRC model has several advantages over the instrumental variables approach. First, it does not require the researcher to choose an instrument and to defend the exclusion restriction. Rather, identification relies on the linear projection of an individual's rate of return onto their complete history of adoption, similar to the correlated random effects method pioneered by Mundlak (1978) and Chamberlain (1984). Second, the approach allows for a test of the importance of the individual's rate of return (comparative advantage in Suri's terminology) in the adoption decision. Third, the approach allows researchers to recover the distribution of the rate of return for post-estimation analysis, which we discuss using the practical example of seed adoption decisions among a sample of farmers in Ethiopia.

## 2 The CRE and CRC models

This section lays out the math behind the calculations of the correlated random effects (CRE) and correlated random coefficients (CRC) models plus the CRC model with an additional endogenous covariate. Two-period versions of these models are developed in Suri (2011). The methodology does not depend on the number of periods in the data but rather on the specifics of the structural equations, reduced form equations, and variance-covariance matrix do vary depending on the number of periods. To fix ideas we outline the calculations of a three-period model. However, the `randcoef` command can accommodate anywhere between two and five periods of panel data.

---

1. In the case of the hybrid maize adoption story presented in Suri (2011) this additional endogenous variable is the choice to apply fertilizer in addition to the choice to adopt hybrid maize.

## 2.1 Three-period CRE model

We start with a simple three-period, no covariates CRE model, for which the data generating process is given by:

$$y_{it} = \delta + \beta h_{it} + \alpha_i + u_{it}, \quad (3)$$

where  $y_{it}$  is the variable of interest for individual  $i$  at time  $t$ . The outcome of interest is a function of a binary indicator of adoption ( $h_{it}$ ), an indicator for the individual ( $\alpha_i$ ) and an idiosyncratic error term  $u_{it} \sim \mathcal{N}(0, \sigma_u^2)$ . We assume strict exogeneity of the error term.

In a fixed effects (FE) model the unique individual indicators are simply dummy variables. In the CRE model we replace the FEs by their linear projections upon the history of the individual's adoption behavior:

$$\alpha_i = \lambda_0 + \lambda_1 h_{i1} + \lambda_2 h_{i2} + \lambda_3 h_{i3} + \nu_i, \quad (4)$$

where  $h_{i1}$  is an indicator that equals one if individual  $i$  adopted the technology at  $t = 1$  and the  $\lambda$ 's are the projection coefficients. Substituting equation (4) into equation (3) we get:

$$y_{it} = \delta + \beta h_{it} + \lambda_0 + \lambda_1 h_{i1} + \lambda_2 h_{i2} + \lambda_3 h_{i3} + \nu_i + u_{it}. \quad (5)$$

Let  $\epsilon_{it} = \nu_i + u_{it}$  where  $\epsilon_{it}$  is strictly exogenous. For each time period we have:

$$y_{i1} = (\delta + \lambda_0) + (\beta + \lambda_1)h_{i1} + \lambda_2 h_{i2} + \lambda_3 h_{i3} + \epsilon_{i1}, \quad (6)$$

$$y_{i2} = (\delta + \lambda_0) + \lambda_1 h_{i1} + (\beta + \lambda_2)h_{i2} + \lambda_3 h_{i3} + \epsilon_{i2}, \quad (7)$$

$$y_{i3} = (\delta + \lambda_0) + \lambda_1 h_{i1} + \lambda_2 h_{i2} + (\beta + \lambda_3)h_{i3} + \epsilon_{i3}. \quad (8)$$

These are the structural equations for each period. Note that we cannot identify the variable of interest ( $\beta$ ) by estimating these equations. Instead, we estimate the following reduced form equations:

$$y_{i1} = \delta_1 + \gamma_1 h_{i1} + \gamma_2 h_{i2} + \gamma_3 h_{i3} + n_{i1}, \quad (9)$$

$$y_{i2} = \delta_2 + \gamma_4 h_{i1} + \gamma_5 h_{i2} + \gamma_6 h_{i3} + n_{i2}, \quad (10)$$

$$y_{i3} = \delta_3 + \gamma_7 h_{i1} + \gamma_8 h_{i2} + \gamma_9 h_{i3} + n_{i3}. \quad (11)$$

We can use the nine reduced form parameters to calculate the four structural parameters.

This requires placing the following restrictions on the parameters:

$$\begin{array}{lll} \gamma_1 = (\beta + \lambda_1) & \gamma_4 = \lambda_1 & \gamma_7 = \lambda_1 \\ \gamma_2 = \lambda_2 & \gamma_5 = (\beta + \lambda_2) & \gamma_8 = \lambda_2 \\ \gamma_3 = \lambda_3 & \gamma_6 = \lambda_3 & \gamma_9 = (\beta + \lambda_3) \end{array}$$

We estimate equations (9)-(11) as a set of seemingly unrelated regressions and then preserve the nine reduced form parameters in a vector  $\pi_{[9 \times 1]}$ . We can also preserve the variance-covariance matrices from the three regressions in a large symmetric block matrix  $\mathbf{V}_{[9 \times 9]}$ . The restrictions on the  $\gamma$ 's can be expressed as  $\pi = \mathbf{H}\delta$  where  $\mathbf{H}_{[9 \times 4]}$  is a restriction matrix embodying the nine restrictions on  $\gamma$  and  $\delta_{[4 \times 1]}$  is a vector of our four structural parameters.

The optimal minimum distance (OMD) function is:

$$\min_{\delta} = \{\pi - \mathbf{H}\delta\}' \mathbf{V}^{-1} \{\pi - \mathbf{H}\delta\}. \quad (12)$$

Solving for  $\delta$  we get:

$$\delta = (\mathbf{H}'\mathbf{V}^{-1}\mathbf{H})^{-1} \mathbf{H}'\mathbf{V}^{-1}\pi, \quad (13)$$

which is the optimal minimum distance estimator. The key to getting the OMD to produce the correct estimates requires making sure the restriction matrix is correctly specified. In the three-period CRE case we have:

$$\begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \gamma_5 \\ \gamma_6 \\ \gamma_7 \\ \gamma_8 \\ \gamma_9 \end{pmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \beta \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix}, \quad (14)$$

which is just  $\pi = \mathbf{H}\delta$ .

Finally, note that the above allows us to calculate the OMD coefficients but not the structural variance-covariance matrix, which requires taking derivatives of the restrictions. In the CRE model this is trivial since all the derivatives are equal to one.

## 2.2 Three-period CRC model

The three-period CRC model is methodologically similar to the CRE model but the specifics of the structural equations, reduced form equations, and restrictions differ. The data generating process is given by:

$$y_{it} = \delta + \beta h_{it} + \theta_i + \phi \theta_i h_{it} + \tau_i + u_{it}. \quad (15)$$

where  $\phi$  is the coefficient on an individual's rate of return or comparative advantage in adoption ( $\theta_i$ ),  $\tau_i$  is an individual's absolute advantage (equivalent to a household FE), and all other terms are as previously defined. To estimate equation (15), we must eliminate the dependence of  $\theta_i$  on  $h_{it}$ . To do this, Suri (2011) and Chamberlain (1984) replace  $\theta_i$  with its linear projection onto the individual's full history of adoption:

$$\theta_i = \lambda_0 + \lambda_1 h_{i1} + \lambda_2 h_{i2} + \lambda_3 h_{i3} + \lambda_4 h_{i1} h_{i2} + \lambda_5 h_{i1} h_{i3} + \lambda_6 h_{i2} h_{i3} + \lambda_7 h_{i1} h_{i2} h_{i3} + \nu_i, \quad (16)$$

Note that we must include the history of interactions because while the projection error  $\nu_i$  is uncorrelated with each individual history it is not necessarily uncorrelated with the product of the histories.

Substituting equation (16) into equation (15) and writing out each time period's function gives:

$$y_{i1} = (\delta + \lambda_0) + [\beta + \phi \lambda_0 + \lambda_1(1 + \phi)]h_{i1} + \lambda_2 h_{i2} + \lambda_3 h_{i3} + [\phi \lambda_2 + \lambda_4(1 + \phi)]h_{i1} h_{i2} + [\phi \lambda_3 + \lambda_5(1 + \phi)]h_{i1} h_{i3} + \lambda_6 h_{i2} h_{i3} + [\phi \lambda_6 + \lambda_7(1 + \phi)]h_{i1} h_{i2} h_{i3} + (\nu_i + \phi \nu_i h_{i1} + u_{i1}),$$

$$y_{i2} = (\delta + \lambda_0) + \lambda_1 h_{i1} + [\beta + \phi \lambda_0 + \lambda_2(1 + \phi)]h_{i2} + \lambda_3 h_{i3} + [\phi \lambda_1 + \lambda_4(1 + \phi)]h_{i1} h_{i2} + \lambda_5 h_{i1} h_{i3} + [\phi \lambda_3 + \lambda_6(1 + \phi)]h_{i2} h_{i3} + [\phi \lambda_5 + \lambda_7(1 + \phi)]h_{i1} h_{i2} h_{i3} + (\nu_i + \phi \nu_i h_{i2} + u_{i2}),$$

$$y_{i3} = (\delta + \lambda_0) + \lambda_1 h_{i1} + \lambda_2 h_{i2} + [\beta + \phi \lambda_0 + \lambda_3(1 + \phi)]h_{i3} + [\phi \lambda_1 + \lambda_5(1 + \phi)]h_{i1} h_{i3} + \lambda_4 h_{i1} h_{i2} + [\phi \lambda_2 + \lambda_6(1 + \phi)]h_{i2} h_{i3} + [\phi \lambda_4 + \lambda_7(1 + \phi)]h_{i1} h_{i2} h_{i3} + (\nu_i + \phi \nu_i h_{i3} + u_{i3}).$$

These are the structural yield equations for each period. From these we can estimate the following reduced form equations:

$$y_{i1} = \delta_1 + \gamma_1 h_{i1} + \gamma_2 h_{i2} + \gamma_3 h_{i3} + \gamma_4 h_{i1} h_{i2} + \gamma_5 h_{i1} h_{i3} + \gamma_6 h_{i2} h_{i3} + \gamma_7 h_{i1} h_{i2} h_{i3} + n_{i1}, \quad (17)$$

$$y_{i2} = \delta_2 + \gamma_8 h_{i1} + \gamma_9 h_{i2} + \gamma_{10} h_{i3} + \gamma_{11} h_{i1} h_{i2} + \gamma_{12} h_{i1} h_{i3} + \gamma_{13} h_{i2} h_{i3} + \gamma_{14} h_{i1} h_{i2} h_{i3} + n_{i2}, \quad (18)$$

$$y_{i3} = \delta_3 + \gamma_{15} h_{i1} + \gamma_{16} h_{i2} + \gamma_{17} h_{i3} + \gamma_{18} h_{i1} h_{i2} + \gamma_{19} h_{i1} h_{i3} + \gamma_{20} h_{i2} h_{i3} + \gamma_{21} h_{i1} h_{i2} h_{i3} + n_{i3}. \quad (19)$$

These equations give 21 reduced form coefficients ( $\gamma_1 - \gamma_{21}$ ) from which we can estimate ten structural parameters ( $\beta, \phi, \lambda_0 - \lambda_7$ ). Note that if we normalize the  $\theta$ 's so that  $\sum \theta_i = 0$  we can effectively eliminate  $\lambda_0$  and only need to estimate nine structural parameters.

The necessary restrictions to identify the structural parameters are:

$$\begin{array}{lll} \gamma_1 = [\beta + \phi\lambda_0 + \lambda_1(1 + \phi)] & \gamma_8 = \lambda_1 & \gamma_{15} = \lambda_1 \\ \gamma_2 = \lambda_2 & \gamma_9 = [\beta + \phi\lambda_0 + \lambda_2(1 + \phi)] & \gamma_{16} = \lambda_2 \\ \gamma_3 = \lambda_3 & \gamma_{10} = \lambda_3 & \gamma_{17} = [\beta + \phi\lambda_0 + \lambda_3(1 + \phi)] \\ \gamma_4 = [\phi\lambda_2 + \lambda_4(1 + \phi)] & \gamma_{11} = [\phi\lambda_1 + \lambda_4(1 + \phi)] & \gamma_{18} = \lambda_4 \\ \gamma_5 = [\phi\lambda_3 + \lambda_5(1 + \phi)] & \gamma_{12} = \lambda_5 & \gamma_{19} = [\phi\lambda_1 + \lambda_5(1 + \phi)] \\ \gamma_6 = \lambda_6 & \gamma_{13} = [\phi\lambda_3 + \lambda_6(1 + \phi)] & \gamma_{20} = [\phi\lambda_2 + \lambda_6(1 + \phi)] \\ \gamma_7 = [\phi\lambda_6 + \lambda_7(1 + \phi)] & \gamma_{14} = [\phi\lambda_5 + \lambda_7(1 + \phi)] & \gamma_{21} = [\phi\lambda_4 + \lambda_7(1 + \phi)] \end{array}$$

Similar to before, we estimate equations (17)-(19) as seemingly unrelated regressions and preserve the 21 reduced form parameters in a vector  $\pi_{[21 \times 1]}$ , the variance-covariance matrices in a large symmetric block matrix  $\mathbf{V}_{[21 \times 21]}$ , the restrictions as  $\mathbf{H}_{[21 \times 9]}$ , and the structural parameters as  $\delta_{[9 \times 1]}$ .

We now need to take derivatives of the elements in  $\mathbf{H}\delta$  with respect to each of the structural parameters. This gives us the 63 derivatives, which we preserve in the structural variance-covariance matrix to facilitate calculation of standard errors.

### 2.3 Three-period CRC model with endogenous covariates

Following Suri (2011), we expand the three-period CRC model to allow for an additional endogenous covariate. This requires that we not only account for the whole history of adoption but those histories interacted with the endogenous variable. We start by including the endogenous covariate,  $f_{it}$ , in the data generating process:

$$y_{it} = \delta + \beta h_{it} + \rho f_{it} + \theta_i + \phi \theta_i h_{it} + \tau_i + u_{it}. \quad (20)$$

We can write out the linear projection of the  $\theta_i$ 's on the history of the individual's adoption behavior of both endogenous technologies:

$$\begin{aligned}
\theta_i = & \lambda_0 + \lambda_1 h_{i1} + \lambda_2 h_{i2} + \lambda_3 h_{i3} + \lambda_4 h_{i1} h_{i2} + \lambda_5 h_{i1} h_{i3} + \lambda_6 h_{i2} h_{i3} + \lambda_7 h_{i1} h_{i2} h_{i3} + \lambda_8 f_{i1} \\
& + \lambda_9 f_{i2} + \lambda_{10} f_{i3} + \lambda_{11} h_{i1} f_{i1} + \lambda_{12} h_{i2} f_{i1} + \lambda_{13} h_{i3} f_{i1} + \lambda_{14} h_{i1} h_{i2} f_{i1} + \lambda_{15} h_{i1} h_{i3} f_{i1} \\
& + \lambda_{16} h_{i2} h_{i3} f_{i1} + \lambda_{17} h_{i1} h_{i2} h_{i3} f_{i1} + \lambda_{18} h_{i1} f_{i2} + \lambda_{19} h_{i2} f_{i2} + \lambda_{20} h_{i3} f_{i2} + \lambda_{21} h_{i1} h_{i2} f_{i2} \\
& + \lambda_{22} h_{i1} h_{i3} f_{i2} + \lambda_{23} h_{i2} h_{i3} f_{i2} + \lambda_{24} h_{i2} h_{i2} h_{i3} f_{i2} + \lambda_{25} h_{i1} f_{i3} + \lambda_{26} h_{i1} h_{i2} f_{i3} + \lambda_{27} h_{i3} f_{i3} \\
& + \lambda_{28} h_{i1} h_{i2} f_{i3} + \lambda_{29} h_{i3} f_{i3} + \lambda_{30} h_{i2} h_{i3} f_{i3} + \lambda_{31} h_{i2} h_{i2} h_{i3} f_{i3} + \nu_i. \tag{21}
\end{aligned}$$

Substituting the above into equation (20) we can write out the structural equations for each time period.<sup>2</sup> From these equations we can estimate the following reduced form equations:

$$\begin{aligned}
y_{i1} = & \delta_1 + \gamma_1 h_{i1} + \gamma_2 h_{i2} + \gamma_3 h_{i3} + \gamma_4 h_{i1} h_{i2} + \gamma_5 h_{i1} h_{i3} + \gamma_6 h_{i2} h_{i3} + \gamma_7 h_{i1} h_{i2} h_{i3} \\
& + \gamma_8 f_{i1} + \gamma_9 f_{i2} + \gamma_{10} f_{i3} + \gamma_{11} h_{i1} f_{i1} + \gamma_{12} h_{i2} f_{i1} + \gamma_{13} h_{i3} f_{i1} + \gamma_{14} h_{i1} h_{i2} f_{i1} \\
& + \gamma_{15} h_{i1} h_{i3} f_{i1} + \gamma_{16} h_{i2} h_{i3} f_{i1} + \gamma_{17} h_{i1} h_{i2} h_{i3} f_{i1} + \gamma_{18} h_{i1} f_{i2} + \gamma_{19} h_{i2} f_{i2} \\
& + \gamma_{20} h_{i3} f_{i2} + \gamma_{21} h_{i1} h_{i2} f_{i2} + \gamma_{22} h_{i1} h_{i3} f_{i2} + \gamma_{23} h_{i2} h_{i3} f_{i2} + \gamma_{24} h_{i1} h_{i2} h_{i3} f_{i2} \\
& + \gamma_{25} h_{i1} f_{i3} + \gamma_{26} h_{i2} f_{i3} + \gamma_{27} h_{i3} f_{i3} + \gamma_{28} h_{i1} h_{i2} f_{i3} + \gamma_{29} h_{i1} h_{i3} f_{i3} \\
& + \gamma_{30} h_{i2} h_{i3} f_{i3} + \gamma_{31} h_{i1} h_{i2} h_{i3} f_{i3} + (\nu_i + \phi \nu_i h_{i1} + u_{i1}), \tag{22}
\end{aligned}$$

$$\begin{aligned}
y_{i2} = & (\delta + \lambda_0) + \gamma_{32} h_{i1} + \gamma_{33} h_{i2} + \gamma_{34} h_{i3} + \gamma_{35} h_{i1} h_{i2} + \gamma_{36} h_{i1} h_{i3} + \gamma_{37} h_{i2} h_{i3} \\
& + \gamma_{38} h_{i1} h_{i2} h_{i3} + \gamma_{39} f_{i1} + \gamma_{40} f_{i2} + \gamma_{41} f_{i3} + \gamma_{42} h_{i1} f_{i1} + \gamma_{43} h_{i2} f_{i1} + \gamma_{44} h_{i3} f_{i1} \\
& + \gamma_{45} h_{i1} h_{i2} f_{i1} + \gamma_{46} h_{i1} h_{i3} f_{i1} + \gamma_{47} h_{i2} h_{i3} f_{i1} + \gamma_{48} h_{i1} h_{i2} h_{i3} f_{i1} + \gamma_{49} h_{i1} f_{i2} \\
& + \gamma_{50} h_{i2} f_{i2} + \gamma_{51} h_{i3} f_{i2} + \gamma_{52} h_{i1} h_{i2} f_{i2} + \gamma_{53} h_{i1} h_{i3} f_{i2} + \gamma_{54} h_{i2} h_{i3} f_{i2} \\
& + \gamma_{55} h_{i1} h_{i2} h_{i3} f_{i2} + \gamma_{56} h_{i1} f_{i3} + \gamma_{57} h_{i2} f_{i3} + \gamma_{58} h_{i3} f_{i3} + \gamma_{59} h_{i1} h_{i2} f_{i3} \\
& + \gamma_{60} h_{i1} h_{i3} f_{i3} + \gamma_{61} h_{i2} h_{i3} f_{i3} + \gamma_{62} h_{i1} h_{i2} h_{i3} f_{i3} + (\nu_i + \phi \nu_i h_{i2} + u_{i2}), \tag{23}
\end{aligned}$$

$$\begin{aligned}
y_{i3} = & (\delta + \lambda_0) + \gamma_{63} h_{i1} + \gamma_{64} h_{i2} + \gamma_{65} h_{i3} + \gamma_{66} h_{i1} h_{i2} + \gamma_{67} h_{i1} h_{i3} + \gamma_{68} h_{i2} h_{i3} \\
& + \gamma_{69} h_{i1} h_{i2} h_{i3} + \gamma_{70} f_{i1} + \gamma_{71} f_{i2} + \gamma_{72} f_{i3} + \gamma_{73} h_{i1} f_{i1} + \gamma_{74} h_{i2} f_{i1} + \gamma_{75} h_{i3} f_{i1} \\
& + \gamma_{76} h_{i1} h_{i2} f_{i1} + \gamma_{77} h_{i1} h_{i3} f_{i1} + \gamma_{78} h_{i2} h_{i3} f_{i1} + \gamma_{79} h_{i1} h_{i2} h_{i3} f_{i1} + \gamma_{80} h_{i1} f_{i2} \\
& + \gamma_{81} h_{i2} f_{i2} + \gamma_{82} h_{i3} f_{i2} + \gamma_{83} h_{i1} h_{i2} f_{i2} + \gamma_{84} h_{i1} h_{i3} f_{i2} + \gamma_{85} h_{i2} h_{i3} f_{i2} \\
& + \gamma_{86} h_{i1} h_{i2} h_{i3} f_{i2} + \gamma_{87} h_{i1} f_{i3} + \gamma_{88} h_{i2} f_{i3} + \gamma_{89} h_{i3} f_{i3} + \lambda_{90} h_{i1} h_{i2} f_{i3} \\
& + \gamma_{91} h_{i1} h_{i3} f_{i3} + \gamma_{92} h_{i2} h_{i3} f_{i3} + \gamma_{93} h_{i1} h_{i2} h_{i3} f_{i3} + (\nu_i + \phi \nu_i h_{i3} + u_{i3}). \tag{24}
\end{aligned}$$

2. Suri (2011) argues for the need to include the full history of adoption in the projection to completely control for its endogeneity and thus ensure that  $E[v_i h_{it} h_{it}] = 0$ . Given this fact, we are not entirely sure why she does not include all possible interactions of the additional endogenous variable,  $f_{it}$ . Since the `randcoef` package is designed to implement the estimation procedure as laid out in Suri (2011), we follow her definition of the projection as in equation (21).

These equations give 93 reduced form coefficients ( $\gamma_1 - \gamma_{93}$ ) from which we can estimate 34 structural parameters ( $\lambda_1 - \lambda_{34}, \rho, \beta, \phi$ ). The restrictions are:

$$\begin{array}{lll}
\gamma_1 = [\lambda_1(1 + \phi) + \beta + \phi\lambda_0] & \gamma_{32} = \lambda_1 & \gamma_{63} = \lambda_1 \\
\gamma_2 = \lambda_2 & \gamma_{33} = [\lambda_2(1 + \phi) + \beta + \phi\lambda_0] & \gamma_{64} = \lambda_2 \\
\gamma_3 = \lambda_3 & \gamma_{34} = \lambda_3 & \gamma_{65} = [\lambda_3(1 + \phi) + \beta + \phi\lambda_0] \\
\gamma_4 = [\lambda_4(1 + \phi) + \phi\lambda_2] & \gamma_{35} = [\lambda_4(1 + \phi) + \phi\lambda_1] & \gamma_{66} = \lambda_4 \\
\gamma_5 = [\lambda_5(1 + \phi) + \phi\lambda_3] & \gamma_{36} = \lambda_5 & \gamma_{67} = [\lambda_5(1 + \phi) + \phi\lambda_1] \\
\gamma_6 = \lambda_6 & \gamma_{37} = [\lambda_6(1 + \phi) + \phi\lambda_3] & \gamma_{68} = [\lambda_6(1 + \phi) + \phi\lambda_2] \\
\gamma_7 = [\lambda_7(1 + \phi) + \phi\lambda_6] & \gamma_{38} = [\lambda_7(1 + \phi) + \phi\lambda_5] & \gamma_{69} = [\lambda_7(1 + \phi) + \phi\lambda_4] \\
\gamma_8 = [\lambda_8 + \rho] & \gamma_{39} = \lambda_8 & \gamma_{70} = \lambda_8 \\
\gamma_9 = \lambda_9 & \gamma_{40} = [\lambda_9 + \rho] & \gamma_{71} = \lambda_9 \\
\gamma_{10} = \lambda_{10} & \gamma_{41} = \lambda_{10} & \gamma_{72} = [\lambda_{10} + \rho] \\
\gamma_{11} = [\lambda_{11}(1 + \phi) + \phi\lambda_8] & \gamma_{42} = \lambda_{11} & \gamma_{73} = \lambda_{11} \\
\gamma_{12} = \lambda_{12} & \gamma_{43} = [\lambda_{12}(1 + \phi) + \phi\lambda_8] & \gamma_{74} = \lambda_{12} \\
\gamma_{13} = \lambda_{13} & \gamma_{44} = \lambda_{13} & \gamma_{75} = [\lambda_{13}(1 + \phi) + \phi\lambda_8] \\
\gamma_{14} = [\lambda_{14}(1 + \phi) + \phi\lambda_{12}] & \gamma_{45} = [\lambda_{14}(1 + \phi) + \phi\lambda_{11}] & \gamma_{76} = \lambda_{14} \\
\gamma_{15} = [\lambda_{15}(1 + \phi) + \phi\lambda_{13}] & \gamma_{46} = \lambda_{15} & \gamma_{77} = [\lambda_{15}(1 + \phi) + \phi\lambda_{11}] \\
\gamma_{16} = \lambda_{16} & \gamma_{47} = [\lambda_{16}(1 + \phi) + \phi\lambda_{13}] & \gamma_{78} = [\lambda_{16}(1 + \phi) + \phi\lambda_{12}] \\
\gamma_{17} = [\lambda_{17}(1 + \phi) + \phi\lambda_{16}] & \gamma_{48} = [\lambda_{17}(1 + \phi) + \phi\lambda_{15}] & \gamma_{79} = [\lambda_{17}(1 + \phi) + \phi\lambda_{14}] \\
\gamma_{18} = [\lambda_{18}(1 + \phi) + \phi\lambda_9] & \gamma_{49} = \lambda_{18} & \gamma_{80} = \lambda_{18} \\
\gamma_{19} = \lambda_{19} & \gamma_{50} = [\lambda_{19}(1 + \phi) + \phi\lambda_9] & \gamma_{81} = \lambda_{19} \\
\gamma_{20} = \lambda_{20} & \gamma_{51} = \lambda_{20} & \gamma_{82} = [\lambda_{20}(1 + \phi) + \phi\lambda_9] \\
\gamma_{21} = [\lambda_{21}(1 + \phi) + \phi\lambda_{23}] & \gamma_{52} = [\lambda_{21}(1 + \phi) + \phi\lambda_{18}] & \gamma_{83} = \lambda_{21} \\
\gamma_{22} = [\lambda_{22}(1 + \phi) + \phi\lambda_{20}] & \gamma_{53} = \lambda_{22} & \gamma_{84} = [\lambda_{22}(1 + \phi) + \phi\lambda_{18}] \\
\gamma_{23} = \lambda_{23} & \gamma_{54} = [\lambda_{23}(1 + \phi) + \phi\lambda_{20}] & \gamma_{85} = [\lambda_{23}(1 + \phi) + \phi\lambda_{19}] \\
\gamma_{24} = [\lambda_{24}(1 + \phi) + \phi\lambda_{23}] & \gamma_{55} = [\lambda_{24}(1 + \phi) + \phi\lambda_{22}] & \gamma_{86} = [\lambda_{24}(1 + \phi) + \phi\lambda_{21}] \\
\gamma_{25} = [\lambda_{25}(1 + \phi) + \phi\lambda_{10}] & \gamma_{56} = \lambda_{25} & \gamma_{87} = \lambda_{25} \\
\gamma_{26} = \lambda_{26} & \gamma_{57} = [\lambda_{26}(1 + \phi) + \phi\lambda_{10}] & \gamma_{88} = \lambda_{26} \\
\gamma_{27} = \lambda_{27} & \gamma_{58} = \lambda_{27} & \gamma_{89} = [\lambda_{27}(1 + \phi) + \phi\lambda_{10}] \\
\gamma_{28} = [\lambda_{28}(1 + \phi) + \phi\lambda_{26}] & \gamma_{59} = [\lambda_{28}(1 + \phi) + \phi\lambda_{25}] & \gamma_{90} = \lambda_{28} \\
\gamma_{29} = [\lambda_{29}(1 + \phi) + \phi\lambda_{27}] & \gamma_{60} = \lambda_{29} & \gamma_{91} = [\lambda_{29}(1 + \phi) + \phi\lambda_{25}] \\
\gamma_{30} = \lambda_{30} & \gamma_{61} = [\lambda_{30}(1 + \phi) + \phi\lambda_{27}] & \gamma_{92} = [\lambda_{30}(1 + \phi) + \phi\lambda_{26}] \\
\gamma_{31} = [\lambda_{31}(1 + \phi) + \phi\lambda_{30}] & \gamma_{62} = [\lambda_{31}(1 + \phi) + \phi\lambda_{29}] & \gamma_{93} = [\lambda_{31}(1 + \phi) + \phi\lambda_{28}]
\end{array}$$

As before, we estimate equations (22) - (24) and preserve the 93 reduced form parameters in a vector  $\pi_{[93 \times 1]}$ , the variance-covariance matrices in a large symmetric block matrix  $\mathbf{V}_{[93 \times 93]}$ , the restrictions as  $\mathbf{H}_{[93 \times 34]}$ , and the structural parameters as  $\delta_{[34 \times 1]}$ . To calculate standard errors we take derivatives of the elements in  $\mathbf{H}\delta$  with respect to each of the structural parameters.

As one can see, the complexity of the problem grows exponentially. The addition of each new period requires the estimation of only one additional structural parameter in the CRE model but  $2^T - 1$  parameters in the CRC models. Table 1 summarizes the number of structural and reduced form parameters that must be estimated in each



model for up to five time periods. We believe that it is this computational intensity that has limited the usefulness of Suri’s approach to estimating CRC models. Of the 320+ papers that cited Suri (2011) as of 2017, only one (a working paper) has actually implemented her method of estimation. The `randcoef` command is designed to lower this hurdle and allow for a broader application of the estimation procedure.

Table 1: Number of parameters estimated in each model

Period	Parameter	CRE	CRC	CRC Endogenous
2	Structural	$2 + 1 = 3$	$3 + 2 = 5$	$(3 \times 3 + 2) + 3 = 14$
	Reduced	$2 \times 2 = 4$	$3 \times 2 = 6$	$11 \times 2 = 22$
3	Structural	$3 + 1 = 4$	$7 + 2 = 9$	$(7 \times 4 + 3) + 3 = 34$
	Reduced	$3 \times 3 = 9$	$7 \times 3 = 21$	$31 \times 3 = 93$
4	Structural	$4 + 1 = 5$	$15 + 2 = 17$	$(15 \times 5 + 4) + 3 = 82$
	Reduced	$4 \times 4 = 16$	$15 \times 4 = 60$	$79 \times 4 = 316$
5	Structural	$5 + 1 = 6$	$31 + 2 = 33$	$(31 \times 6 + 5) + 3 = 194$
	Reduced	$5 \times 5 = 25$	$31 \times 5 = 155$	$191 \times 5 = 955$

### 3 Data Considerations

Including the full history of adoption (the adoption decision in each year plus all interactions) in the projection of  $\theta_i$  allows for the assumption that there is no correlation between  $\nu_i$  and the decision to adopt (Chamberlain 1984; Suri 2011). However, identification of  $\phi$  in Suri’s CRC model when adoption is binary requires that all possible adoption histories are observed in the data. If this is not the case, then the model will suffer from multicollinearity. This is both a result of how Suri constructed the projection of  $\theta_i$  and of measuring adoption as a binary variable. For relatively small data sets, or as the number of time periods in a panel grows, there will be a higher probability that one of the adoption histories is not observed in the data. This is not an estimation or coding problem in the command, but rather an inherent limitation of the method.

To illustrate this, assume we have two time periods in the panel. The projection of  $\theta_i$  onto adoption will be:

$$\theta_i = \lambda_0 + \lambda_1 h_{i1} + \lambda_2 h_{i2} + \lambda_3 h_{i1} h_{i2} + \nu_i. \quad (25)$$

In this case, there will be four potential adoption histories: 1) those that never adopt (Never-Adopters), 2) those that always adopt (Always-Adopters), 3) those that adopted in period two, but not in period one (Adopters), and 4) those that adopted in period one, but dis-adopted in period two (Dis-adopters). If adoption corresponds to a 1 and non-adoption a 0, then each history corresponds to a particular tuple of choices. For the two-period case, Always-Adopters would be represented by the tuple (1, 1), Never-Adopters represented by the tuple (0, 0), Adopters by the tuple (0, 1), and Dis-adopters

by the tuple (1,0).

Table 2 summarizes the histories as they correspond to the variables in the projection. Note that the use of the intercept ( $\lambda_0$ ) implicitly means that the Never-Adopters are captured by this intercept and so are not included in the table.

Table 2: Adoption histories

	$h_{i1}$	$h_{i2}$	$h_{i1}h_{i2}$
(1,1)	1	1	1
(1,0)	1	0	0
(0,1)	0	1	0

From Table 2, we can see that if any one of the histories is not observed in the data (equivalent to a row disappearing from the table), then at least two of the independent variables becomes collinear by construction. For instance, if there were no Adopters in the data, then the last row would be missing. In this case  $h_{i2}$  and  $h_{i1}h_{i2}$  would now be perfectly collinear (the highlighted cells in the table). As such, estimation of Suri's CRC model requires a large enough dataset to ensure all adoption histories are present or, alternatively, must measure adoption as a continuous.

## 4 Fitting the CRE and CRC models using Stata

### 4.1 Syntax

The `randcoef` command estimates both the CRE (default) and the CRC models.<sup>3</sup> The command follows the same general syntax as the `sureg` command. However, it does not allow for equation naming since the independent variables are the same in all regressions. The command's syntax requires the order of both the outcome and choice variables to be listed chronologically. The command uses the number of dependent variables to know how many time periods the CRE or CRC model contains and then generates the necessary interaction terms and calculates the appropriate restriction matrix. If the user wants to change the order of the variables, they can use the `matrix` option accordingly<sup>4</sup>

```
randcoef (depvar1 depvar2 depvar3 ...) [if], choice(indepvar1 indepvar2  
indepvar3 ...) endo(endovar1 endovar2 endovar3 ...) [options]
```

3. Note that `randcoef` requires that the `tuple` package be installed.

4. Applies only for the CRE method

## 4.2 Options

**choice**(*indepvar1 indepvar2 indepvar3 ...*) specifies the variables of interest organized in the same order as the dependent variables. In the case of the CRC model, interactions should not be included since they are created automatically by the program. This option is required.

**method**(*string*) determines the method to be used: **cre** (*default*) or **crc**.

**controls**(*varlist*) allows for the addition of exogenous covariates as controls to the underlying **sureg** regression.

**showreg** prints the output of the SUR regression.

**matrix** allows the user to specify the restriction matrix (for changing the order of coefficients). Note that the restriction's order is important since it must match the order in which the parameters of interest are inputted in the SUR regressions. This option is only allowed if the CRE method is specified. The restriction matrix for the CRC method is preprogrammed automatically following that outlined in Suri (2011) depending on the number of dependent variables.

**endogenous**(*varlist*) adds the ability to allow for one additional endogenous variable to be added to the estimation. The interactions needed between the choice variables and the endogenous variable are automatically created.

**keep** allows the user to preserve the interactions between the choice variables and the endogenous ones (in the case the endogenous option is used) in the command line after the estimation has taken place. If specified, running the command several times requires dropping those variables.

**weighting**(*string*) specifies the weighting matrix to be used in estimation. This option is only allowed if the CRC method is specified.

**omd**(*default*) Optimal Minimum Distance: uses the inverse of the variance-covariance matrix of the reduced form estimates.

**ewmd** Equally Weighted Minimum Distance: uses the identity matrix.

**dwm** Diagonally Weighted Minimum Distance: uses the OMD matrix, but with zeros on the off-diagonals.

## 5 Interpreting Output

As an example of estimating and interpreting the CRC model we use household-level adoption of improved chickpea in Ethiopia.<sup>5</sup> The dependent variables are income per capita in each of the three survey rounds (**lnp08**, **lnp10**, **lnp14**) while the choice variables are adoption of improved chickpea in each period (**icp08**, **icp10**, **icp14**).

---

5. For a description of the data see Verkaart et al. (2017).

We also include a set of exogenous covariates as controls (`'controls1' ...`).<sup>6</sup> Note that panels in a “long” format must be reshaped as “wide” prior to estimation. Output from the CRC method using the OMD estimator is:

```
. randcoef lnp08 lnp10 lnp14, choice(icp08 icp10 icp14)
> controls(`controls1' `controls2' `controls3') meth(CRC) keep
  RUNNING MODEL WITH OMD WEIGHTING MATRIX
  Equations used in sureg:lnp08 lnp10 lnp14 = icp08 icp10 icp14
  > __00000C __00000D __00000E __00000F `controls1' `controls2' `controls3'
```

The model used is : method: CRC

The variables of interest are : icp08 icp10 icp14 \_\_00000C \_\_00000D \_\_00000E  
> \_\_00000F `controls1' `controls2' `controls3'

Minimum Distance Estimator is being calculated  
(output omitted)

With corresponding Parameters matrix:

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
11	.2241552	.1797628	1.25	0.212	-.1281735	.5764839
12	.3692923	.1521381	2.43	0.015	.0711071	.6674774
13	.164945	.1257764	1.31	0.190	-.0815722	.4114622
14	-.97277	.575561	-1.69	0.091	-2.100849	.1553089
15	.0193329	.2340247	0.08	0.934	-.4393471	.4780129
16	-.2653933	.1585105	-1.67	0.094	-.5760682	.0452815
17	.8201766	.6114319	1.34	0.180	-.3782078	2.018561
b	.4408669	.235167	1.87	0.061	-.020052	.9017859
phi	-.2547399	.6437777	-0.40	0.692	-1.516521	1.007041

In this case, the interactions for the choice variable are automatically created. Given that `keep` is used, these interactions are available after the estimation for further analysis. Note that the magnitudes of the structural parameters from equation (16) are:  $\lambda_1 = 0.224$ ,  $\lambda_2 = 0.369$ ,  $\lambda_3 = 0.165$ ,  $\lambda_4 = -0.973$ ,  $\lambda_5 = 0.019$ ,  $\lambda_6 = -0.265$ , and  $\lambda_7 = 0.820$ . The aggregate returns to improved chickpea adoption is  $\beta = 0.441$  while  $\phi = -0.255$  is the coefficient on the individual's rate of return or comparative advantage in adoption. In this example the aggregate returns to adoption are significant but the comparative advantage term is not. One would interpret this as a lack of significant differences in comparative advantage in this economy. Returns to improved chickpea adoption, at least in this sample, are homogeneous not heterogeneous.

Despite a lack of significant differences in the returns to adoption based on household unobservables, we can still explore heterogeneity within the population by predicting the  $\hat{\theta}$  term for a given adoption history. We can recover  $\hat{\theta}$  using equation (16) and our structural OMD estimates. Given that each history is binary, and given that we observe at least one household in each history, the projection is fully saturated. This procedure results in eight mass points for the  $\hat{\theta}$ s.

6. Because our interest lies in the coefficient estimates for the structural parameters, and in the interests of parsimony, we have only listed the locals for the control variables. A complete list of the control variables and variable definitions is contained in the accompanying data and .do file, which also allows for replication of results.

Once we have recovered the  $\hat{\theta}_s$ , we can predict the average returns for a given adoption history. This involves calculating  $\hat{\beta} + \hat{\phi}\hat{\theta}_i$ , where  $\hat{\beta}$  is the average return to improved varieties and each  $i$  is a specific adoption history. The results can be viewed as the counterfactual returns for non-adopting households using weighted averages of all possible returns. Again, since the histories are binary and the projection is fully saturated, the process produces eight mass points, which we graph in Figures 1.

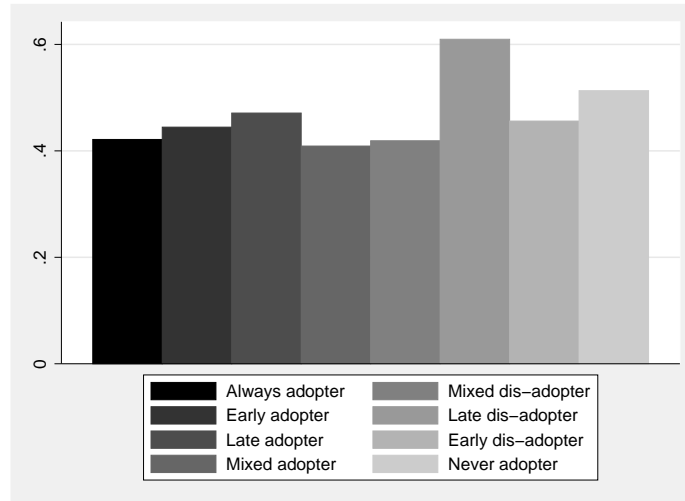


Figure 1: Distribution of returns to adoption

As the bars in Figure 1 make clear, returns to adoption are relatively homogeneous across adoption history. Since  $\phi$  is not significant, the weighted average across these groups would yield  $\beta$  or the aggregate returns to adoption. If  $\phi$  were significant we would expect noticeable differences across the bars as returns would differ significantly by adoption history. See Suri (2011) for graphs in which returns differ across groups.

## 6 Conclusions

Correlated random coefficient models can be used to understand a variety of activities in which the return on the activity to the individual is correlated with the individual's choice to participate. Several methods exist to consistently estimate these models but most rely on an instrumental variables approach. While the structural approach first developed in Suri (2006) has been around for a decade, its use has been limited by the computationally intensive nature of the estimation procedure. The `randcoef` package allows users to estimate both CRE and CRC models and provides them with a variety of estimation options. Additionally, the package allows for estimation using up to five rounds of data and can accommodate an additional endogenous regressor. The hope is that the package will lower the hurdle for implementing this approach and thereby add another tool for estimating CRC models.

## 7 References

- Chamberlain, G. 1984. Panel Data. In *Hanbook of Econometrics, Vol. II*, ed. Z. Griliches and M. Intriligator. Amsterdam: North-Holland.
- Heckman, J., and E. Vytlacil. 1998. Instrumental Variables Methods for the Correlated Random Coefficient Model: Estimating the Average Rate of Return to Schooling When the Return is Correlated with Schooling. *Journal of Human Resources* 33(4): 974–87.
- Mundlak, Y. 1978. On the Pooling of Time Series and Cross Section Data. *Econometrica* 46(1): 69–85.
- Suri, T. 2006. Selection and Comparative Advantage in Technology Adoption. Economic Growth Center Discussion Paper No. 944, Yale University.
- . 2011. Selection and Comparative Advantage in Technology Adoption. *Econometrica* 79(1): 159–209.
- Verkaart, S., B. G. Munyua, K. Mausch, and J. D. Michler. 2017. Welfare impacts of improved chickpea adoption: A pathway for rural development in Ethiopia? *Food Policy* 66(1): 50–61.
- Wooldridge, J. M. 2003. Further Results on Instrumental Variables Estimation of Average Treatment Effects in the Correlated Random Coefficient Model. *Economic Letters* 79: 185–191.

### About the authors

Oscar Barriga Cabanillas is a Ph.D. student in the Department of Agricultural and Resource Economics at the University of California, Davis.

Jeffrey D. Michler is a Postdoctoral Research Fellow in the Department of Agricultural and Consumer Economics at the University of Illinois at Urbana-Champaign.

Aleksandr Michuda is a Ph.D. student in the Department of Agricultural and Resource Economics at the University of California, Davis.

Emilia Tjernström is an Assistant Professor in the Robert M. La Follette School of Public Affairs and the Department of Agricultural and Applied Economics at the University of Wisconsin-Madison.