A Group Random Coefficient Approach to Modeling Heterogeneous Returns to Technology Adoption*

Emilia Tjernström[1], Dalia Ghanem[2], Oscar Barriga Cabanillas[2], Travis J. Lybbert[2], Jeffrey D. Michler[3], and Aleksandr Michuda[2]

[1] *University of Sydney*
[2] *University of California, Davis*
[3] *University of Arizona*

This version: July 2020

Abstract

Our paper revisits the econometric model that Suri (2011) (S2011) used in her study of heterogeneous returns to agricultural technology adoption. We propose an alternative group random coefficient (GRC) estimation strategy and revisit the empirical puzzle of why relatively few sub-Saharan farmers adopt modern technologies. Drawing on recent developments in the nonparametric panel identification literature, we start with an unrestricted GRC model that nonparametrically identifies the returns to adoption under time homogeneity. We show that the parameters of the S2011 correlated random coefficient model (CRC) can be identified from a restricted version of the GRC method. Specifically, the model in S2011 implies a key restriction that we call linearity in comparative advantage (LCA). Our unrestricted GRC model can be used to detect identification concerns for key structural parameters from the CRC model. We illustrate our method using the same data set as the original study and find that the motivating empirical puzzle remains unsolved.

JEL Classification: C12, C14, O33,Q16
Keywords: GMM, technology adoption, heterogeneity, correlated random coefficients

*We thank the Data Management Team at Tegemeo Institute of Agricultural Policy and Development for providing us with access to the data and Tavneet Suri for data access, code, and helpful comments.

1 Introduction

In an influential paper, Suri (2011) (hereinafter S2011) develops an alternative approach to identifying correlated random coefficient (CRC) models and uses the method to address a longstanding development puzzle: why do many sub-Saharan farmers continue to use traditional farming techniques when more modern, higher-return agricultural technologies are available? To the many explanations explored in the existing literature, S2011 adds one that hinges on a specific form of heterogeneity in the returns and costs associated with the technology. S2011 uses data from Kenyan farm households to estimate a CRC model that imposes a parametric relationship between farmers' absolute advantage (i.e., the component of productivity that is independent of technology) and their comparative advantage (i.e., their relative productivity with the hybrid technology over that without).

A key innovation in the S2011 approach compared to earlier methods (Wooldridge, 1997; Heckman and Vytlacil, 1998) is that identification does not rely on the existence of a valid instrumental variable. Instead, identification comes from a linear projection of an individual's returns to adoption onto the observed history of her adoption, an approach similar to the correlated random effects (CRE) method in Chamberlain (1984). The S2011 application of a CRC model to development economics is novel and the paper is widely cited—yet the influence of the empirical approach has been limited. Citations to S2011 suggest its impact has largely been to document the existence and relevance of heterogeneous returns to technology adoption in developing countries, a finding that is not unique to S2011. Few, if any, citations center on the use or interpretation of the specific form of heterogeneity implied by S2011's linearity in comparative advantage assumption.

We revisit the econometric model in S2011 and propose an alternative group random coefficient (GRC) strategy that draws on recent developments in the nonparametric panel identification literature (Altonji and Matzkin, 2005; Chernozhukov *et al.*, 2013; Bester and Hansen, 2009). Our approach provides a convenient estimation strategy for several reasons. First, we show that the S2011 CRC model may be viewed as a restricted version of a GRC model that nonparametrically identifies returns to adoption assuming time homogeneity (Chernozhukov *et al.*, 2013). Unlike the reduced form of the S2011 CRC model, the parameters in the unrestricted GRC model have a clear economic interpretation. Second, the unrestricted GRC model has another practical advantage: it can elucidate potential identification concerns for the CRC model to the practitioner. Finally, our methodology is more easily extended to multiple time periods than the S2011 CRC approach, which can suffer from multi-collinearities when certain adoption histories are unobserved in a given dataset.

Throughout this paper, we illustrate our method using the same empirical example as S2011. Section 2 presents our model in the two-period case, clarifies its relationship with the S2011 model, and extends the approach to the multiple-period case. Section 3 describes and presents the empirical analysis revisiting hybrid adoption in Kenya and Section 4 discusses the relevance of our findings for the broader questions surrounding technology adoption in low-income agriculture.

2 A Group Random Coefficient Approach

Suppose that yield is a function of hybrid adoption, h_{it}, farmer ability, a_i, and idiosyncratic shocks, ϵ_{it}, which is given by the following, for $i = 1, \ldots, n$ and $t = 1, \ldots, T$.

$$y_{it} = f(h_{it}, a_i) + \epsilon_{it}. \tag{1}$$

In this paper as in S2011, we maintain the strict exogeneity assumption, i.e. $E[\epsilon_{it}|h_i, a_i] = 0$, where $h_i = (h_{i1}, \ldots, h_{iT})$ denotes individual i's adoption history and each h_{it} is a binary indicator of adoption. We impose no restriction on the distribution of farmer ability conditional on adoption history $(a_i|h_i)$, thereby treating a_i as a "fixed effect" (Chernozhukov *et al.*, 2013). Note that a_i may be any finite-dimensional vector. Furthermore, as per arguments in Chernozhukov *et al.* (2013), $f(h_{it}, a_i)$ may be viewed as the conditional mean function of a fully nonseparable model, $\phi(h_{it}, a_i, u_{it})$, where we assume time homogeneity, i.e. $u_{it}|h_i, a_i \stackrel{d}{=} u_{i1}|h_i, a_i$. For simplicity, we consider a model without covariates, but our approach to identification extends to the inclusion of additively separable covariates.

Following Chernozhukov *et al.* (2013), we can express the above equation equivalently as a random coefficient model by letting $\mu_i \equiv f(0, a_i)$ and $\Delta_i \equiv f(1, a_i) - f(0, a_i)$,

$$y_{it} = \mu_i + \Delta_i h_{it} + \epsilon_{it}. \tag{2}$$

In our empirical context, where a farmer decides whether or not to adopt a new agricultural technology, y_{it} represents maize yields, μ_i denotes farmer i's expected yield without adoption and Δ_i her returns to adoption, respectively. The econometric model of adoption in S2011 is given by the following,

$$y_{it} = \tau_i + \theta_i + (\beta + \phi\theta_i)h_{it} + \epsilon_{it}, \tag{3}$$

where τ_i is farmer i's absolute advantage assumed to be (mean) independent of technology adoption (i.e., $E[\tau_i|h_i] = E[\tau_i]$). Selection into technology adoption is determined by θ_i, farmer i's comparative advantage, which admits the normalization $E[\theta_i] = 0$. We note that (3) is a restricted version of (2), where $\mu_i = \tau_i + \theta_i$ and $\Delta_i = \beta + \phi\theta_i$. We call this key restriction—that the returns to adoption are linear in comparative advantage—the linearity in comparative advantage (LCA) restriction.

In the following, we present our group random coefficient (GRC) approach and show how the parameters of the S2011 model can be identified from a restricted version of the GRC method. We first outline our method for the two-period case and then generalize it to any finite $T \geq 2$.

2.1 Two-period Case

2.1.1 GRC Unrestricted Model

The GRC approach relies on the insight that with a binary variable and fixed T, there is a finite number of adoption histories. For $T = 2$, there are four possible realizations of h_i, with support given by $\mathcal{H} = \{0,1\}^2 = \{(0,0),(0,1),(1,0),(1,1)\}$. Since the adoption histories may imply different distributions of ability, it is natural to define subpopulations in terms of adoption histories. S2011 refers to the four subpopulations in the two-period case as never-adopters, joiners, leavers and always-adopters. We use $\mathcal{H}_S = \{(0,1),(1,0)\}$ to denote the set of switcher subpopulations. Its complement set, $\mathcal{H}_S^c = \{(0,0),(1,1)\}$, is composed of the two stayer subpopulations. As in S2011, we are interested in how the average returns to adoption varies across the different subpopulations. We therefore integrate the unrestricted model (2) with respect to $a_i|h_i$, which yields the following conditional mean model under strict exogeneity

$$E[y_{it}|h_i = \underline{h}] = \mu_{\underline{h}} + \Delta_{\underline{h}} h_{it}, \tag{4}$$

where $\mu_{\underline{h}} \equiv E[\mu_i|h_i = \underline{h}]$ and $\Delta_h \equiv E[\Delta_i|h_i = \underline{h}]$.

By the time homogeneity of $\mu_{\underline{h}}$ and $\Delta_{\underline{h}}$, we can identify the average return to adoption for subpopulations that we observe with and without hybrid adoption in our data. Hence, $\Delta_{\underline{h}}$ is only nonparametrically identified for the switcher subpopulations, $\underline{h} \in \mathcal{H}_S$. For stayer subpopulations, we can only identify their average yield with *or* without adoption. For the never-adopters, we can identify the average yield without adoption, $\mu_{(0,0)}$. For the always-adopters, we can identify the average yield with adoption, which we denote $\kappa_{(1,1)} = \mu_{(1,1)} + \Delta_{(1,1)}$. Without further restrictions, we cannot separately identify $\mu_{(1,1)}$ and $\Delta_{(1,1)}$. Hence, the returns to adoption is not nonparametrically identified for the stayer subpopulations.

All of the aforementioned identifiable objects can be estimated consistently using the following GRC model,

$$y_{it} = \sum_{\underline{h} \in \mathcal{H}\setminus(1,1)} \mu_{\underline{h}} 1\{h_i = \underline{h}\} + \sum_{\underline{h} \in \mathcal{H}_S} \Delta_{\underline{h}} h_{it} 1\{h_i = \underline{h}\} + \kappa_{(1,1)} h_{it} 1\{\underline{h} = (1,1)\} + \epsilon_{it}. \tag{5}$$

Unlike the reduced form of the CRC model in S2011, all of the coefficients in the above model have economic meaning; $\mu_{\underline{h}}$ is the average yield without hybrid adoption for subpopulation $\underline{h}$, Δ_h is the average return to adoption for switcher subpopulation $\underline{h}$, and $\kappa_{(1,1)}$ is the average yield with hybrid for the alwayers-adopters. Next, we impose the LCA restriction on the above model and illustrate how the unrestricted model can indicate potential identification concerns for ϕ, a key parameter in the S2011 model.

2.1.2 GRC Model with the LCA Restriction

The following proposition establishes the relationship between parameters in the unrestricted GRC model and those in the S2011 model, which imposes the LCA restriction. Let $\theta_{\underline{h}} = E[\theta_i | h_i = \underline{h}]$.

Proposition 1. *Let* $y_{it} = \mu_i + \Delta_i h_{it} + \epsilon_{it}$. *Assume* $\mu_i = \tau_i + \theta_i$, $\Delta_i = \beta + \phi\theta_i$, $E[\theta_i] = 0$, $E[\tau_i | h_i] = E[\tau_i]$, $E[\epsilon_{it} | h_i, \tau_i, \theta_i] = 0$, *the following equalities hold for* $\underline{h}, \underline{h}' \in \mathcal{H} = \{0, 1\}^T$,

(i) $\Delta_{\underline{h}} = \beta + \phi\theta_{\underline{h}}$,

(ii) $\mu_{\underline{h}} - \mu_{\underline{h}'} = \theta_{\underline{h}} - \theta_{\underline{h}'}$,

(iii) $\Delta_{\underline{h}} - \Delta_{\underline{h}'} = \phi\left(\theta_{\underline{h}} - \theta_{\underline{h}'}\right)$, *for* $\underline{h} \neq \underline{h}'$.

The derivation of the above results is provided in a supplementary appendix. The above identities imply that we can re-write ϕ as the ratio of different subpopulations' differences in returns to adoption to their differences in yield without adoption. Since we can identify both $\mu_{\underline{h}}$ and $\Delta_{\underline{h}}$ for switcher subpopulations, ϕ is identified from the following in the two-period case, assuming $\mu_{(1,0)} \neq \mu_{(0,1)}$,

$$\phi = \frac{\Delta_{(1,0)} - \Delta_{(0,1)}}{\mu_{(1,0)} - \mu_{(0,1)}}. \tag{6}$$

The above equality illustrates that ϕ would not be identified if joiners and leavers had the same average yield without adoption. Since the unrestricted GRC model enables us to estimate both parameters without imposing the LCA restriction, we can use these estimates to evaluate potential concerns regarding the identification of ϕ.

We can write the restricted GRC model by solving (6) for $\Delta_{(1,0)}$ using the above identity and plugging it into (5):

$$\begin{aligned} y_{it} = &\sum_{\underline{h} \in \mathcal{H} \setminus (1,1)} \mu_{\underline{h}} + \Delta_{(0,1)} h_{it} + \phi(\mu_{(1,0)} - \mu_{(0,1)}) h_{it} 1\{h_i = (1,0)\} \\ &+ \left(\mu_{(1,1)} + \phi\left(\mu_{(1,1)} - \mu_{(0,1)}\right)\right) h_{it} 1\{h_i = (1,1)\} + \epsilon_{it}, \end{aligned} \tag{7}$$

where the restriction on the coefficient on $h_{it}1\{h_i = (1,1)\}$ follows from noting that $\kappa_{(1,1)} - \Delta_{(0,1)} = \mu_{(1,1)} + \Delta_{(1,1)} - \Delta_{(0,1)}$ and using Proposition 1 *(ii)-(iii)*. The restricted GRC model can be estimated using nonlinear least squares or method of moments.

Since the LCA restriction allows us to extrapolate to stayers as pointed out in Verdier (2020), we can identify $\mu_{(1,1)}$. Further, this allows identification of β and $\theta_{\underline{h}}$ for all $\underline{h} \in \mathcal{H}$. Let $\pi_{\underline{h}} = P(h_i = \underline{h})$ for $\underline{h} \in \mathcal{H}$. Note that $E[\mu_i] = \sum_{\underline{h} \in \mathcal{H}} \pi_{\underline{h}} \mu_{\underline{h}}$. Since $E[\theta_i] = 0$, $E[\mu_i] = E[\tau_i]$. As a result,

$$\theta_{\underline{h}} = E[\mu_i | h_i = \underline{h}] - E[\tau_i | h_i = \underline{h}] = E[\mu_i | h_i = \underline{h}] - E[\tau_i] = \mu_{\underline{h}} - \sum_{\underline{h} \in \mathcal{H}} \pi_{\underline{h}} \mu_{\underline{h}}.$$

Since $\Delta_{\underline{h}} = \beta + \phi\theta_{\underline{h}}$, we can therefore also identify $\beta = \Delta_{(0,1)} - \phi\theta_{(0,1)} = \Delta_{(1,0)} - \phi\theta_{(1,0)}$.

2.2 Multiple-Period Case

For the general case, where $T \geq 2$, we first generalize our notation. Let $\underline{h} = (h_1, \ldots, h_T) \in \mathcal{H} = \{0,1\}^T$, $\mathcal{H}_S = \{\underline{h} \in \mathcal{H}_S : 0 < \sum_{t=1}^T h_{it} < T\}$, and $\mathcal{H}_S^c = \mathcal{H} \backslash \mathcal{H}_S$. We can now write our unrestricted GRC model for any $T \geq 2$

$$y_{it} = \sum_{\underline{h} \in \mathcal{H}: \sum_{t=1}^T h_t < T} \mu_{\underline{h}} 1\{h_i = \underline{h}\} + \sum_{\underline{h} \in \mathcal{H}_S} \Delta_{\underline{h}} h_{it} 1\{h_i = \underline{h}\} + \kappa_{\underline{h}_T} h_{it} 1\left\{\sum_{t=1}^T h_{it} = T\right\} + \epsilon_{it}. \quad (8)$$

where $\underline{h}_T$ denotes the always-adopter trajectory.

Using Proposition 1, we can obtain a restricted version of the above model,

$$\begin{aligned} y_{it} = & \sum_{\underline{h} \in \mathcal{H}: \sum_{t=1}^T h_t < T} \mu_{\underline{h}} + \Delta_{\underline{h}_0} h_{it} + \sum_{\underline{h} \in \mathcal{H}_S \backslash \underline{h}_0} \phi(\mu_{\underline{h}} - \mu_{\underline{h}_0}) h_{it} 1\{h_i = \underline{h}\} \\ & + \left(\mu_{\underline{h}_T} + \phi\left(\mu_{\underline{h}_T} - \mu_{\underline{h}_0}\right)\right) h_{it} 1\left\{\sum_{t=1}^T h_{it} = T\right\} + \epsilon_{it}, \end{aligned} \quad (9)$$

for some baseline trajectory $\underline{h}_0 \in \mathcal{H}_S$. We use efficient GMM to estimate the above model using all regressors as instruments. When $T > 2$, then there are more than two switcher subpopulations, $|\mathcal{H}_S| > 2$. Since the identification of ϕ only requires the presence of two switcher subpopulations, we would have over-identifying restrictions for this parameter and can use Hansen's J-statistic to test them.

In sum, the GRC approach provides several appealing features, especially when $T > 2$. First, the S2011 approach to estimating the CRC model is relatively cumbersome to adapt to the multiple-period case. This is due to multicollinearities that arise in the reduced form whenever some adoption histories are unobserved in a given dataset.[1] Since the regressors in our GRC approach consist of dummy variables for the adoption histories, this issue is circumvented by the inclusion of dummy variables for the observed trajectories. Second, the unrestricted GRC model, unlike the reduced form of the CRC model, has an economic interpretation and provides the practitioner with insights on potential identification concerns pertaining to the parameter ϕ. Third, the GRC approach allows us to test the two key restrictions in the S2011 model—specifically, the time homogeneity assumption and the LCA restriction. Finally, relating the S2011 model with the panel identification literature provides other testable identifying assumptions, including exchangeability and other non-parametric correlated random effects restrictions (Altonji and Matzkin, 2005; Bester and Hansen,

[1]To obtain the reduced form of the CRC model in S2011, θ_i is projected onto a fully saturated model of h_{it} for all $t = 1, \ldots, T$. As soon as any adoption history is unobserved, then at least two of the independent variables in this projection become collinear.

2009; Ghanem, 2017).

3 Revisiting Hybrid Maize Adoption in Kenya

We demonstrate the advantages of the GRC approach using the same empirical application as S2011. Specifically, we compare the GRC and CRC estimators in the context of hybrid maize adoption among Kenyan farmers using the same nationwide panel dataset used in S2011.[2] We construct our sample and generate variables according to instructions provided by the author of the original study.[3] This provides a working dataset that is very close, but not identical, to the data used in S2011.

We begin by replicating the OLS and fixed effects (FE) specifications that launch the analysis in S2011. These results, shown in Table 1, provide an estimate of the yield advantages of hybrid maize compared to non-hybrid maize varieties. While the results in Panel A without district fixed effects and without controls generate estimates that are statistically identical to those provided in S2011 for the coefficient on hybrid, our point estimates are smaller in magnitude than hers once we include these fixed effects and controls.[4] Whereas OLS results in S2011 with controls suggest hybrids are associated with 54% higher yield, ours suggest a more modest 23%. As in S2011, we cannot reject that the FE results in Panel A are different from zero.

In Panel B, we estimate the same specifications including data from the year 2000, which was dropped from the analysis reported in S2011. While not directly comparable to results presented in S2011, the patterns in Panel B include similar hybrid point estimates for the OLS regressions, but the FE estimates are larger and statistically different from zero.

We replicate the CRC results in S2011 in Panel A of Table 2 using the `Stata` package from (Barriga Cabanillas *et al.*, 2018).[5] Most of the important coefficient estimates in this panel are statistically indistinguishable from S2011. One exception is that our point estimates of β, the average return to hybrid, are smaller than those in S2011. Further, our estimates of ϕ, the key sorting parameter in the S2011 CRC model are substantially different from the original. Whereas S2011 finds that ϕ is consistently negative and often significantly different from zero, our ϕ point estimates are small and statistically insignificant. This is an important discrepancy given that this parameter drives the main conclusion in S2011: it defines the specific form of heterogeneity

[2]The dataset is the Tegemeo Agricultural Policy Research and Analysis (TAPRA) Rural Household survey. This nationwide survey was collected in five waves in 1997, 2000, 2004, 2007, and 2010. The agricultural variables cover the 1996/97, 1999/2000, 2003/04, 2006/07, and 2009/10 crop years. The first wave of the survey interviewed 1,500 agricultural households, covering 22 districts and 107 villages across eight agroecological zones. Sampling probabilities were proportional to village size with reference to census data.

[3]Details on the dataset and variable construction are in the supplementary appendix as well as the cleaning code, which is available upon request.

[4]The results in Panel A correspond to Table IIIA in S2011.

[5]We omit the results in columns (3) and (6) in S2011 Table VIIIA, which include both covariates and interactions of the covariates with the hybrid decision.

that drives hybrid adoption. Our inability to replicate this key finding in S2011 is unexpected and can only be attributed to seemingly trivial differences in the working data we construct from the Tegemeo panel dataset as discussed in the supplementary appendix. When we extend the analysis to $T = 3$ in Panel B, however, our CRC estimate of ϕ becomes negative and significant.

The unrestricted GRC estimates help explain the results in Table 1 and the inconsistent CRC results in Table 2. For $T = 2$ (Table 3), the yield returns to hybrid adoption are very different for joiners (Δ_{01}) and leavers (Δ_{10}). This explains why the estimated coefficient on hybrid for the $T = 2$ FE regression is insignificant, as it pools these two switcher subgroups. However, the average yield without adoption for these two subgroups (μ_{01} and μ_{10}) is statistically indistinguishable, especially when we add control variables in column (2). As noted in (6) (see also Proposition 1 *(ii)-(iii)*), ϕ is not identified when joiners and leavers have the same average yield without adoption; this complicates the identification and estimation of ϕ. This may explain why very small discrepancies in our working data seem to have disproportionate effects on the results for this key parameter. With this complication in mind, we report the results from the restricted GRC model, where we impose the LCA restriction, in columns (3) and (4) of Table 3. These restricted GRC estimates suggest, like S2011, that ϕ is large and negative.

Extending the GRC to $T = 3$ provides further insights to resolve this inconsistent pattern of results. In Table 4, we continue to see heterogeneous yield effects associated with hybrid adoption, i.e. the estimated Δ coefficients. With the parametric sorting restriction imposed, we find negative and significant estimates of ϕ, but with $T = 3$ we can also use over-identification tests of this restriction. When we do, we clearly reject the LCA restriction.

In sum, using the same underlying dataset as S2011, we do not see consistent evidence that ϕ is negative in the two-year panel. We attribute this inconsistent evidence to weak identification of this sorting parameter—likely stemming from the fact that yield outcomes in the absence of hybrid adoption are indistinguishable for joiners and leavers. When we extend our analysis to the three-year panel, we reject the parametric restriction on which this parameter is based, suggesting that the ϕ-based form of heterogeneous returns is inappropriate in the context of hybrid maize adoption among Kenyan farmers.

4 Discussion

Despite being well-known and widely-cited in development economics, S2011 has had surprisingly limited methodological impact: rather than leading to widespread use of the CRC method and nuanced discussions of the specific form of heterogeneity it implies, the article is largely cited for showing the presence of heterogeneity in returns to new technologies. This is unfortunate since optimal policy and product design will often vary with the particular form that heterogeneity of returns takes rather than its existence. We revive the methodological contribution of S2011 by generalizing her model to make it more flexible and more accessible for empirical analysis.

Our results lead us to different conclusions about the form of heterogeneity than S2011 and therefore (re-)open important questions regarding the factors that drive technology adoption in agriculture in low-income countries. While we find evidence of heterogeneity in the returns to adoption across joiners and leavers, we cannot consistently replicate the ϕ estimates in S2011 using the two-year panel. Using the extended three-year panel dataset, the CRC and restricted GRC approaches produce similar ϕ estimates. However, we reject the parametric restriction that yields this parameter, which suggests that the linearity of response in comparative advantage does not hold in this empirical context.

References

Altonji, J. and Matzkin, R. (2005). Cross-section and panel data estimators for nonseparable models with endogenous regrssors. *Econometrica*, **73** (3), 1053–1102.

Barriga Cabanillas, O., Michler, J. D., Michuda, A. and Tjernström, E. (2018). Fitting and Interpreting Correlated Random Coefficient (CRC) Models Using Stata. *Stata Journal*, **18** (1), 159–173.

Bester, C. A. and Hansen, C. (2009). Identification of marginal effects in a nonparametric correlated random effects model. *Journal of Business & Economic Statistics*, **27** (2), 235–250.

Chamberlain, G. (1984). Panel data. In Z. Griliches and M. Intriligator (eds.), *Handbook of Econometrics*, Elsevier, pp. 1247–1318.

Chernozhukov, V., Fernandez-Val, I., Hahn, J. and Newey, W. (2013). Average and quantile effects in nonseparable panel data models. *Econometrica*, **81** (2), pp.535–580.

Ghanem, D. (2017). Testing identifying assumptions in nonseparable panel data models. *Journal of Econometrics*, **197** (2), 202–217.

Heckman, J. and Vytlacil, E. (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), 974–987.

Suri, T. (2011). Selection and comparative advantage in technology adoption. *Econometrica*, **79** (1), 159–209.

Verdier, V. (2020). Average treatment effects for stayers with correlated random coefficient models of panel data. *Journal of Applied Econometrics*, **forthcoming**.

Wooldridge, J. (1997). On two stage least squares estimation of the average treatment effect in a random coefficient model. *Economics Letters*, **56** (2), 129–133.

Table 1: OLS and FE (Table IIIA in Suri 2011)

	OLS			FE	
Panel A: 1997 and 2004					
Hybrid	1.042***	0.277***	0.231***	0.0139	0.0308
	(0.0467)	(0.0503)	(0.0475)	(0.0696)	(0.0662)
Acres			-0.00679		-0.0637***
			(0.00880)		(0.0217)
Seed rate (kg/acre) x 10			0.219***		0.177***
			(0.0310)		(0.0309)
Land prep (Ksh/acre) x 1000			0.0111***		0.0190***
			(0.00300)		(0.00517)
Fertilizer (Ksh/acre) x 1000			0.0229***		0.0116***
			(0.00308)		(0.00390)
Hired labor (Ksh/acre) x 1000			0.0372***		0.0240***
			(0.00968)		(0.00926)
Family labor (hours/acre) x 1000			0.234***		0.241**
			(0.0855)		(0.108)
N	1197	1197	1197	1197	1197
$N \times T$	2394	2394	2394	2394	2394
District FE	No	Yes	Yes		
Controls	No	No	Yes	No	Yes
Adj. R^2	0.21	0.41	0.51	0.49	0.56
Panel B: 1997, 2000, and 2004					
Hybrid	0.994***	0.343***	0.263***	0.124**	0.105**
	(0.0413)	(0.0425)	(0.0401)	(0.0543)	(0.0518)
Acres			-0.0132		-0.0706***
			(0.00851)		(0.0205)
Seed rate (kg/acre) x 10			0.262***		0.233***
			(0.0207)		(0.0217)
Land prep (Ksh/acre) x 1000			0.0176***		0.0180***
			(0.00427)		(0.00657)
Fertilizer (Ksh/acre) x 1000			0.0263***		0.0130***
			(0.00352)		(0.00282)
N	1197	1197	1197	1197	1197
$N \times T$	3591	3591	3591	3591	3591
District FE	No	Yes	Yes		
Controls	No	No	Yes	No	Yes
Adj. R^2	0.19	0.36	0.43	0.45	0.50

Notes: Dependent variable is ln yield. Covariates follow Suri (2011): All regressions include the following household-level demographic controls: household size, the number of boys (males<age 16), the number of girls (females<age 16), the number of men (aged 17 to 39), the number of women, and the number of older men (> age 40). All regressions additionally control for the number of maize acres, the seed rate (kg per acre), land preparation expenditures (Ksh per acre), and fertilizer expenditure (Ksh per acre), as well as main season rainfall. When $t = 2$, we additionally include hired labor and family labor (both measured in hours per acre), but these variables were not collected in the 2000 data collection wave.
OLS specification: $y_{it} = \delta + \beta h_{it} + X_{it}^{'}\gamma + \varepsilon_{it}$
FE specification: $y_{it} = \delta + \alpha_i + \beta h_{it} + X_{it}^{'}\gamma + \epsilon_{it}$

Table 2: CRC results (Table VIIIA in Suri 2011)

	Full sample		Without HIV districts	
Panel A: 1997 and 2004				
λ_1	0.716***	0.639***	0.630***	0.545***
	(0.0761)	(0.0710)	(0.0782)	(0.0720)
λ_2	0.923***	0.701***	0.937***	0.727***
	(0.108)	(0.0993)	(0.116)	(0.106)
λ_3	-0.334	-0.465	-0.297	-0.322
	(0.405)	(0.443)	(0.305)	(0.295)
β	0.0195	0.0161	-0.0151	0.0576
	(0.0842)	(0.478)	(0.133)	(0.104)
ϕ	0.104	0.854	-0.0443	-0.00475
	(0.935)	(4.116)	(0.605)	(0.952)
Observations	1197	1197	1057	1057
Controls	No	Yes	No	Yes
χ^2	8927.8	8753.7	25298.1	26736.0
Panel B: 1997, 2000, and 2004				
λ_1	0.443***	0.388***	0.371***	0.323***
	(0.0759)	(0.0718)	(0.0796)	(0.0759)
λ_2	0.0641	-0.0217	0.0389	-0.0512
	(0.0779)	(0.0729)	(0.0868)	(0.0811)
λ_3	0.513***	0.361***	0.539***	0.363***
	(0.107)	(0.103)	(0.116)	(0.114)
λ_4	0.260**	0.289**	0.216	0.240*
	(0.128)	(0.122)	(0.132)	(0.126)
λ_5	-0.141	-0.0697	-0.286*	-0.211
	(0.169)	(0.168)	(0.167)	(0.171)
λ_6	0.473***	0.514***	0.331*	0.435**
	(0.170)	(0.161)	(0.180)	(0.174)
λ_7	-0.0626	-0.0695	0.115	0.191
	(0.233)	(0.273)	(0.236)	(0.332)
β	0.00396	-0.0570	-0.0313	-0.150
	(0.106)	(0.128)	(0.122)	(0.181)
ϕ	-0.333**	-0.494***	-0.295	-0.527***
	(0.156)	(0.151)	(0.194)	(0.181)
Observations	1197	1197	1057	1057
Controls	No	Yes	No	Yes
χ^2	27248.6	27375.9	57411.2	51857.4

Notes: Dependent variable is ln yield. Covariates follow Suri (2011): All regressions include the following household-level demographic controls: household size, the number of boys (males<age 16), the number of girls (females<age 16), the number of men (aged 17 to 39), the number of women, and the number of older men (> age 40). All regressions additionally control for the number of maize acres, the seed rate (kg per acre), land preparation expenditures (Ksh per acre), and fertilizer expenditure (Ksh per acre), as well as main season rainfall. When $t = 2$, we additionally include hired labor and family labor (both measured in hours per acre), but these variables were not collected in the 2000 data collection wave.
The λ coefficients correspond to the following projections:
For $t = 2$, $\theta_i = \lambda_0 + \lambda_1 h_{i1} + \lambda_2 h_{i2} + \lambda_3 h_{i1} h_{i2} + \nu_i$
For $t = 3$ the projection is $\theta_i = \lambda_0 + \lambda_1 h_{i1} + \lambda_2 h_{i2} + \lambda_3 h_{i3} + \lambda_4 h_{i1} h_{i2} + \lambda_5 h_{i1} h_{i3} + \lambda_6 h_{i2} h_{i3} + \lambda_7 h_{i1} h_{i2} h_{i3} + \nu_i$.

Table 3: GRC models, 2 periods

	Unrestricted		Restricted	
μ_{00}	5.246***	0.210*	5.246***	4.775***
	(0.0459)	(0.111)	(0.0459)	(0.0780)
μ_{01}	5.942***	0.711***	5.942***	5.364***
	(0.0916)	(0.217)	(0.0916)	(0.103)
μ_{10}	6.215***	0.704***	6.215***	5.685***
	(0.0746)	(0.166)	(0.0746)	(0.0972)
μ_{11}			5.870***	5.230***
			(0.137)	(0.180)
Δ_{01}	0.508***	-0.269	0.508***	0.436***
	(0.0978)	(0.263)	(0.0978)	(0.0889)
Δ_{10}	-0.476***	0.363**		
	(0.0950)	(0.165)		
ϕ			-3.602***	-2.547***
			(1.337)	(0.737)
Derived parameters:				
$\bar{\mu}$			5.76	5.18
			(0.079)	(0.11)
β			1.18	0.91
			(0.24)	(0.080)
Hypothesis tests ($p-$values):				
H_0: All μ equal	0.021	0.98	0.021	0.0047
H_0: All Δ equal			5.6e-13	1.3e-10
$H_0 : \phi = 0$			0.0070	0.00055
Observations	2394	2394	2394	2394
Controls	No	Yes	No	Yes

Notes: Dependent variable is ln yield. Covariates follow Suri (2011). All regressions include the following household-level demographic controls: household size, the number of boys (males<age 16), the number of girls (females<age 16), the number of men (aged 17 to 39), the number of women, and the number of older men (> age 40). All regressions additionally control for the number of maize acres, the seed rate (kg per acre), land preparation expenditures (Ksh per acre), and fertilizer expenditure (Ksh per acre), as well as main season rainfall. When $t = 2$, we additionally include hired labor and family labor (both measured in hours per acre), but these variables were not collected in the 2000 data collection wave.

Table 4: GRC models, 3 periods

	Unrestricted		Restricted	
μ_{000}	5.288***	0.438***	5.288***	4.842***
	(0.0484)	(0.113)	(0.0484)	(0.0658)
μ_{001}	5.677***	0.566*	5.790***	5.203***
	(0.143)	(0.309)	(0.137)	(0.144)
μ_{010}	5.341***	0.343*	5.352***	4.859***
	(0.0850)	(0.204)	(0.0835)	(0.0953)
μ_{011}	6.127***	1.212***	6.216***	5.633***
	(0.0929)	(0.227)	(0.0905)	(0.104)
μ_{100}	5.856***	0.621***	5.836***	5.400***
	(0.0935)	(0.197)	(0.0936)	(0.101)
μ_{101}	5.948***	0.999***	6.149***	5.651***
	(0.198)	(0.287)	(0.166)	(0.169)
μ_{110}	6.318***	1.071***	6.226***	5.791***
	(0.103)	(0.215)	(0.101)	(0.107)
μ_{111}			7.007***	6.699***
			(0.262)	(0.398)
Δ_{001}	0.521***	-0.137	0.114	0.174*
	(0.165)	(0.382)	(0.0789)	(0.0958)
Δ_{010}	0.325***	0.0189		
	(0.125)	(0.187)		
Δ_{011}	0.269**	-0.442**		
	(0.118)	(0.217)		
Δ_{100}	-0.269**	0.465**		
	(0.117)	(0.193)		
Δ_{101}	0.298	0.408*		
	(0.201)	(0.239)		
Δ_{110}	-0.336***	-0.160		
	(0.107)	(0.192)		
ϕ			-0.392***	-0.569***
			(0.142)	(0.130)
Derived parameters:				
$\bar{\mu}$			6.33	5.94
			(0.13)	(0.21)
β			-0.097	-0.24
			(0.12)	(0.18)
Hypothesis tests ($p-$values):				
H_0: All μ equal	4.0e-13	0.022	9.4e-13	1.7e-13
H_0: All Δ equal			0.35	0.021
$H_0 : \phi = 0$			0.0057	0.000011
H_0 : Over-id restrictions valid			0.0000018	0.000034
Observations	3591	3591	3591	3591
Controls	No	Yes	No	Yes

Notes: Dependent variable is ln yield. Covariates follow Suri (2011). All regressions include the following household-level demographic controls: household size, the number of boys (males<age 16), the number of girls (females<age 16), the number of men (aged 17 to 39), the number of women, and the number of older men (> age 40). All regressions additionally control for the number of maize acres, the seed rate (kg per acre), land preparation expenditures (Ksh per acre), and fertilizer expenditure (Ksh per acre), as well as main season rainfall. When $t = 2$, we additionally include hired labor and family labor (both measured in hours per acre), but these variables were not collected in the 2000 data collection wave.