# Smoothed LSDV Estimation of Functional-Coefficient Panel Data Models with Two-Way Fixed Effects*

Shaymal C. Halder[1]  Emir Malikov[2]

[1]Auburn University
[2]University of Nevada, Las Vegas

April 20, 2020

**Abstract**

The existing semiparametric estimators for varying-coefficient fixed-effects models exclusively assume one-way fixed effects, typically in the dimension of cross-sectional units. However, more often than not applied researchers wish to control for both the individual and time fixed effects in their panel regressions, with the latter included to account for common unobservable factors correlated with regressors. While rather trivial in a linear model, controlling for time effects by explicitly including time-period dummies as additional regressors does not provide a straight-forward estimation procedure in the case of a semiparametric model. We provide an alternative by extending the Sun et al. (2009) smoothed least-squares dummy variable (LSDV) estimator to the case of a functional-coefficient model with *two*-way fixed effects whereby we allow for unobservable heterogeneity in both dimensions of the data: cross-section and time. Both fixed effects are concentrated out of the model via locally smoothed two-dimensional within transformation. Simulations show that the estimator performs well in finite samples. We also showcase its practical usefulness by revisiting the role of management as a factor of production.

**Keywords**: fixed effect, local linear, LSDV, semiparametric, smooth coefficient, time effect

**JEL Classification**: C23

**Correspondence to* emir.malikov@unlv.edu (Malikov). We would like to thank an anonymous referee for many insightful comments and suggestions that helped improve this paper. Any remaining errors are our own.

# 1 Introduction

A linear fixed-effects model is a workhorse of applied research in economics. The model's broad popularity chiefly stems from its ability (given the availability of panel data) to control for unobservable confounders that may correlate with regressors. However, just like all parametric models, a linear fixed-effects regression is prone to misspecification owing to its reliance on the parametric form assumption (here, linearity) the violation of which may lead to inconsistency and thus misleading inference. The semiparametric functional-coefficient[1] fixed-effects model of Sun et al. (2009) provides a means to robustify the conventional linear fixed-effects regression by letting its coefficients be unspecified nonparametric functions of relevant contextual variables which, among other benefits, accommodates potential heterogeneity in marginal effects of linear regressors. While not as flexible as a fully nonparametric fixed-effects model (Henderson et al., 2008), such a semiparametric specification is attractive because of its ability to alleviate the so-called "curse of dimensionality" associated with nonparametric estimation and thus to achieve better convergence rates.

The existing semiparametric estimators for varying-coefficient fixed-effects models (Sun et al., 2009; Rodriguez-Poo & Soberón, 2014, 2015), however, exclusively assume *one*-way fixed effects, typically in the dimension of individual cross-sectional units.[2] This is unfortunate because more often than not researchers wish to control for *both* unit- and time-specific fixed effects in their panel regressions, especially with the widespread popularity of the difference-in-difference based identification strategies. In addition to the conventional time-invariant individual effects, the time effects are included with the intent to control for common unobservable factors correlated with regressors. In linear models, this is rather simple. Since most microeconomic studies use short panels ($n \gg T$), practitioners customarily control for time effects by explicitly including time-period dummies as additional regressors, whereas the individual fixed effects are usually transformed out of the equation by either the within or first-difference transformation. The same procedure is however not as trivial in the case of a semiparametric functional-coefficient model because the direct estimation of constant coefficients on such time-period dummies renders the model a partially linear, functional-coefficient regression thereby necessitating more than a single-step estimation which, besides theoretical complications, is also more computationally demanding.

To fill this practically important gap, we contribute to the literature by extending the Sun et al. (2009) smoothed least-squares dummy variable (SLSDV) estimator to the case of a functional-coefficient panel data model with *two*-way fixed effects whereby we allow for unobservable heterogeneity in both dimensions of the data: cross-section and time. Both fixed effects are concentrated out of the model via *local* kernel-smoothed two-dimensional within transformation. One of the benefits of the proposed estimation procedure is that the unknown functional coefficients are estimated in a single step. Using local-linear fitting, the simulation experiments show that the two-way SLSDV

[1] Also sometimes referred to as a "varying-coefficient" or "smooth-coefficient" model.

[2] The same applies to other types of semiparametric fixed-effects panel data models such as a partially linear specification (see Su & Ullah, 2006; Qian & Wang, 2012; Lin et al., 2014).

estimator performs well in finite samples. We also showcase its practical usefulness by revisiting the role of management as a factor of production, with the focus on variability in its contribution to firm output in the face of technological and institutional changes.

## 2 The Model

We consider the problem of estimating the following semiparametric functional-coefficient panel data model with *two*-way fixed effects:

$$y_{it} = \mathbf{x}_{it}' \boldsymbol{\beta}(\mathbf{z}_{it}) + \mu_i + \lambda_t + u_{it}, \tag{2.1}$$

where $i = 1, \ldots, n$, $t = 1, \ldots, T$. The outcome $y_{it}$ is a scalar; the covariate column vector $\mathbf{x}_{it}$ is of dimension $p$ and excludes time-invariant regressors for identification purposes; $\boldsymbol{\beta}(\cdot)$ is a $p \times 1$ vector of the corresponding unknown functional coefficients which are assumed to be smooth functions; and the vector of smoothing variables $\mathbf{z}_{it}$ is of dimension $q$. The remaining variables in the equation are all scalars. We assume the following about unobservables in the model. The individual-specific effects $\{\mu_i\}$ are an *i.i.d.* sequence (over $i$) with a zero mean and finite variance. Analogously, the time-specific effect $\lambda_t$ is *i.i.d.* over $t$ and also has a zero mean and finite variance. We let both $\mu_i$ and $\lambda_t$ correlate with $\mathbf{x}_{it}$ and/or $\mathbf{z}_{it}$ in an arbitrary unspecified way. Such a nonparametric treatment of both unobservable effects and their correlation with the regressors in the equation renders them "fixed." Lastly, the random disturbance $u_{it}$ is assumed to be *i.i.d.* over both $i$ and $t$ with a zero mean and finite variance, and $\mathbb{E}[u_{it}|\mathbf{w}_i] = 0$ with $\mathbf{w}_i = (\mu_i, \lambda_t, \mathbf{x}_{i1}', \ldots, \mathbf{x}_{iT}', \mathbf{z}_{i1}', \ldots, \mathbf{z}_{iT}')'$. Thus, eq. (2.1) is a fixed-effects model with strictly exogenous regressors.

The salient feature of our model is that it accommodates heterogeneity in *both* dimensions of the data: cross-section ($\mu_i$) and time ($\lambda_t$). In that, it nests the more traditional functional-coefficient one-way fixed-effects panel data model with individual effects only. Namely, when $\lambda_t = 0$ for all $t$, model (2.1) reduces to the more standard specification first studied by Sun et al. (2009).

The consistent estimation of (2.1) is complicated by the presence of unobservable $\{\mu_i\}$ and $\{\lambda_t\}$ which cannot be ignored due to their correlation with the regressors. A popular approach to tackling fixed effects is to transform the model to rid it of the former. For instance, this is the route taken by Rodriguez-Poo & Soberón (2014, 2015) in their estimation of the one-way model whereby they remove individual effects via cross-time first-difference/within transformation *before* taking a local approximation of the transformed equation. Such an approach is non-trivial in the case with two fixed effects because (i) if first-differencing, it is not obvious what cross-sections to difference $\lambda_t$ over given that there is typically no natural ordering across units and (ii) if within-transforming, the local approximation of the transformed two-way model would need to be $(n+T)q$-variate which would deliver an exceptionally poor convergence rate and is practically prohibitive.[3] We therefore

[3]In case of the two-way first-differencing, the local approximation would need to be $4q$-variate which is also likely practically prohibitive.

proceed with the alternative method whereby both fixed effects are removed *after* taking the local approximation of (2.1).

Building on Sun et al.'s (2009) idea, we generalize their smoothed LSDV approach to tackling *more than one* fixed effect in the functional-coefficient model. This extension is rather natural. Writing the model in (2.1) in the matrix from, we have

$$\mathbf{y} = \mathbb{M}\{\mathbf{X}, \boldsymbol{\beta}(\mathbf{Z})\} + \mathbf{D}\boldsymbol{\mu} + \mathbf{P}\boldsymbol{\lambda} + \mathbf{u}, \tag{2.2}$$

where $\mathbf{y} = (\mathbf{y}_1', \ldots, \mathbf{y}_n')'$ and $\mathbf{u} = (\mathbf{u}_1', \ldots, \mathbf{u}_n')'$ are $(nT) \times 1$ vectors, with $\mathbf{y}_i = (y_{i1}, \ldots, y_{iT})'$ and $\mathbf{u}_i = (u_{i1}, \ldots, u_{iT})'$. The operator $\mathbb{M}\{\cdot\}$ stacks $\mathbf{x}_{it}'\boldsymbol{\beta}(\mathbf{z}_{it})$ into a $nT \times 1$ vector with the $(i, t)$ subscripts matching those of $\mathbf{y}$ and $\mathbf{u}$. Further, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)'$ is an $n \times 1$ vector of individual fixed effects, and $\mathbf{D} = \mathbf{I}_n \otimes \mathbf{i}_T$ is an $(nT) \times n$ design matrix, where $\mathbf{I}_n$ is an identity matrix of dimension $n$ and $\mathbf{i}_T$ denotes an $T \times 1$ vector of ones. Analogously, $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_T]'$ is a $T \times 1$ vector of time fixed effects, with $\mathbf{P} = \mathbf{i}_n \otimes \mathbf{I}_T$.

As the name suggests, the smoothed LSDV approach involves local smoothing around $\mathbf{z}_{it} = z$. First, define a $nT \times nT$ diagonal matrix of local kernel weights $\mathbf{W}_H(z) = \text{diag}\{\mathbf{K}_H(\mathbf{z}_1, z), \cdots, \mathbf{K}_H(\mathbf{z}_n, z)\}$, with $\mathbf{K}_H(\mathbf{z}_i, z) = \text{diag}\{k_H(z_{i1}, z), \cdots, k_H(z_{iT}, z)\}$ being a $T \times T$ diagonal matrix for each $i$, where $k_H(\mathbf{z}_{it}, z) = k\{\mathbf{H}^{-1}(\mathbf{z}_{it} - z)\}$ is the $q$-variate product kernel and $\mathbf{H} = \text{diag}\{h_1, \ldots, h_q\}$ is a diagonal bandwidth matrix of dimension $q$. To derive the estimator for unknown functional coefficients, we then solve the following locally weighted least-squares problem:

$$\min_{\boldsymbol{\beta}(z), \boldsymbol{\mu}, \boldsymbol{\lambda}} [\mathbf{y} - \mathbb{M}\{\mathbf{X}, \boldsymbol{\beta}(z)\} - \mathbf{D}\boldsymbol{\mu} - \mathbf{P}\boldsymbol{\lambda}]'\mathbf{W}_H(z)[\mathbf{y} - \mathbb{M}\{\mathbf{X}, \boldsymbol{\beta}(z)\} - \mathbf{D}\boldsymbol{\mu} - \mathbf{P}\boldsymbol{\lambda}]. \tag{2.3}$$

The first-order conditions of (2.3) with respect to both fixed effects $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ are

$$\mathbf{D}'\mathbf{W}_H(z)[\mathbf{y} - \mathbb{M}\{\mathbf{X}, \boldsymbol{\beta}(z)\} - \mathbf{D}\boldsymbol{\mu} - \mathbf{P}\boldsymbol{\lambda}] = 0 \tag{2.4}$$
$$\mathbf{P}'\mathbf{W}_H(z)[\mathbf{y} - \mathbb{M}\{\mathbf{X}, \boldsymbol{\beta}(z)\} - \mathbf{D}\boldsymbol{\mu} - \mathbf{P}\boldsymbol{\lambda}] = 0, \tag{2.5}$$

which can be solved for $\widehat{\boldsymbol{\mu}}$, i.e.,

$$\begin{aligned}\widehat{\boldsymbol{\mu}}(z) = \; & \left[\mathbf{I}_n - (\mathbf{D}'\mathbf{W}_H(z)\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}_H(z)\mathbf{P}(\mathbf{P}'\mathbf{W}_H(z)\mathbf{P})^{-1}\mathbf{P}'\mathbf{W}_H(z)\mathbf{D}\right]^{-1}(\mathbf{D}'\mathbf{W}_H(z)\mathbf{D})^{-1} \times \\ & \mathbf{D}'\mathbf{W}_H(z)\left[\mathbf{I}_{nT} - \mathbf{P}(\mathbf{P}'\mathbf{W}_H(z)\mathbf{P})^{-1}\mathbf{P}'\mathbf{W}_H(z)\right](\mathbf{y} - \mathbb{M}\{\mathbf{X}, \boldsymbol{\beta}(z)\}),\end{aligned} \tag{2.6}$$

and for $\widehat{\boldsymbol{\lambda}}$:

$$\begin{aligned}\widehat{\boldsymbol{\lambda}}(z) = \; & \left[\mathbf{I}_T - (\mathbf{P}'\mathbf{W}_H(z)\mathbf{P})^{-1}\mathbf{P}'\mathbf{W}_H(z)\mathbf{D}(\mathbf{D}'\mathbf{W}_H(z)\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}_H(z)\mathbf{P}\right]^{-1}(\mathbf{P}'\mathbf{W}_H(z)\mathbf{P})^{-1} \times \\ & \mathbf{P}'\mathbf{W}_H(z)\left[\mathbf{I}_{nT} - \mathbf{D}(\mathbf{D}'\mathbf{W}_H(z)\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}_H(z)\right](\mathbf{y} - \mathbb{M}\{\mathbf{X}, \boldsymbol{\beta}(z)\}).\end{aligned} \tag{2.7}$$

Define the two *local* within transformation matrices: $\mathbf{N}_H(z) \equiv \mathbf{I}_{nT} - \mathbf{P}[\mathbf{P}'\mathbf{W}_H(z)\mathbf{P}]^{-1}\mathbf{P}'\mathbf{W}_H(z)$

and $\mathbf{M}_H(z) = \mathbf{I}_{nT} - \mathbf{D}[\mathbf{D}'\mathbf{\Omega}_H(z)\mathbf{D}]^{-1}\mathbf{D}'\mathbf{\Omega}_H(z)$, where $\mathbf{\Omega}_H(z) \equiv \mathbf{N}'_H(z)\mathbf{W}_H(z)\mathbf{N}_H(z)$. Then, substituting $\widehat{\boldsymbol{\mu}}(z)$ and $\widehat{\boldsymbol{\lambda}}(z)$ from (2.6)–(2.7) for $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$, respectively, in the objective function in (2.3) yields a concentrated locally weighted least-squares problem from which both unknown fixed effects are removed:

$$\min_{\boldsymbol{\beta}(z)} \left(\mathbf{y} - \mathbb{M}\{\mathbf{X}, \boldsymbol{\beta}(z)\}\right)'\mathbf{\Sigma}_H(z) \left(\mathbf{y} - \mathbb{M}\{\mathbf{X}, \boldsymbol{\beta}(z)\}\right), \tag{2.8}$$

where $\mathbf{\Sigma}_H(z) \equiv \mathbf{M}_H(z)'\mathbf{\Omega}_H(z)\mathbf{M}_H(z)$. The riddance of fixed effects from the model is ensured by $\mathbf{N}_H(z)\mathbf{P}\boldsymbol{\lambda} = \mathbf{0}_{nT\times 1}$ and $\mathbf{M}_H(z)\mathbf{D}\boldsymbol{\mu} = \mathbf{0}_{nT\times 1}$ for all $z$. Upon a close examination of the local weighting matrix $\mathbf{\Sigma}_H(z)$, it is evident that $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ are removed via local $z$-specific kernel-weighted two-way within transformation. That is, the locally approximated model is transformed by subtracting the smoothed (around $\mathbf{z}_{it} = z$) version of cross-time and cross-individual averages and adding the smoothed pooled (grand) average. Note that the two-way within transformation implied by the smoothed LSDV procedure is performed *after* the local approximation is taken thereby removing $\mu_i$ and $\lambda_t$ "asymptotically" as opposed to purging them from (2.1) by directly within-transforming the model using unsmoothed (global) averages before deriving the estimator for functional coefficients.

To operationalize the estimator of $\boldsymbol{\beta}(z)$ from the profiled problem in (2.8), we rely on local-polynomial kernel approximators. For each $s = 1, \ldots, p$, the local Taylor expansion of the unknown functional coefficient around $\mathbf{z}_{it} = z$ is

$$\beta_s(\mathbf{z}_{it}) = \beta_s(z) + (\mathbf{z}_{it} - z)'\nabla_{\mathbf{z}}\beta_s(z) + (\mathbf{z}_{it} - z)'\nabla^2_{\mathbf{z}}\beta_s(z)(\mathbf{z}_{it} - z) + \ldots, \tag{2.9}$$

where $\nabla_{\mathbf{z}}\beta_s(z) = (\partial\beta_s(z)/\partial z_{1,it}, \ldots, \partial\beta_s(z)/\partial z_{q,it})'$ is a $q \times 1$ vector of first-order gradients and $\nabla^2_{\mathbf{z}}\beta_s(z) = \nabla_{\mathbf{z}'}(\nabla_{\mathbf{z}}\beta_s(z))$ is the $q \times q$ Hessian matrix of the second-order derivatives, etc. The first-order (local linear) approximation is arguably the most popular among practitioners (Henderson & Parmeter, 2015), and we adopt it here too and so do Sun et al. (2009). Thus, in what follows, we make use of $\beta_s(\mathbf{z}_{it}) \approx \beta_s(z) + (\mathbf{z}_{it} - z)\nabla_{\mathbf{z}}\beta_s(z)$ around $\mathbf{z}_{it} = z$.

Define a $(q+1) \times 1$ vector $\boldsymbol{\theta}_s(z) = (\beta_s(z), \nabla_{\mathbf{z}}\boldsymbol{\beta}_s(z)')'$ of unknown local parameters for each $s = 1, \ldots, p$. Then, the unknown $p\times(q+1)$ parameter matrix is defined as $\mathbf{\Theta}(z) = [\boldsymbol{\theta}_1(z) \ \ldots \ \boldsymbol{\theta}_p(z)]'$:

$$\mathbf{\Theta}(z) \equiv \begin{bmatrix} \boldsymbol{\theta}_1(z)' \\ \vdots \\ \boldsymbol{\theta}_p(z)' \end{bmatrix} = \begin{bmatrix} \beta_1(z) & \nabla\boldsymbol{\beta}_1(z)' \\ \vdots & \vdots \\ \beta_p(z) & \nabla\boldsymbol{\beta}_p(z)' \end{bmatrix} = \begin{bmatrix} \boldsymbol{\beta}(z) & \nabla\boldsymbol{\beta}(z)' \end{bmatrix},$$

where the first column of the above matrix is $\boldsymbol{\beta}(\cdot)$ evaluated at $z$ which is of primary interest. Next, define a $(q+1) \times 1$ vector of deviations from $z$, i.e., $\mathcal{Z}_{it}(z) = (1, (\mathbf{z}_{it} - z)')'$. For $\mathbf{z}_{it}$ close to $z$, we replace $\boldsymbol{\beta}(\mathbf{z}_{it})$ in (2.8) with $\mathbf{\Theta}(z)\mathcal{Z}_{it}(z)$ and obtain the local-linear estimator of functional coefficients $\boldsymbol{\beta}(z)$ from the locally concentrated minimization problem:

$$\min_{\mathbf{\Theta}(z)} \left(\mathbf{y} - \mathcal{X}(z)\text{vec}\{\mathbf{\Theta}(z)\}\right)'\mathbf{\Sigma}_H(z) \left(\mathbf{y} - \mathcal{X}(z)\text{vec}\{\mathbf{\Theta}(z)\}\right), \tag{2.10}$$

where we stack (by columns) the unknown parameter matrix $\mathbf{\Theta}(z)$ into a $p(q+1)\times 1$ vector denoted by the operator vec$\{\cdot\}$, and $\mathcal{X}(z) = (\mathcal{X}_i'(z), \ldots, \mathcal{X}_n'(z))'$ is an $nT \times p(q+1)$ data matrix, with each $T \times p(q+1)$ block given by

$$\mathcal{X}_i(z) = \begin{bmatrix} \mathcal{Z}_{i1}'(z) \otimes \mathbf{x}_{i1}' \\ \vdots \\ \mathcal{Z}_{iT}'(z) \otimes \mathbf{x}_{iT}' \end{bmatrix}.$$

Lastly, solving the first-order condition of (2.10) for the unknown $\mathbf{\Theta}(z)$ yields the following local-linear two-way fixed-effects estimator:

$$\text{vec}\left\{\widehat{\mathbf{\Theta}}(z)\right\} = \left(\mathcal{X}(z)'\mathbf{\Sigma}_H(z)\mathcal{X}(z)\right)^{-1}\mathcal{X}(z)'\mathbf{\Sigma}_H(z)\mathbf{y}. \tag{2.11}$$

A few remarks are in order. Since we assume that all elements in $\mathbf{x}_{it}$ are time-varying, the identification of $\boldsymbol{\beta}(\cdot)$ in our fixed-effects model does not require additional restrictions on $\mu_i$ and $\lambda_t$ such as the popular "zero sum" normalization à la Su & Ullah (2006) or Malikov et al. (2016) oftentimes imposed in semiparametric fixed-effects models.[4] The latter is normally necessary if the model admits time-invariant regressors, although one can sometimes achieve identification even without it: e.g., Sun et al. (2009) identify unknown smooth coefficient functions in their original one-way fixed-effects model that permits one time-invariant regressor by relying on the assumption that fixed effects are an *i.i.d.* sequence of random variables with a zero mean and finite variance. Similar arguments are used by Lin et al. (2014) for a fully nonparametric one-way fixed-effects model. However, even if the identification of $\boldsymbol{\beta}(\cdot)$ can be achieved without restricting fixed effects, to make the local-polynomial smoothed LSDV *estimation* of the functional coefficient on the time-invariant regressor (essentially, a nonparametric intercept function) feasible, one has to restrict the unobserved effects nonetheless. This happens because the within transformation is meant to remove *any* time-invariant term. The latter is the reason why Sun et al.'s (2009) one-way estimator is operationalized using the restricted matrix $\mathbf{D}$ under the zero-sum normalization of fixed effects. In our case however, we derive the estimator under no such restriction.[5] Having said that, should one be interested in allowing for a time-invariant $x$ in our two-way model, the estimator in (2.11) may be made feasible by replacing $\mathbf{D}$ and $\mathbf{P}$ with their restricted counterparts $\mathbf{D}_R = [-\mathbf{i}_{n-1}\mathbf{I}_{n-1}]' \otimes \mathbf{i}_m$ and $\mathbf{P}_\mathbf{R} = \mathbf{i}_n \otimes [-\mathbf{i}_{m-1}\mathbf{I}_{m-1}]'$ (under $\sum_{i=1}\mu_i = 0$ and $\sum_{t=1}\lambda_t = 0$) and accordingly redefining matrices $\mathbf{N}_H(z)$, $\mathbf{M}_H(z)$, $\mathbf{\Omega}_H(z)$ and $\mathbf{\Sigma}_H(z)$.

[4]The assumption of no time-invariant regressors in functional-coefficient fixed-effects models is not without precedent, e.g., see Rodriguez-Poo & Soberón (2014).

[5]Even though unknown parameter functions $\boldsymbol{\beta}(\cdot)$ in our model are identified without any normalization of fixed effects, clearly, the separable identification of $\{\mu_i\}$ and $\{\lambda_t\}$ does require a restriction. We are not concerned with the latter since the estimator removes both effects from the estimating equation.

Table 1. Simulation results for the two-way SLSDV estimator ($p = 1$, $q = 1$)

| | $T = 3$ | | | $T = 5$ | | |
|---|---|---|---|---|---|---|
| | $n = 50$ | $n = 100$ | $n = 200$ | $n = 50$ | $n = 100$ | $n = 200$ |
| | *Case 1: $x$ and $z$ are correlated* | | | | | |
| RMSE | 0.1763 | 0.1398 | 0.1063 | 0.1233 | 0.0981 | 0.0747 |
| MAE | 0.1237 | 0.0967 | 0.0713 | 0.0836 | 0.0637 | 0.0478 |
| | *Case 2: $x$ and $z$ are uncorrelated* | | | | | |
| RMSE | 0.1850 | 0.1459 | 0.1096 | 0.1274 | 0.1011 | 0.0788 |
| MAE | 0.1267 | 0.0989 | 0.0723 | 0.0849 | 0.0647 | 0.0488 |

Reported are the results for the functional coefficient estimator $\widehat{\beta}_1(\cdot)$.

## 3 Finite-Sample Performance

We first examine the finite-sample performance of the proposed estimator in (2.11) in a series of Monte Carlo experiments. All data generating processes follow model (2.1).

We begin with the DGP with $p = q = 1$ whereby $y_{it} = x_{it}\beta(z_{it}) + \mu_i + \lambda_t + u_{it}$, where the variables are drawn for all $i = 1, \dots, n$ and $t = 1, \dots, T$ as follows: $z_{it} = 0.5(\omega_{it} + \omega_{it-1})$, where $\omega_{it} \sim$ i.i.d. $\mathcal{U}(0, 0.5\pi)$; $x_{it} = 0.5(bz_{it} + x_{it-1}) + \zeta_{it}$, where $\zeta_{it} \sim$ i.i.d. $\mathcal{N}(0, 1)$; and $u_{it} \sim$ i.i.d. $\mathcal{N}(0, 0.5)$. We consider two cases: (1) $x_{it}$ and $z_{it}$ are correlated with $b = 1$ and (2) $x_{it}$ and $z_{it}$ are uncorrelated with $b = 0$. The outcome is generated with the following specification of individual and time effects: $\mu_i = c_1(\overline{z}_i + \overline{x}_i) + \rho_i$ with $\rho_i \sim$ i.i.d. $\mathcal{N}(0, 0.5)$ for all $i = 2, \dots, n$, and $\lambda_t = c_2(\overline{z}_t + \overline{x}_t) + \varrho_t$ with $\varrho_t \sim$ i.i.d. $\mathcal{N}(0, 0.5)$ for all $t = 2, \dots, T$, where $c_1$ and $c_2$ control the degree of correlation with regressors. We set $c_1 = c_2 = 0.5$ for "fixed" effects. The functional coefficient is $\beta(z_{it}) = \sin(\pi z_{it})$.

We consider cross-sectional sample sizes $n = \{50,\ 100,\ 200\}$ with the number of time periods $T = \{3,\ 5\}$. For each $(n, T)$, we simulate the model 500 times. We use the popular Silverman (1986) rule-of-thumb bandwidth for the smoothing variables. The kernel function of choice is second-order Gaussian. For each simulation, we compute the average (over $z_{it}$) root mean squared error (RMSE) and the average (over $z_{it}$) mean absolute error (MAE) for each functional coefficient function: $RMSE\big(\widehat{\beta}(\cdot)\big) = \sqrt{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\big[\widehat{\beta}(z_{it}) - \beta(z_{it})\big]^2}$ and $MAE\big(\widehat{\beta}(\cdot)\big) = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\big|\widehat{\beta}(z_{it}) - \beta(z_{it})\big|$, and then report their respective averages computed over 500 simulations in Table 1. The results in Table 1 are encouraging and indicate that, in both cases, the estimation of $\beta(\cdot)$ becomes more stable as the sample size increases for both fixed $T$ and fixed $n$. Both the RMSE and MAE decline significantly.

Next, we examine the performance of our estimator in higher-dimensional models. First, we consider the fixed-effects DGP with $p = 2$ regressors and $q = 1$ smoothing variable: $y_{it} = x_{1,it}\beta_1(z_{it}) + x_{2,it}\beta_2(z_{it}) + \mu_i + \lambda_t + u_{it}$, where $z_{it} = 0.5(\omega_{it} + \omega_{it-1})$; $x_{s,it} = 0.5(z_{it} + x_{s,it-1}) + \zeta_{s,it}$ for $s = 1, \dots, p$; $\mu_i = c_1(\overline{z}_i + 0.5\overline{x}_{1,i} + 0.5\overline{x}_{2,i}) + \rho_i$ and $\lambda_t = c_2(\overline{z}_t + 0.5\overline{x}_{1,t} + 0.5\overline{x}_{2,t}) + \varrho_t$ with $c_1 = c_2 = 0.5$. The random terms $\omega_{it}$, $\zeta_{s,it}$, $u_{it}$, $\rho_i$ and $\varrho_t$ are drawn as before. The functional coeffi-

Table 2. Simulation results for the two-way SLSDV estimator ($p = 2$, $q = 1$)

| | | $n = 50$ | $n = 100$ | $n = 200$ |
|---|---|---|---|---|
| $\beta_1(\cdot)$ | RMSE | 0.1798 | 0.1427 | 0.1046 |
| | MAE | 0.1202 | 0.0933 | 0.0685 |
| $\beta_2(\cdot)$ | RMSE | 0.1900 | 0.1489 | 0.1129 |
| | MAE | 0.1301 | 0.1015 | 0.0759 |

Reported are the results for the functional coefficient estimators $\widehat{\beta}_1(\cdot)$ and $\widehat{\beta}_2(\cdot)$. $T = 3$ throughout.

Table 3. Simulation results for the two-way SLSDV estimator ($p = 1$, $q = 2$)

| | $n = 50$ | $n = 100$ | $n = 200$ |
|---|---|---|---|
| RMSE | 0.3153 | 0.2592 | 0.2160 |
| MAE | 0.2047 | 0.1659 | 0.1351 |

Reported are the results for the functional coefficient estimator $\widehat{\beta}_1(\cdot)$. $T = 3$ throughout.

cients are specified as $\beta_1(z_{it}) = 1 + z_{it}^3/3$ and $\beta_2(z_{it}) = \sin(\pi z_{it})$. Table 2 summarizes these results. Second, we increase the number of variables that enter the model nonparametrically. The fixed-effects DGP with $p = 1$ regressor but $q = 2$ smoothing variables is $y_{it} = x_{it}\beta(z_{it})+\mu_i+\lambda_t+u_{it}$, where $z_{l,it} = 0.5(\omega_{l,it}+\omega_{l,it-1})$ for $l = 1, \ldots, q$; $x_{it} = 0.5(z_{q,it}+x_{it-1})+\zeta_{it}$; $\mu_i = c_1(0.5\overline{z}_{1,i}+0.5\overline{z}_{2,i}+\overline{x}_i)+\rho_i$ and $\lambda_t = c_2(0.5\overline{z}_{1,t}+0.5\overline{z}_{2,t}+\overline{x}_t)+\varrho_t$ with $c_1 = c_2 = 0.5$. The remaining random terms $\omega_{l,it}$, $\zeta_{it}$, $u_{it}$, $\rho_i$ and $\varrho_t$ are drawn as before. The functional coefficient is specified as $\beta(z_{it}) = 1 + z_{1,it}z_{2,it} + z_{2,it}^2$. The corresponding results are reported in Table 3. From Tables 2–3, we see that the estimator continues to perform well when the dimensionality of the model rises. As expected of a consistent estimator, RMSE declines with the increase in $n$.

**Empirical Application.** We also showcase the practical usefulness of our two-way SLSDV estimator by revisiting the role of management for production. Researchers in business economics have long argued that management is an important intangible input to production contributing significantly to productivity differences across firms and, when aggregated, across countries. However, until recently, the data on systematic measurements of management practices have been lacking. The World Management Survey developed and since expanded by Bloom & Van Reenen (2007) and their coauthors was the major step forward in this regard, and the empirical management literature has ever since been growing. Here, we revisit the estimation of management-augmented firm production functions from Bloom et al. (2017) that are aimed to measure the association between the firm's management practices and its production performance. Our focus is on relaxing the routinely assumed time-invariance of the relationship between management and firm output by letting the input elasticities, including that corresponding to the managerial input, smoothly vary with (discretized) time in an unspecified way. It is only natural to expect the role of management practices for firm production to change over time due to both the technological and institutional changes. By letting the production-function coefficients vary with time in an arbitrary nonparametric fashion, we are able to accommodate potential temporal variation in the relationship between management and output in a flexible way that is robust to misspecification. Out of space considerations, we relegate the details of our empirical application to the Supplementary Web Appendix.

## 4 Concluding Remarks

The existing semiparametric estimators for varying-coefficient fixed-effects models exclusively assume one-way fixed effects, typically in the dimension of cross-sectional units. We extend the Sun et al. (2009) estimator to the case of a functional-coefficient model with *two*-way fixed effects whereby we allow for unobservable heterogeneity in both dimensions of the data: cross-section and time. Both fixed effects are concentrated out of the model via locally smoothed two-dimensional within transformation. The estimator performs well in finite samples.

## References


Bloom, N., Sadun, R., & Van Reenen, J. (2017). Management as a technology? NBER Working Paper 22327.

Bloom, N. & Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *Quarterly Journal of Economics*, *122*(4), 1341–1408.

Henderson, D. J., Carroll, R., & Li, Q. (2008). Nonparametric estimation and testing of fixed effects panel data models. *Journal of Econometrics*, *144*, 257–275.

Henderson, D. J. & Parmeter, C. F. (2015). *Panel Data Econometrics*. Cambridge University Press.

Lin, Z., Li, Q., & Sun, Y. (2014). A consistent nonparametric test of parametric regression functional form in fixed effects panel data models. *Journal of Econometrics*, *178*, 167–179.

Malikov, E., Kumbhakar, S. C., & Sun, Y. (2016). Varying coefficient panel data model in the presence of endogenous selectivity and fixed effects. *Journal of Econometrics*, *190*(2), 233–251.

Qian, J. & Wang, L. (2012). Estimating semiparametric panel data models by marginal integration. *Journal of Econometrics*, *167*, 483–493.

Rodriguez-Poo, J. M. & Soberón, A. (2014). Direct semi-parametric estimation of fixed effects panel data varying coefficient models. *Econometrics Journal*, *17*, 107–138.

Rodriguez-Poo, J. M. & Soberón, A. (2015). Nonparametric estimation of fixed effects panel data varying coefficient models. *Journal of Multivariate Analysis*, *133*, 95–122.

Silverman, B. W. (1986). *Density Estimation*. London: Chapman and Hall.

Su, L. & Ullah, A. (2006). Profile likelihood estimation of partially linear panel data models with fixed effects. *Economics Letters*, *92*(1), 75–81.

Sun, Y., Carroll, R. J., & Li, D. (2009). Semiparametric estimation of fixed effects panel data varying coefficient models. *Advances in Econometrics*, *25*, 101–129.