



The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

The Stata Journal (2018)
18, Number 1, pp. 159–173

Fitting and interpreting correlated random-coefficient models using Stata

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

Jeffrey D. Michler
University of Saskatchewan
Saskatoon, Canada
jeffrey.michler@usask.ca

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

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

Abstract. In this article, we introduce the community-contributed command `randcoef`, which fits the correlated random-effects and correlated random-coefficient models discussed in Suri (2011, *Econometrica* 79: 159–209). While this approach has been around for a decade, its use has been limited by the computationally intensive nature of the estimation procedure that relies on the optimal minimum distance estimator. `randcoef` can accommodate up to five rounds of panel data and offers several options, including alternative weight matrices for estimation and inclusion of additional endogenous regressors. We also present postestimation analysis using sample data to facilitate understanding and interpretation of results.

Keywords: st0517, `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 Vytlačil 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 Vytlačil [1998], maize yields in Suri [2011]) is a function of the rate of return, α_{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 ϵ_{0i} and ϵ_{1i} are zero in expectation. We can then rewrite (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 α_{1i} , the rate of return for the individual, is correlated with the individual's choice of h_i . In the case of schooling, this would

be the case if one's return to education influences one's choice to invest in education. In the case of technology adoption, the classic example is when one's return to adoption influences one's choice to adopt the new technology.

The literature has produced several alternative approaches to identifying the correlated random-coefficient (CRC) model in (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 (FE) 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 (OMD) estimator. `randcoef` 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.¹

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 defend the exclusion restriction. Rather, identification relies on the linear projection of an individual's rate of return onto his or her complete history of adoption, similar to the correlated random-effects (CRE) 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 postestimation 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 explains the math behind the calculations of the CRE and CRC models, as well as the CRC model with an additional endogenous covariate. Two-period versions of these models are developed in Suri (2011). To fix these ideas, we outline the calculations of a three-period model. 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, which do vary depending on the number of periods. Though we discuss only a three-period model, `randcoef` can accommodate anywhere between two and five periods of panel data.

1. In the case of the hybrid maize adoption story 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

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

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 (α_i), and an idiosyncratic error term $u_{it} \sim \mathcal{N}(0, \sigma_u^2)$. We assume strict exogeneity of the error term.

In an FE model, the unique individual indicators are simply dummy variables. In the CRE model, we replace the FE with 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_{it} for $t = 1, 2, 3$ is an indicator that equals 1 if individual i adopts at time t and the λ 's are the projection coefficients. By the nature of the projection, ν_i is uncorrelated with h_{it} (Suri 2011). Substituting (4) into (3), we get

$$y_{it} = \zeta + \beta h_{it} + \lambda_0 + \lambda_1 h_{i1} + \lambda_2 h_{i2} + \lambda_3 h_{i3} + \nu_i + u_{it}$$

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

$$\begin{aligned} y_{i1} &= (\zeta + \lambda_0) + (\beta + \lambda_1)h_{i1} + \lambda_2 h_{i2} + \lambda_3 h_{i3} + \epsilon_{i1} \\ y_{i2} &= (\zeta + \lambda_0) + \lambda_1 h_{i1} + (\beta + \lambda_2)h_{i2} + \lambda_3 h_{i3} + \epsilon_{i2} \\ y_{i3} &= (\zeta + \lambda_0) + \lambda_1 h_{i1} + \lambda_2 h_{i2} + (\beta + \lambda_3)h_{i3} + \epsilon_{i3} \end{aligned}$$

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

$$y_{i1} = \zeta_1 + \gamma_1 h_{i1} + \gamma_2 h_{i2} + \gamma_3 h_{i3} + n_{i1} \quad (5)$$

$$y_{i2} = \zeta_2 + \gamma_4 h_{i1} + \gamma_5 h_{i2} + \gamma_6 h_{i3} + n_{i2} \quad (6)$$

$$y_{i3} = \zeta_3 + \gamma_7 h_{i1} + \gamma_8 h_{i2} + \gamma_9 h_{i3} + n_{i3} \quad (7)$$

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 (5)–(7) as a set of seemingly unrelated regressions (SUR) 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 γ 's can be expressed as $\boldsymbol{\pi} = \mathbf{H}\boldsymbol{\delta}$, where $\mathbf{H}_{[9 \times 4]}$ is a restriction matrix embodying the nine restrictions on γ , and $\boldsymbol{\delta}_{[4 \times 1]}$ is a vector of our four structural parameters.

The OMD function is

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

Solving for $\boldsymbol{\delta}$, we get

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

which is the OMD estimator. The key to getting the OMD to produce the correct estimates requires ensuring that the restriction matrix is correctly specified. In the three-period CRE case, we have

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

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

Alternatively, one could replace the inverse of the variance–covariance matrix (\mathbf{V}^{-1}) with either the identity matrix (\mathbf{I}) to generate the EWMD estimator or the inverse of the variance–covariance matrix with all off-diagonals set to zero ($\text{diag}[\mathbf{V}^{-1}]$) to generate the DWMD estimator. As [Altonji and Segal \(1996\)](#) discuss, OMD may generate significant bias in small samples, leading some researchers to prefer EWMD. However, as [Blundell, Pistaferri, and Preston \(2008\)](#) point out, EWMD does not allow for heteroskedasticity, which can be addressed by using DWMD.

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 because 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

$$y_{it} = \zeta + \beta h_{it} + \theta_i + \phi \theta_i h_{it} + \tau_i + u_{it} \quad (8)$$

where ϕ is the coefficient on an individual's rate of return or comparative advantage in adoption (θ_i), τ_i is an individual's absolute advantage (equivalent to an FE), and all other terms are as previously defined. To estimate (8), we must eliminate the dependence of θ_i on h_{it} . To do this, [Suri \(2011\)](#) and [Chamberlain \(1984\)](#) replace θ_i with its linear projection onto the individual's full history of adoption:

$$\theta_i = \lambda_0 + \lambda_1 h_{i1} + \lambda_2 h_{i2} + \lambda_3 h_{i3} + \lambda_4 h_{i1} h_{i2} + \lambda_5 h_{i1} h_{i3} + \lambda_6 h_{i2} h_{i3} + \lambda_7 h_{i1} h_{i2} h_{i3} + \nu_i \quad (9)$$

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

Substituting (9) into (8) and writing out each time period's function gives

$$\begin{aligned} y_{i1} = & (\zeta + \lambda_0) + \{\beta + \phi\lambda_0 + \lambda_1(1 + \phi)\}h_{i1} + \lambda_2 h_{i2} + \lambda_3 h_{i3} + \{\phi\lambda_2 + \lambda_4(1 + \phi)\}h_{i1} h_{i2} \\ & + \{\phi\lambda_3 + \lambda_5(1 + \phi)\}h_{i1} h_{i3} + \lambda_6 h_{i2} h_{i3} + \{\phi\lambda_6 + \lambda_7(1 + \phi)\}h_{i1} h_{i2} h_{i3} + (\nu_i \\ & + \phi\nu_i h_{i1} + u_{i1}) \end{aligned}$$

$$\begin{aligned} y_{i2} = & (\zeta + \lambda_0) + \lambda_1 h_{i1} + \{\beta + \phi\lambda_0 + \lambda_2(1 + \phi)\}h_{i2} + \lambda_3 h_{i3} + \{\phi\lambda_1 + \lambda_4(1 + \phi)\}h_{i1} h_{i2} \\ & + \lambda_5 h_{i1} h_{i3} + \{\phi\lambda_3 + \lambda_6(1 + \phi)\}h_{i2} h_{i3} + \{\phi\lambda_5 + \lambda_7(1 + \phi)\}h_{i1} h_{i2} h_{i3} + (\nu_i \\ & + \phi\nu_i h_{i2} + u_{i2}) \end{aligned}$$

$$\begin{aligned} y_{i3} = & (\zeta + \lambda_0) + \lambda_1 h_{i1} + \lambda_2 h_{i2} + \{\beta + \phi\lambda_0 + \lambda_3(1 + \phi)\}h_{i3} + \{\phi\lambda_1 + \lambda_5(1 + \phi)\}h_{i1} h_{i3} \\ & + \lambda_4 h_{i1} h_{i2} + \{\phi\lambda_2 + \lambda_6(1 + \phi)\}h_{i2} h_{i3} + \{\phi\lambda_4 + \lambda_7(1 + \phi)\}h_{i1} h_{i2} h_{i3} \\ & + (\nu_i + \phi\nu_i h_{i3} + u_{i3}) \end{aligned}$$

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

$$\begin{aligned} y_{i1} = & \zeta_1 + \gamma_1 h_{i1} + \gamma_2 h_{i2} + \gamma_3 h_{i3} + \gamma_4 h_{i1} h_{i2} + \gamma_5 h_{i1} h_{i3} + \gamma_6 h_{i2} h_{i3} \\ & + \gamma_7 h_{i1} h_{i2} h_{i3} + n_{i1} \end{aligned} \quad (10)$$

$$\begin{aligned} y_{i2} = & \zeta_2 + \gamma_8 h_{i1} + \gamma_9 h_{i2} + \gamma_{10} h_{i3} + \gamma_{11} h_{i1} h_{i2} + \gamma_{12} h_{i1} h_{i3} + \gamma_{13} h_{i2} h_{i3} \\ & + \gamma_{14} h_{i1} h_{i2} h_{i3} + n_{i2} \end{aligned} \quad (11)$$

$$\begin{aligned} y_{i3} = & \zeta_3 + \gamma_{15} h_{i1} + \gamma_{16} h_{i2} + \gamma_{17} h_{i3} + \gamma_{18} h_{i1} h_{i2} + \gamma_{19} h_{i1} h_{i3} + \gamma_{20} h_{i2} h_{i3} \\ & + \gamma_{21} h_{i1} h_{i2} h_{i3} + n_{i3} \end{aligned} \quad (12)$$

These equations give 21 reduced-form coefficients ($\gamma_1 - \gamma_{21}$), from which we can estimate 10 structural parameters ($\beta, \phi, \lambda_0 - \lambda_7$). Note that if we normalize the θ 's so that $\sum \theta_i = 0$, we can effectively eliminate λ_0 and must estimate only nine structural parameters.

The necessary restrictions to identify the structural parameters are

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

Similar to before, we estimate (10)–(12) as SUR and preserve the 21 reduced-form parameters in a vector $\boldsymbol{\pi}_{[21 \times 1]}$, 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 $\boldsymbol{\delta}_{[9 \times 1]}$.

We now must take derivatives of the elements in $\mathbf{H}\boldsymbol{\delta}$ 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 account not only for the whole history of adoption but also for those histories interacted with the endogenous variable. We start by including the endogenous covariate, f_{it} , in the data-generating process:

$$y_{it} = \zeta + \beta h_{it} + \rho f_{it} + \theta_i + \phi \theta_i h_{it} + \tau_i + u_{it} \quad (13)$$

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

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

Substituting the above into (13), we can write out the structural equations for each time period.² From these equations, we can estimate the following reduced-form equations:

$$\begin{aligned}
 y_{i1} = & \zeta_1 + \gamma_1 h_{i1} + \gamma_2 h_{i2} + \gamma_3 h_{i3} + \gamma_4 h_{i1} h_{i2} + \gamma_5 h_{i1} h_{i3} + \gamma_6 h_{i2} h_{i3} + \gamma_7 h_{i1} h_{i2} h_{i3} \\
 & + \gamma_8 f_{i1} + \gamma_9 f_{i2} + \gamma_{10} f_{i3} + \gamma_{11} h_{i1} f_{i1} + \gamma_{12} h_{i2} f_{i1} + \gamma_{13} h_{i3} f_{i1} + \gamma_{14} h_{i1} h_{i2} f_{i1} \\
 & + \gamma_{15} h_{i1} h_{i3} f_{i1} + \gamma_{16} h_{i2} h_{i3} f_{i1} + \gamma_{17} h_{i1} h_{i2} h_{i3} f_{i1} + \gamma_{18} h_{i1} f_{i2} + \gamma_{19} h_{i2} f_{i2} \\
 & + \gamma_{20} h_{i3} f_{i2} + \gamma_{21} h_{i1} h_{i2} f_{i2} + \gamma_{22} h_{i1} h_{i3} f_{i2} + \gamma_{23} h_{i2} h_{i3} f_{i2} + \gamma_{24} h_{i1} h_{i2} h_{i3} f_{i2} \\
 & + \gamma_{25} h_{i1} f_{i3} + \gamma_{26} h_{i2} f_{i3} + \gamma_{27} h_{i3} f_{i3} + \gamma_{28} h_{i1} h_{i2} f_{i3} + \gamma_{29} h_{i1} h_{i3} f_{i3} \\
 & + \gamma_{30} h_{i2} h_{i3} f_{i3} + \gamma_{31} h_{i1} h_{i2} h_{i3} f_{i3} + (\nu_i + \phi \nu_i h_{i1} + u_{i1}) \tag{14}
 \end{aligned}$$

$$\begin{aligned}
 y_{i2} = & (\zeta + \lambda_0) + \gamma_{32} h_{i1} + \gamma_{33} h_{i2} + \gamma_{34} h_{i3} + \gamma_{35} h_{i1} h_{i2} + \gamma_{36} h_{i1} h_{i3} + \gamma_{37} h_{i2} h_{i3} \\
 & + \gamma_{38} h_{i1} h_{i2} h_{i3} + \gamma_{39} f_{i1} + \gamma_{40} f_{i2} + \gamma_{41} f_{i3} + \gamma_{42} h_{i1} f_{i1} + \gamma_{43} h_{i2} f_{i1} + \gamma_{44} h_{i3} f_{i1} \\
 & + \gamma_{45} h_{i1} h_{i2} f_{i1} + \gamma_{46} h_{i1} h_{i3} f_{i1} + \gamma_{47} h_{i2} h_{i3} f_{i1} + \gamma_{48} h_{i1} h_{i2} h_{i3} f_{i1} + \gamma_{49} h_{i1} f_{i2} \\
 & + \gamma_{50} h_{i2} f_{i2} + \gamma_{51} h_{i3} f_{i2} + \gamma_{52} h_{i1} h_{i2} f_{i2} + \gamma_{53} h_{i1} h_{i3} f_{i2} + \gamma_{54} h_{i2} h_{i3} f_{i2} \\
 & + \gamma_{55} h_{i1} h_{i2} h_{i3} f_{i2} + \gamma_{56} h_{i1} f_{i3} + \gamma_{57} h_{i2} f_{i3} + \gamma_{58} h_{i3} f_{i3} + \gamma_{59} h_{i1} h_{i2} f_{i3} \\
 & + \gamma_{60} h_{i1} h_{i3} f_{i3} + \gamma_{61} h_{i2} h_{i3} f_{i3} + \gamma_{62} h_{i1} h_{i2} h_{i3} f_{i3} + (\nu_i + \phi \nu_i h_{i2} + u_{i2}) \tag{15}
 \end{aligned}$$

$$\begin{aligned}
 y_{i3} = & (\zeta + \lambda_0) + \gamma_{63} h_{i1} + \gamma_{64} h_{i2} + \gamma_{65} h_{i3} + \gamma_{66} h_{i1} h_{i2} + \gamma_{67} h_{i1} h_{i3} + \gamma_{68} h_{i2} h_{i3} \\
 & + \gamma_{69} h_{i1} h_{i2} h_{i3} + \gamma_{70} f_{i1} + \gamma_{71} f_{i2} + \gamma_{72} f_{i3} + \gamma_{73} h_{i1} f_{i1} + \gamma_{74} h_{i2} f_{i1} + \gamma_{75} h_{i3} f_{i1} \\
 & + \gamma_{76} h_{i1} h_{i2} f_{i1} + \gamma_{77} h_{i1} h_{i3} f_{i1} + \gamma_{78} h_{i2} h_{i3} f_{i1} + \gamma_{79} h_{i1} h_{i2} h_{i3} f_{i1} + \gamma_{80} h_{i1} f_{i2} \\
 & + \gamma_{81} h_{i2} f_{i2} + \gamma_{82} h_{i3} f_{i2} + \gamma_{83} h_{i1} h_{i2} f_{i2} + \gamma_{84} h_{i1} h_{i3} f_{i2} + \gamma_{85} h_{i2} h_{i3} f_{i2} \\
 & + \gamma_{86} h_{i1} h_{i2} h_{i3} f_{i2} + \gamma_{87} h_{i1} f_{i3} + \gamma_{88} h_{i2} f_{i3} + \gamma_{89} h_{i3} f_{i3} + \lambda_{90} h_{i1} h_{i2} f_{i3} \\
 & + \gamma_{91} h_{i1} h_{i3} f_{i3} + \gamma_{92} h_{i2} h_{i3} f_{i3} + \gamma_{93} h_{i1} h_{i2} h_{i3} f_{i3} + (\nu_i + \phi \nu_i h_{i3} + u_{i3}) \tag{16}
 \end{aligned}$$

2. Note that the logic of identification of the second endogenous variable does not precisely follow the logic of identification of the adoption decision. In [Suri \(2011\)](#), identification of the adoption decision requires controlling for the full history of adoption, including all of its own interaction terms. However, identification of the second endogenous variable relies only on the second variable's interaction with the full history of adoption and not the second variable's own interaction terms (that is, $f_{i1} f_{i2}$).

These equations give 93 reduced-form coefficients ($\gamma_1 - \gamma_{93}$), from which we can estimate 37 structural parameters ($\lambda_1 - \lambda_{34}, \rho, \beta, \phi$). The restrictions are the following:

$$\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 (14)–(16) and preserve the 93 reduced-form parameters in a vector $\boldsymbol{\pi}_{[93 \times 1]}$, the variance–covariance matrices in a large symmetric block matrix $\mathbf{V}_{[93 \times 93]}$, the restrictions as $\mathbf{H}_{[93 \times 37]}$, and the structural parameters as $\boldsymbol{\delta}_{[37 \times 1]}$. To calculate standard errors, we take derivatives of the elements in $\mathbf{H}\boldsymbol{\delta}$ 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 fitting CRC models. Of the 420-plus articles that cited Suri (2011) as of 2018, only 2 (both working articles) have 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 θ_i allows for the assumption that there is no correlation between ν_i and the decision to adopt (Chamberlain 1984; Suri 2011). However, identification of ϕ in Suri's CRC model when adoption is binary requires that all possible adoption histories be observed in the data. If this is not the case, then the model will suffer from multicollinearity. This is a result both of how Suri constructed the projection of θ_i and of measuring adoption as a binary variable. For relatively small datasets, 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, let's assume we have two time periods in the panel. The projection of θ_i onto adoption will be

$$\theta_i = \lambda_0 + \lambda_1 h_{i1} + \lambda_2 h_{i2} + \lambda_3 h_{i1} h_{i2} + \nu_i \quad (17)$$

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 disadopted in period two (disadopters). If adoption corresponds to a 1 and nonadoption to 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 by the tuple (0, 0), adopters by the tuple (0, 1), and disadopters 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 (λ_0) in (17) captures the never adopters, so they 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 continuous.

4 Fitting the CRE and CRC models using Stata

4.1 Description

The `randcoef` command fits the CRE (default) and CRC models.³ The command follows the same general syntax as the `sureg` command. However, it does not allow for equation naming, because 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. `randcoef` 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 users want to change the order of the variables, they can use the `matrix()` option accordingly.⁴

4.2 Syntax

The syntax to fit the CRE model is

```
randcoef (depvar1 depvar2 depvar3 ...) [if],
        choice(indepvar1 indepvar2 indepvar3 ...)
        [method(CRE) controls(varlist) showreg matrix(string)]
```

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

4. Applies only for the CRE method.

The syntax to fit the CRC model is

```
randcoef (depvar1 depvar2 depvar3 ...) [if],
        choice(indepvar1 indepvar2 indepvar3 ...)
        [method(CRC) controls(varlist) showreg endogenous(varlist) keep
         weighting(string)]
```

4.3 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, because they are created automatically by the program. These variables must be defined as dummies. `choice()` 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` allows the user to see the SUR output. By default, the program does not show the SUR output for faster computation and to avoid displaying unnecessary information.

`matrix(string)` allows the user to specify the restriction matrix (for changing the order of coefficients). Note that the restriction's order is important, because it must match the order in which the parameters of interest are input in the SUR. This option is allowed only 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)` allows for endogenous variables to be added to the CRC estimation. This must be a dummy variable. The interactions needed between the choice variables and the endogenous variables are automatically created.

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

`weighting(string)` specifies the weighting matrix to be used in estimation. This option is allowed only if the CRC method is specified. *string* may be one of the following:

`omd` (default) specifies OMD, which uses the inverse of the variance–covariance matrix of the reduced-form estimates.

`ewmd` specifies EWMD, which uses the identity matrix.

`dwmd` specifies DWMD, which uses the OMD matrix but with zeros on the off-diagonals.

5 Interpreting output

As an example of fitting and interpreting the CRC model, we use household-level adoption of improved chickpea in Ethiopia.⁵ 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`). We also include a set of exogenous covariates as controls (``controls1' ...`).⁶ 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') method(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'
Minimun Distance Estimator is being calculated
(output omitted)
With corresponding Parameters matrix:
```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
11	.1902277	.1583402	1.20	0.230	-.1201135	.5005689
12	.3902405	.1481763	2.63	0.008	.0998202	.6806608
13	.1208645	.1233118	0.98	0.327	-.1208222	.3625512
14	-.6114457	.3197189	-1.91	0.056	-1.238083	.0151918
15	-.0115502	.1943865	-0.06	0.953	-.3925407	.3694403
16	-.2721276	.14402	-1.89	0.059	-.5544015	.0101463
17	.4747094	.3116379	1.52	0.128	-.1360896	1.085508
b	.2158545	.0746901	2.89	0.004	.0694647	.3622443
phi	.9196741	1.500435	0.61	0.540	-2.021124	3.860472

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 (9) are $\lambda_1 = 0.190$, $\lambda_2 = 0.390$, $\lambda_3 = 0.121$, $\lambda_4 = -0.611$, $\lambda_5 = -0.012$, $\lambda_6 = -0.272$, and $\lambda_7 = 0.475$. The aggregate returns to improved chickpea adoption is $\beta = 0.216$, while $\phi = 0.920$ 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.

5. For a description of the data, see [Verkaart et al. \(2017\)](#).

6. Because our interest lies in the coefficient estimates for the structural parameters, and in the interests of parsimony, we have listed only 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.

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 (9) and our structural OMD estimates. This requires generating variables equal to the eight mass points for each of the $\hat{\theta}$'s.

```
. generate theta1 = 10 + _b[l1]*0 + _b[l2]*0 + _b[l3]*0 + _b[l4]*0 + _b[l5]*0
> + _b[l6]*0 + _b[l7]*0
. label variable theta1 "never adopt"
. generate theta2 = 10 + _b[l1]*1 + _b[l2]*1 + _b[l3]*1 + _b[l4]*1 + _b[l5]*1
> + _b[l6]*1 + _b[l7]*1
. label variable theta2 "always adopt"
. generate theta3 = 10 + _b[l1]*0 + _b[l2]*1 + _b[l3]*1 + _b[l4]*0 + _b[l5]*0
> + _b[l6]*1 + _b[l7]*0
. label variable theta3 "early adopters"
. generate theta4 = 10 + _b[l1]*0 + _b[l2]*0 + _b[l3]*1 + _b[l4]*0 + _b[l5]*0
> + _b[l6]*0 + _b[l7]*0
. label variable theta4 "late adopters"
. generate theta5 = 10 + _b[l1]*1 + _b[l2]*0 + _b[l3]*0 + _b[l4]*0 + _b[l5]*0
> + _b[l6]*0 + _b[l7]*0
. label variable theta5 "early disadopters"
. generate theta6 = 10 + _b[l1]*1 + _b[l2]*1 + _b[l3]*0 + _b[l4]*1 + _b[l5]*0
> + _b[l6]*0 + _b[l7]*0
. label variable theta6 "late disadopters"
. generate theta7 = 10 + _b[l1]*1 + _b[l2]*0 + _b[l3]*1 + _b[l4]*0 + _b[l5]*1
> + _b[l6]*0 + _b[l7]*0
. label variable theta7 "mixed adopters"
. generate theta8 = 10 + _b[l1]*0 + _b[l2]*1 + _b[l3]*0 + _b[l4]*0 + _b[l5]*0
> + _b[l6]*0 + _b[l7]*0
. label variable theta8 "mixed disadopters"
```

Given that each history is binary, and given that we observe at least one household in each history, the projection is fully saturated.

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, such as `generate r1 = _b[b] + _b[phi]*theta1`.

The results can be viewed as the counterfactual returns for nonadopting households using weighted averages of all possible returns. Again, because the histories are binary and the projection is fully saturated, the process produces eight mass points, which we graph in figure 1.⁷

7. Complete code for generating figure 1 is available as part of the Stata code associated with this article.

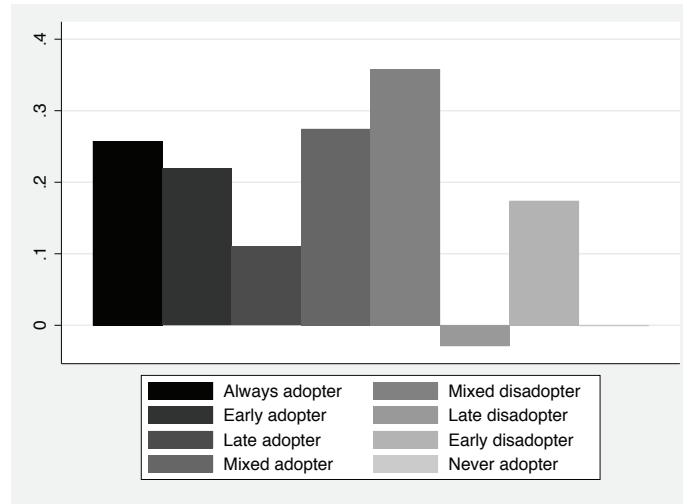


Figure 1. Distribution of returns to adoption

As the bars in figure 1 reveal, returns to adoption appear heterogeneous, at least for some adoption histories. What the coefficient on ϕ tells us is that ultimately these differences are not significant. Because ϕ is not significant, the weighted average across these groups would yield β or the aggregate returns to adoption. If ϕ were significant, we would expect larger differences across the bars or differences between groups based on adoption history. See Suri (2011) for graphs in which returns differ significantly and consistently 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 fit 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. `randcoef` allows users to fit both CRE and CRC models and provides them with a variety of estimation options. Additionally, it allows for estimation using up to five rounds of data and can accommodate an additional endogenous regressor. We hope `randcoef` will lower the hurdle for implementing this approach and thereby add another tool for fitting CRC models.

7 References

- Altonji, J. G., and L. M. Segal. 1996. Small-sample bias in GMM estimation of covariance structures. *Journal of Business and Economic Statistics* 14: 353–366.
- Blundell, R., L. Pistaferri, and I. Preston. 2008. Consumption inequality and partial insurance. *American Economic Review* 98: 1887–1921.
- Chamberlain, G. 1984. Panel data. In *Handbook of Econometrics*, vol. 2, ed. Z. Griliches and M. D. Intriligator, 1247–1318. Amsterdam: Elsevier.
- 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: 974–987.
- Mundlak, Y. 1978. On the pooling of time series and cross section data. *Econometrica* 46: 69–85.
- Suri, T. 2006. Selection and comparative advantage in technology adoption. Discussion Paper 944, Economic Growth Center, Yale University. http://www.econ.yale.edu/growth_pdf/cdp944.pdf.
- . 2011. Selection and comparative advantage in technology adoption. *Econometrica* 79: 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: 50–61.
- Wooldridge, J. M. 2003. Further results on instrumental variables estimation of average treatment effects in the correlated random coefficient model. *Economics Letters* 79: 185–191.

About the authors

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

Jeffrey D. Michler is an assistant professor in the Department of Agricultural and Resource Economics at the University of Saskatchewan.

Aleksandr Michuda is a PhD 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.