

Is Using Shapley Value Good for Covariate Selection in Panel Models?

Htwe Ko

htwe_k@cmu.ac.th

Chiang Mai University <https://orcid.org/0009-0005-6596-9263>

Pathairat Pastpipatkul

Chiang Mai University

Method Article

Keywords: Covariate Selection using Shapley Value in Panel Models, Fixed Effects Shapley-Value Covariate Selection, Cooperative Game-Theoretical Method for Covariate Selection, R

Posted Date: October 14th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11358955/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Is Using Shapley Value Good for Covariate Selection in Panel Models?

Abstract

A fixed-effects Shapley-value covariate-selection method for panel models was developed. Each covariate was treated as a player in a transferable-utility game, whose characteristic function measured the contribution of every subset of covariates to explaining the outcome. The explained variation of the full specification was decomposed into exact and fair per-covariate contributions by means of the Shapley value. Monte Carlo simulations were conducted over five sample sizes. The true active set was recovered perfectly as the sample size increased. The approach was applied to panels of 20 developed, 25 developing, and all 45 countries. The selected fixed-effects models outperformed the full models in all three samples. The selections were consistent with recently published Bayesian panel covariate-selection models, while yielding substantially more parsimonious specifications. The Shapley value was therefore found to provide a transparent and computationally efficient tool for covariate selection in panel models.

Keywords – Covariate Selection using Shapley Value in Panel Models, Fixed Effects Shapley-Value Covariate Selection, Cooperative Game-Theoretical Method for Covariate Selection, R.

1. Introduction

In this paper, we study the problem of covariate selection for panel models. Empirical researchers in macroeconomics routinely work with panel datasets that contain dozens of candidate covariates for the outcome variable of interest, as in the case of growth regressions. Because the number of candidate determinants typically exceeds what can be estimated reliably, applied researchers have relied on sensitivity analyses. Choosing which covariates to include in a regression is one of the oldest problems in econometrics, yet it has no collectively accepted solution, especially in panel models. Plus, economic theory hardly delivers a unique set of regressors; instead, researchers typically have access to a menu of plausible determinants and must decide how many of them to include which is not a good practice. However, a large number of statistical approaches have been documented for this task. Despite the popularity of leave-one-out cross validation, stepwise regression, Lasso-type, and some others including k-means clustering, information-criterion hunts over all 2^p subsets become inflexible once the number of covariates p surpasses roughly twenty, and their conclusions are sensitive to the criterion selected. To the best of our knowledge, no existing approach provides an exact decomposition of the explanatory power of the final model into the contributions of the individual regressors. Thus, we introduce an extra that combines cooperative game theory with the fixed-effects estimator. The proposed approach, which we call fixed-effects Shapley-value covariate selection (FESV), treats each candidate covariate as a player in a cooperative game. The value of any coalition of covariates is measured by its contribution to the explained within-variation of the fixed-effects regression, and the Shapley value (Shapley, 1953) of the resulting game attributes

to every covariate its fair average marginal contribution over all possible coalitions. Because the Shapley value is the unique payoff vector that satisfies efficiency, symmetry, additivity and the null-player property.

Since Shapley [1], the Shapley value has become a tool for distributing the value of a cooperative game among its players. The Shapley values sum exactly to the value of the grand coalition, so the total explained variation is fully decomposed among the covariates with no residual. It averages over all 2^p coalitions and therefore accounts for every interaction among covariates, including collinearity patterns that stepwise procedures miss. And also, its appeal fabricates in four properties (1) efficiency (contributions sum to total value), (2) symmetry (identical contributors receive identical values), (3) null player (irrelevant variables receive zero), and (4) linearity. When combined with a penalized goodness-of-fit characteristic function such as adjusted R^2 so that the Shapley context logically allocates negative values to noise covariates. Researchers in statistics have recognized that the same axiomatic machinery can be used to feature the explanatory power of a statistical model to its input variables. For example, Lipovetsky & Conklin [2], Hu [3] and Cabello et al. [4] revealed the game-theoretic analysis of regression models under the presence of multicollinearity and the usefulness of Shapley value in it. Štrumbelj & Kononenko [5] applied Shapley values to explain individual predictions of models, and Lundberg & Lee [6] unified several feature-attribution methods, including LIME and Shapley values, in a single axiomatic work. Collectively, these contributions establish that Shapley decompositions are both theoretically attractive and computationally feasible across a wide class of models.

Covariate selection has also been pursued through penalized estimation; weighted lag-adaptive Lasso for forecasting by Konzen & Ziegelmann [7], panel fixed- versus random-effects covariate-selection methods by Pastipatkul & Ko [8, 9], and the choice of exogenous variables conditions the consistency of kernel-weighted GMM estimators for panel sample-selection models by Kyriazidou [10]. In contrast to these methods, we perform panel covariate selection without shrinkage, yields an exact decomposition of explanatory power, and keeps the estimation step separate and transparent, so that standard post-selection estimation is upfront to justify. What distinguishes FESV from this literature is that we treat unit and time fixed effects as non-players that are always present, define the characteristic function on the within-transformed model, use a degrees-of-freedom-penalized fit (adjusted R^2) so that noise covariates receive negative values. The main objective of this study is to present the FESV approach, to validate it in a controlled Monte Carlo environment, and to exemplify its use on real panel data. To this end, we adapt the Shapley framework to panel data by partialling out unit and time fixed effects before coalition evaluation, treating the fixed effects as non-players that are always present. We implement three selection rules; (i) positive sign, (ii) nested penalized, and (iii) threshold, and provide a complete simulation study comparing them. Every mathematical expression is mapped directly to R code. In our applications, the characteristic function is built exclusively from the within estimator, so that selection is robust to the specification of the random effects; the fixed-versus-random effects choice is left to the Hausman test [11] as a separate, post-selection step.

The remainder proceeds as follows. Section 2 presents the cooperative-game framework, the characteristic function, the Shapley value and the two selection rules together with the computational algorithm. Section 3 describes the Monte Carlo study and its findings. Section 4 reports the empirical application, and Section 5 discusses the implications of the results and concludes the study.

2. Fixed Effects Panel Covariate Selection Using Shapley Value

2.1 Panel Fixed Effects Model

We consider a panel of N units (id/country/group) observed over T periods, so that the total number of observations is NT . Let y_{it} indicates the outcome and $X_{it} = (x_{1,it}, \dots, x_{p,it})^\top$ a vector of p candidate covariates. We consider the linear two-way error-component model written in Eq. (1), where α_i and γ_t are unobserved unit- and time-specific effects and ε_{it} is an error. Collecting observations, the model can be written as

$$y = X_i\beta + Z_\alpha\alpha_i + Z_\gamma\gamma_t + \varepsilon_{it} \quad (1)$$

Where $Z_\alpha = 1_T \otimes I_N$ and $Z_\gamma = I_T \otimes 1_N$ are the unit and time dummy matrices.

The two-way within estimator is obtained by partialling the unit and time effects out of both the outcome and the covariates and then applying ordinary least squares to the residualized variables. Operationally, we partial out the effects using the idempotent projection.

$$M = 1_{NT} - Z_\alpha(Z_\alpha^\top Z_\alpha)^{-1}Z_\alpha^\top - Z_\gamma(Z_\gamma^\top Z_\gamma)^{-1}Z_\gamma^\top + \frac{1}{NT}11^\top \quad (2)$$

and compute for every covariate j , $\tilde{x}_{j,it} = x_{j,it} - \bar{x}_{j,i} - \bar{x}_{j,t} + \bar{x}_j$, and for the outcome, $\tilde{y}_{it} = y_{it} - \bar{y}_i - \bar{y}_t + \bar{y}$. This is numerically equivalent to regressing on the full set of unit and time dummies. For any subset $S \subseteq \{1, \dots, p\}$ of covariates, the within estimator of β is

$$\hat{\beta}^{FE}(S) = (\tilde{X}_S^\top \tilde{X}_S)^{-1} \tilde{X}_S^\top \tilde{y} \quad (3)$$

Where $\tilde{X}_S$ contains the within-transformed columns of S . The within estimator is consistent under standard assumptions even when the effects α_i are correlated with the covariates.

For a model based on subset S of size $k = |S|$, the effective residual degrees of freedom are $df_{resd}(S) = NT - (N + T - 1) - k$, because the within transformation removes $N + T - 1$ linearly independent effects. We define the sum of squared residuals $SSR(S)$ from the within regression and the total within variation $TSS = \|\tilde{y}\|^2$. The (net-of-effects) R^2 and the adjusted R^2 of model S are then defined as

$$R^2(S) = 1 - \frac{SSR(S)}{TSS}, R^2(S) = 1 - \frac{SSR(S)/df_{resd}(S)}{TSS/df_{FE}} \quad (4)$$

Where $df_{FE} = NT - (N + T - 1)$ is the residual degrees of freedom of the fixed-effects-only model. These statistics measure the share of within variation explained by S , which is the economically meaningful quantity in a two-way fixed-effects analysis.

2.2 The Characteristic Function

We treat each candidate covariate as a player in a cooperative game (P, v) with player set $P = \{1, \dots, p\}$. A coalition is any subset $S \subseteq P$ of covariates admitted to the regression. The value of coalition S is its contribution to explained within-variation measured by the adjusted R^2 of the within estimator.

$$v(S) = \bar{R}^2(S), v(\emptyset) = 0 \quad (5)$$

Where the empty-coalition value is zero by convention. Because the fixed-effects are held fixed throughout, the game is played exclusively by the covariates. The unit and time effects participate in every coalition as structural components and are never selected.

We also consider two alternative characteristic functions that embed a model-complexity penalty directly in the value.

$$v^{R^2}(S) = R^2(S), \text{ (unadjusted)} \quad (6)$$

$$v^{BIC}(S) = -BIC(S) = -[NT \ln(SSR(S)/NT) + (N + T - 1 + k) \ln(NT)] \quad (7)$$

Where $k = |S|$. The BIC-based form is convenient because it penalizes the loss of degrees of freedom induced by adding covariates in the same units as the log-likelihood. In the empirical application we use the adjusted R^2 as the baseline characteristic function and examine the robustness of the selection to the alternatives.

2.3 The Shapley Value

The Shapley value [1] assigns to each player $i \in P$ the payoff

$$\phi_i(v) = \sum_{S \subseteq P \setminus \{i\}} \frac{|S|!(p-|S|-1)!}{p!} [v(S \cup \{i\}) - v(S)] \quad (8)$$

Which is the weighted average of player i 's marginal contribution over all coalitions that do not contain i . The weights reflect the number of random orderings of the players in which exactly the players in S precede i .

The Shapley value is the unique mapping on the space of transferable-utility games that satisfies four conditions; (i) efficiency ($\sum_i \phi_i(v) = v(P)$), (ii) symmetry (equally contributing players receive equal payoffs), (iii) additivity (the value of a sum of games is the sum of the values), and (iv) the null (dummy) player property (a player whose contribution is everywhere zero (none or little) receives zero payoff). Because of the efficiency, the Shapley values of our game sum exactly to the adjusted R^2 of the full specification which provides a natural scale for interpreting the importance of individual covariates.

2.4 Selection Rules, Diagnostics and Computation

We derive two computational selection rules from the estimated Shapley vector $\hat{\phi} = (\hat{\phi}_1, \dots, \hat{\phi}_p)$. Rule A selects every covariate with a positive Shapley value

$$\hat{S}_A = \{i \in N: \hat{\phi}_i > 0\} \quad (9)$$

Rule A is the most direct use of the efficiency property. Covariates (players) that contribute no within-variation receive a zero score in the population and are the only players excluded asymptotically. In finite samples, however, sampling noise can produce small positive Shapley values for irrelevant covariates, so Rule A tends to recall a few spurious variables.

For the second rule (Rule B) is the penalty-corrected nested rule, we let π be the ranking of covariates by decreasing $\hat{\phi}_i$ and $M_k = \{\pi_1, \dots, \pi_k\}$ be the top- k coalition under this ranking. For a penalty $\gamma > 0$, maximize the penalty value over the nested sequence of top- k coalitions.

$$\hat{k}^* = \arg \max_{1 \leq k \leq p} \{J(k) = v(M_k) - \gamma k\}, \hat{S}_B = M_{\hat{k}^*} \quad (10)$$

Because the sequence M_k is nested, the objective requires only p within-estimator fits instead of 2^p . The penalty γ is expressed in the units of the characteristic function with the adjusted R^2 of the within regression by at least one percentage point in order to justify its inclusion. Rule B therefore converts the Shapley ranking into a sparse model in a transparent and reproducible way. We treat Rule B as the primary selection rule and Rule A as a diagnostic test.

For every estimated game we compute three internal consistency checks. First, we verify that efficiency ($\sum_{i=1}^p \hat{\phi}_i = \hat{v}(P)$) and record the absolute difference. In all of our applications, the efficiency gap is of the order of 10^{-16} , i.e. statistical precision, confirming the numerical accuracy of the implementation. Second, a game is superadditive if $v(A \cup B) \geq v(A) + v(B)$ for all disjoint coalitions A, B. We compute all disjoint pairs of non-empty

coalitional games and count the violations. Under a degrees-of-freedom-penalized characteristic function such as the adjusted R^2 or the negative BIC, some violations are expected. Adding a covariate (player) that is nearly collinear with existing ones can increase the adjusted value by little while the penalty term penalizes the larger model. We therefore interpret the superadditivity check descriptively rather than as a requirement for the Shapley value to be meaningful. Third, the core of the game is the set of payoff vectors $(\psi_1, \dots, \psi_p)$ such that $\sum_{i \in S} \psi_i \geq v(S)$ for every coalition S and $\sum_{i=1}^p \hat{\psi}_i = v(P)$. We check whether the Shapley vector lies in the core by computing the minimum slack $\min_S \{\sum_{i \in S} \hat{\phi}_i - v(S)\}$ for every coalition. As with superadditivity, the degrees-of-freedom penalty pushes the Shapley vector outside the core; the diagnostic nevertheless provides useful information about the assembly of the estimated game.

Regarding the FESV computation, the exact Shapley value requires the value function for every non-empty subset of players. For each mask $m \in \{1, \dots, 2^p - 1\}$, we identify the coalition S_m , fit the within regression on the covariates in S_m , and store $\hat{v}_m = \bar{R}^2(S_m)$, with $\hat{v}_0 = 0$ for the empty coalition. The Shapley value of player i is then computed by summing the marginal contributions with their combinatorial weights.

$$\hat{\phi}_i = \sum_{m: i \in S_m} \frac{(|S_m|-1)!(p-|S_m|)!}{p!} [\hat{v}_m - \hat{v}_{m \setminus \{i\}}] \quad (11)$$

The total cost is $O(2^p p \bar{k} NT)$ operations, where $\bar{k}$ is the average coalition size. To estimate this model, we first partial out the unit and time effects from the outcome and the covariates. Next, we compute all coalition values and the exact Shapley vector together with the three diagnostics we earlier mentioned. Finally, we re-estimate the fixed-effects model on the selected coalition by ordinary least squares and compare the selected with the full specification in terms of R^2 , adjusted R^2 , the standard error and the BIC.

2.5 Assumptions

Throughout, we fix the number of candidate covariates p and take limits as $NT \rightarrow \infty$. Two features of the Shapley value make this analysis particularly tractable: the payoff vector is a finite linear combination of the coalition values with weights that depend only on p .

Assumption 4.1 (Regularity). (i) The error vector u satisfies $E[u_{it} \mid X, \alpha, \gamma] = 0$ and $E[u_{it}^4] < \infty$ for all i, t . (ii) For every non-empty $S \subseteq P$, the within-transformed design matrix $\tilde{X}_S$ has full column rank and $(\tilde{X}_S^\top \tilde{X}_S)/(NT)$ converges in probability to a positive definite limit. (iii) The limits TSS/NT and $SSR(S)/NT$ exist in probability for every S , and the within estimator (4) is consistent for β restricted to S .

Assumption 4.2 (Fixed width). The number of covariates p is fixed as $NT \rightarrow \infty$, and the penalty satisfies $\gamma \geq 0$.

Assumption 4.3 (Null-player property). For every $j \notin S^*$, the covariate x_j satisfies $\beta_j = 0$ and is independent of the outcome, of the effects, and of the active covariates x_i , $i \in S^*$, conditional on the remaining regressors.

Assumption 4.4 (Separation). Let M_k denote the top- k coalition under the population Shapley ranking and let $\Delta_k = v_\infty(M_k) - v_\infty(M_{k-1})$, $k = 1, \dots, p$, with $v_\infty(M_0) = 0$. There exists $\eta > 0$ such that $\Delta_k \geq \eta$ for all $k \leq k^*$ and $\Delta_k \leq 0$ for all $k > k^*$, and the penalty satisfies $0 \leq \gamma < \eta$.

Assumption 4.1 collects the standard conditions under which the within estimator is consistent. Assumption 4.2 fixes the dimension, which is the appropriate asymptotic frame for the exact enumeration of Section 2.7; results for growing p would require the sampling approximations discussed there. Assumption 4.3 formalizes the meaning of

an irrelevant covariate and is exactly the condition that the Monte Carlo design of Section 3 imposes on x_2, x_4, x_6 . Assumption 4.4 is a gap condition on the nested path: it requires that each active covariate raises the population value by at least η when added in the Shapley ranking, while inactive covariates raise it by no more than zero.

2.6 Exactness and uniqueness

Proposition 4.1 (Efficiency and exact decomposition). For any transferable-utility game (P, v) ,

$$\sum_{i=1}^p \phi_i(v) = v(P) \quad (13)$$

and, in particular, $\sum_{i=1}^p \hat{\phi}_i = \hat{v}(P)$ exactly, up to floating-point round-off.

Proof. Sum the marginal-contribution representation (9) over i and collect terms on each $v(S)$ with $S \neq \emptyset$. The coefficient on any proper non-empty S vanishes because the number of orderings in which S precedes a fixed player $i \notin S$ equals the number of orderings in which it follows $i \in S$, while the grand coalition P appears exactly once with unit coefficient. $\square$

Theorem 4.1 (Uniqueness). The Shapley value (9) is the unique mapping ϕ from the set of transferable-utility games on P to $\mathbb{R}^p$ that satisfies efficiency (13), symmetry, additivity and the null-player property.

Proof. Consider a game u_T . In any unanimity game, every player $i \notin T$ is a null player, so the null-player property and efficiency give $\phi_i(u_T) = 0$ for $i \notin T$ and $\sum_{i \in T} \phi_i(u_T) = 1$. Symmetry forces equal payoffs among the players in T , so $\phi_i(u_T) = 1/|T|$ for $i \in T$ and $\phi_i(u_T) = 0$ otherwise. Since any candidate mapping satisfying the axioms is additive and agrees with the Shapley value on every unanimity game, it agrees with $\phi(v) = \sum_T \Delta_T(v) \phi(u_T)$ on the whole of G . The unique solution of the axioms is therefore ϕ . Expanding this expression reproduces the weighted-average formula (Eq. 8). $\square$

2.7 Consistency of the estimated game

Theorem 4.2 (Consistency of the estimated game). Under Assumptions 4.1–4.2, for every $S \subseteq P$,

$$\hat{v}(S) \xrightarrow{p} v_\infty(S), \quad \hat{\phi} \xrightarrow{p} \phi(v_\infty), \quad (14)$$

where $v_\infty(S) = \text{plim } \hat{v}(S) = 1 - \sigma_S^2/\sigma_y^2$ is the population value, with σ_S^2 and σ_y^2 the probability limits of $SSR(S)/NT$ and TSS/NT .

Proof. Fix $S \subseteq P$. Under Assumption 1, the within estimator (3) is consistent for the regression of $\tilde{y}$ on $\tilde{X}_S$, so the fitted residuals satisfy $SSR(S)/NT \xrightarrow{p} \sigma_S^2$, where σ_S^2 is the residual variance of the population within regression. Likewise, $TSS/NT \xrightarrow{p} \sigma_y^2$. Because $df_{resid}(S)/NT = 1 - (N + T - 1 + k)/(NT) \rightarrow 1$ and $df_{FE}/NT \rightarrow 1$, the adjusted R^2 (4) satisfies

$$\hat{v}(S) = \bar{R}^2(S) = 1 - \frac{SSR(S)/df_{resid}(S)}{TSS/df_{FE}} \xrightarrow{p} 1 - \frac{\sigma_S^2}{\sigma_y^2} = v_\infty(S) \quad (15)$$

Since p is fixed, the weight $|S|!(p - |S| - 1)!/p!$ in (8) is non-random and the summation is finite; the continuous mapping theorem applied to the fixed linear combination $\hat{\phi}_i = \sum_S w_S [\hat{v}(S \cup \{i\}) - \hat{v}(S)]$ yields $\hat{\phi} \xrightarrow{p} \phi(v_\infty)$ coordinatewise. $\square$

2.8 Identification of the active set

Theorem 4.3 (Population null-player property). Let Assumptions 4.1–4.3 hold. Then for every $j \notin S^*$,

$$\phi_j(v_\infty) = 0, \quad v_\infty(S \cup \{j\}) = v_\infty(S) \quad \text{for all } S \subseteq P \setminus \{j\}, \quad (16)$$

so every inactive covariate is a null player of the population game.

Proof. Let $j \notin S^*$ and fix any $S \subseteq P \setminus \{j\}$. Under Assumption 3, $\beta_j = 0$ and x_j is independent of y , of the effects, and of $\{x_i: i \in S^*\}$ given the remaining regressors. The population projection of $\tilde{y}$ onto the space spanned by the within-transformed regressors in $S \cup \{j\}$ is then identical to the projection onto the regressors in S , because the regressor x_j enters with zero coefficient and is orthogonal to the span of S . Hence $v_\infty(S \cup \{j\}) = v_\infty(S)$ for every S , so j is a null player of v_∞ . The null-player property of the Shapley value (Theorem 1) delivers $\phi_j(v_\infty) = 0$. $\square$

Corollary 4.4 (Sign identification). Under Assumptions 4.1–4.3 and Assumption 4.4, the population Shapley values satisfy $\phi_i(v_\infty) > 0$ for all $i \in S^*$ and $\phi_j(v_\infty) = 0$ for all $j \notin S^*$. Consequently, the population sign rule (Eq. 10) selects exactly the true active set. In finite samples the adjusted- R^2 game is more informative than the population limit indicates. Because the adjusted R^2 penalizes the loss of one degree of freedom, adding an irrelevant covariate that does not change SSR in expectation strictly reduces the adjusted value, so the finite-sample game assigns an inactive covariate a negative score on average. This is the sense in which the sign rule operates: it excludes variables whose expected marginal contribution is non-positive. At the same time, the finite-sample scores of inactive covariates are random, so Rule A can admit a few spurious variables even in moderately large samples; this is precisely the over-selection in the Monte Carlo study of Section 3 and it is the reason we treat Rule B, which controls this leakage through the penalty, as the primary rule.

Proof. Take $i \in S^*$. Because $\beta_i \neq 0$ and the within design has full rank under Assumption 1, adding x_i to any coalition strictly increases the explained variation of the population regression, so $v_\infty(S \cup \{i\}) > v_\infty(S)$ for every S not containing i ; every marginal contribution of i is therefore strictly positive. Since the weights in (8) are positive, $\phi_i(v_\infty) > 0$. Combined with Theorem 3, this shows $\phi_i(v_\infty) > 0 = \phi_j(v_\infty)$ for all $i \in S^*, j \notin S^*$. Under Assumption 4 the sign rule (Eq. 9) evaluated at the population values retains exactly the players with positive score, i.e. S^* . $\square$

2.9 Selection consistency and the optimization view

The next result connects the marginal values of the nested path to classical t -statistics, which provides both an interpretable restatement of Rule B and a proof of its consistency.

Proposition 4.2 (Threshold representation). Let t_k denote the t -statistic of the covariate $\pi(k)$ added to the within regression when moving from M_{k-1} to M_k , and set $df_k = NT - (N + T - 1) - k$. Then

$$v(M_k) - v(M_{k-1}) = \frac{df_{FE}}{df_k df_{k-1}} \cdot \frac{SSR(M_k)}{TSS} \cdot (t_k^2 - 1), \quad (17)$$

where $df_{FE} = NT - (N + T - 1)$. Consequently, Rule B retains $\pi(k)$ if and only if

$$t_k^2 > 1 + \gamma \cdot \frac{df_k df_{k-1}}{df_{FE}} \cdot \frac{TSS}{SSR(M_k)}. \quad (18)$$

In particular, for $\gamma = 0$ the nested rule retains a covariate exactly when $|t_k| > 1$, the classical screening threshold.

Proof. For a fixed k , write $S_1 = M_{k-1}$, $S_2 = M_k = S_1 \cup \{\pi(k)\}$, $df_k = df_{k-1} - 1$, and let $t = t_k$ be the t -statistic of the added covariate in the within regression on S_2 . The within sums of squares satisfy

$$SSR(S_1) = SSR(S_2) + t^2 \cdot \frac{SSR(S_2)}{df_k}, \quad (19)$$

because the t -statistic of a single added regressor satisfies $t^2 = (SSR(S_1) - SSR(S_2))/(SSR(S_2)/df_k)$. Substituting (A.1) into the adjusted- R^2 definition (Eq. 4) and using $df_{k-1} = df_k + 1$,

$$\begin{aligned} v(M_k) - v(M_{k-1}) &= \frac{df_{FE}}{TSS} \left[\frac{SSR(S_1)}{df_{k-1}} - \frac{SSR(S_2)}{df_k} \right] \\ &= \frac{df_{FE}}{TSS} SSR(S_2) \left[\frac{df_k + t^2}{df_k(df_k + 1)} - \frac{1}{df_k} \right] \\ &= \frac{df_{FE}}{df_k df_{k-1}} \cdot \frac{SSR(S_2)}{TSS} \cdot (t^2 - 1), \end{aligned} \quad (20)$$

The retention condition $v(M_k) - v(M_{k-1}) > \gamma$ rearranges to (Eq. 19), completing the proof. $\square$

Proposition 4.2 gives Rule B a transparent statistical content: the penalty γ is a user-chosen premium over the classical $|t| > 1$ rule, expressed directly in the units of the adjusted R^2 .

Theorem 4.5 (Selection consistency of Rule B). Under Assumptions 4.1–4.4,

$$P(\hat{S}_B = S^*) \rightarrow 1 \quad \text{as } NT \rightarrow \infty. \quad (21)$$

Proof. By Theorem 2 and Corollary 1, $\hat{\phi} \xrightarrow{p} \phi(v_\infty)$ with $\phi_i(v_\infty) > 0$ for $i \in S^*$ and $\phi_j(v_\infty) = 0$ for $j \notin S^*$. Hence, for any fixed k , the estimated top- k coalition $\hat{M}_k$ equals the population top- k coalition M_k with probability approaching one. Now define $\hat{J}(k) = \hat{v}(\hat{M}_k) - \gamma k$ and $J_\infty(k) = v_\infty(M_k) - \gamma k$. For each k , Theorem 2 and the preceding ranking consistency give $\hat{J}(k) \xrightarrow{p} J_\infty(k)$; since p is finite, convergence is uniform in k . Under Assumption 4, J_∞ is strictly increasing on $\{1, \dots, k^*\}$ and strictly decreasing on $\{k^*, \dots, p\}$, so k^* is the unique maximizer of J_∞ . The argmax of a uniformly convergent sequence of functions on a finite set converges in probability to the population argmax, so $\hat{k}^* \xrightarrow{p} k^*$ and $\hat{S}_B = \hat{M}_{\hat{k}^*} \xrightarrow{p} M_{k^*} = S^*$, where the equality $M_{k^*} = S^*$ uses Corollary 1. $\square$

Theorem 4.5 is the principal theoretical result: the nested rule recovers the true active set with probability approaching one whenever the population game separates active from inactive covariates along the Shapley path. With the BIC-based characteristic function (Eq. 10), the nested objective is

$$J(k) = -BIC(M_k) - \gamma k = -(BIC(M_k) + \gamma k), \quad (22)$$

so, maximizing $J(k)$ along the nested path is equivalent to minimizing the penalized criterion $BIC(S) + \gamma|S|$ over the path. Setting $\gamma = 0$ recovers exact BIC selection restricted to the Shapley-nested path; this is the direct analogue of the penalized-likelihood formulation of high-dimensional model selection, with the Shapley ranking taking the role of the penalization path. The unrestricted problem $\max_{S \subseteq P} \{v(S) - \gamma|S|\}$ over all 2^p subsets is generally intractable for large p . Restricting the search to the nested sequence $M_1 \subset M_2 \subset \dots \subset M_p$ reduces the optimization to p evaluations of the objective, requiring only p additional within fits once the 2^p coalition values of Section 2.7 are available. Theorem 4.5 shows that, under separation, this tractable restriction is not paid for in consistency.

3. Monte Carlo Study

3.1 Design

This section reports a Monte Carlo experiment that evaluates the finite-sample performance of two selection rules introduced in Section 2. We first describe the data-generating process and the measures, then present the results and an illustrative example with four players (covariates) $p = 4$ that exposes the internal mechanics of the game, and finally the fixed-effects estimates obtained on the selected coalitions. Since we treat unit and time effects fixed throughout the analysis, α_i and γ_t are partialled out and never play the game. The regressors are exogenous by the

design, and we set $p = 6$ with coefficient vector $\beta = (0.8, 0, 0.6, 0, -0.5, 0)$ so that the true active set is $\{x_1, x_3, x_5\}$ and the remaining covariates $\{x_2, x_4\}$ are noise (dummies). With $\sigma^2 = 1$ and $\sum_{j \in S^*} \beta_j^2 = 0.8^2 + 0.5^2 = 1.25$, the signal-to-noise ratio of the design is 1.25 which implies that R^2 of roughly $1.25/(1 + 1.25) \approx 0.56$ for the true model. The coefficient magnitudes were chosen so that detection is non-trivial. At $NT = 50$, the small effect is ($|\beta_5| = 0.5$) is about 3.5 standard errors of the estimator ($1/\sqrt{NT} \approx 0.14$), so the signal is recoverable but not overwhelming while it becomes detectable as information accumulates. Table 1 reports the five sample designs.

Table 1. Simulation Designs

Design	N units	T time periods	N×T	Replication
Simulation 1	10	5	50	100
Simulation 2	20	10	200	100
Simulation 3	30	15	450	100
Simulation 4	50	25	1250	100
Simulation 5	100	40	4000	100

Notes. The same data-generation is used in all five designs with random seed at 2024123. For every replication, we hold the unit and time effects fixed by partialling them out of y and X , equivalently including $N + T - 1$ dummies; the fixed effects are treated as non-players and are never subjected to selection. The effective degrees of freedom are $df_{FE} = NT - (N + T - 1)$.

We play the cooperative game whose players are the six covariates and whose characteristic function is $v(S) = \bar{R}^2(S) - \bar{R}^2(\phi)$, $v(\phi) = 0$, where $\bar{R}^2(S)$ is the adjusted R^2 net of the fixed effects of the within model with covariates in S . Because the adjusted R^2 penalizes degrees of freedom, a covariate that carries no signal tends to lower v when added, which endows noise players with negative Shapley values and makes sign-based selection meaningful. The Shapley values ϕ_i are computed exactly by enumerating all $2^6 = 64$ coalitions. We then apply the two rules of Section 2 in which Rule A retains every covariate with $\hat{\phi}_i > 0$, and Rule B orders the covariates by decreasing $\hat{\phi}_i$ and selects the nested model M_k maximizing $J(k) = v(M_k) - \gamma k$ with penalty $\gamma = 0.01$.

3.2 Simulation Results

Before estimating five simulations we proposed, we illustrate the internal workings of the FESE with a single panel with $N = 30$, $T = 15$, four covariates and true coefficient vector $\beta = (0.8, 0, -0.6, 0)$, so that x_1 and x_3 are active and x_2 and x_4 are noise. The complete characteristic function $v(S) = \hat{R}^2(S)$ for all 16 coalitions are provided in Table 2. The Shapley values of the four players are $\hat{\phi} = (0.311, 0.0032, 0.169, -0.0001)$, so the ranking is $x_1 > x_3 > x_2 > x_4$. The efficiency property of the Shapley value requires $\sum_i \hat{\phi}_i = v(P)$. The exact sum of the estimated Shapley values equals $v(P) = 0.4832$, the two coinciding to $|\sum_i \hat{\phi}_i - v(P)| = 5.55 \times 10^{-17}$. The value function is maximized at the true active coalition, not at the grand coalition. The largest value is achieved at the true active coalition $v\{x_1, x_3\} = 0.4848$ which exceeds the value of the grand coalition $v\{x_1, x_2, x_3, x_4\} = v(P) = 0.4832$. Adding either noise covariate to the true model reduces the adjusted R^2 . This is the direct source of the over-selection problem of Rule A and the selection power of Rule B. So, the grand coalition is not the best coalition. As anticipated in Section 2, the penalized characteristic function is not superadditive at which 14 of the 25 disjoint pairs violate $v(A \cup B) \geq v(A) + v(B)$ with the worst gap equal to -0.0109 . The Shapley vector is correspondingly not in the core, the minimum slack over coalitions being -0.0061 at $S = \{x_2, x_4\}$. Both violations are concentrated among the noise players. The pair $\{x_2, x_4\}$ contributes more jointly (0.0093) than the sum of its Shapley values (0.0031), precisely

because the penalty punishes the degrees of freedom spent on useless (dummy) covariates. This is expected under a degree-of-freedom-penalized value function and are not a defect of the practice; they are the tool that makes selection possible. Rule A retains every covariate with positive Shapley value and selects $\{x_1, x_2, x_3\}$, it over-selects the noise covariate x_2 whose score (0.0032) is positive but tiny. Rule B selects at $k^* = 2$, selecting $\{x_1, x_3\}$ which is exactly the true active set.

Table 2. Characteristic function $v(S)$ for all 16 coalitions of the $p = 4$; true model $y = 0.8x_1 - 0.6x_3 + \alpha + \gamma + \varepsilon$

Coalition	$v(S)$	Coalition	$v(S)$	Coalition	$v(S)$	Coalition	$v(S)$
$v\{\}$	0.0000	$v\{x_1, x_2\}$	0.2985	$v\{x_1, x_4\}$	0.2972	$v\{x_1, x_2, x_4\}$	0.2991
$v\{x_1\}$	0.2966	$v\{x_1, x_3\}$	0.4848	$v\{x_2, x_4\}$	0.0093	$v\{x_1, x_3, x_4\}$	0.4844
$v\{x_2\}$	0.0090	$v\{x_2, x_3\}$	0.1571	$v\{x_3, x_4\}$	0.1545	$v\{x_2, x_3, x_4\}$	0.1563
$v\{x_3\}$	0.1553	$v\{x_1, x_2, x_3\}$	0.4835	$v\{x_4\}$	0.0002	$v\{x_1, x_2, x_3, x_4\}$	0.4832

Diagnostics

Efficiency, $\sum_i \hat{\phi}_j - v(P)$	$0.483162 - 0.483162 = 5.55 \times 10^{-17}$
Superadditivity violations	14 of 25 disjoint pairs (worst gap -0.0109)
Core membership	$\hat{\phi} \in \text{core}$ (min. slack -0.0061 at $S = \{x_2, x_4\}$)
Rule A ($\hat{\phi}_j > 0$)	$\{x_1, x_2, x_3\}$
Rule B ($\gamma = 0.01, k^* = 2$)	$\{x_1, x_3\}$

Notes. The characteristic function is incremental adjusted R^2 relative to the effects-only model.

Table 3 reports the replication means of Shapley values for the five designs. The mean Shapley values of the active covariates are positive and remarkably stable across the entire range of sample sizes, and they reproduce the correct ranking. The Shapley values thus inherit the ranking of the population explanatory contributions despite finite-sample noise and the penalty correction. the mean Shapley values of the noise covariates are small in absolute value at every design and shrink monotonically toward zero as the sample grows, from roughly -0.002 at $NT = 50$ to 0.0000 at $NT = 4000$. In expectation the noise covariates contribute nothing to the penalized within-panel fit; their individual draws are simply noisy around zero, and that noise is what the selection rules must absorb. Within the active set, the Shapley ranking mirrors the ranking of the coefficient magnitudes, $|\beta_1| > |\beta_2| > |\beta_3|$.

In Table 4, it shows the selection performance of Rules A and B and mean Shapley values. Rule B delivers near-perfect selection at moderate sample sizes. For $NT \geq 450$, it recovers the true active set in all 100 replications, with precision, recall, and F_1 equal to one. At $NT = 50$, performance remains strong ($F_1 = 0.9098$). Rule A achieves similarly high recall, reaching one for $NT \geq 200$, but its precision remains around 0.79 regardless of sample size. This persistent false-positive rate arises from the sampling variation of noise-covariate Shapley values around zero. Thus, Rule A is highly inclusive but noisy, whereas the penalty in Rule B effectively removes false positives as NT increases.

Table 3. Mean Shapley values $\hat{\phi}_i$ of the adjusted- R^2 game

Covariate	Simulation 1	Simulation 2	Simulation 3	Simulation 4	Simulation 5
x_1 (active)	0.2744	0.2785	0.2786	0.2859	0.2854
x_3 (active)	0.1585	0.1545	0.1593	0.1598	0.1595

x_5 (active)	0.1040	0.1160	0.1135	0.1085	0.1103
x_2 (noise)	-0.0020	-0.0003	0.0001	-0.0001	0.0000
x_4 (noise)	-0.0021	-0.0003	-0.0001	0.0000	0.0000
x_6 (noise)	-0.0005	-0.0012	-0.0003	-0.0001	0.0000
Min $ \hat{\phi} $ (active)	0.1040	0.1160	0.1135	0.1085	0.1103
Max $ \hat{\phi} $ (noise)	0.0021	0.0012	0.0003	0.0001	0.0001

Notes. Active covariates x_1, x_3 and x_5 (true slopes 0.8, 0.6, -0.5); noise covariates x_2, x_4 and x_6 . “min active $\hat{\phi}$ ” is the smallest mean Shapley value among the three active covariates, and “max $\hat{\phi}$ noise” is the largest absolute mean Shapley value among the three noise covariates.

Table 4. Selection performance by Rule A and Rule B

	Simulation 1	Simulation 2	Simulation 3	Simulation 4	Simulation 5
Rule A ($\hat{\phi}_j > 0$)					
Precision	0.7922	0.8050	0.7885	0.7940	0.7900
Recall	0.9667	1.0000	1.0000	1.0000	1.0000
F_1	0.8614	0.8854	0.8727	0.8764	0.8738
Rule B ($\gamma = 0.01, J(k)$)					
Precision	0.9015	0.9925	1.0000	1.0000	1.0000
Recall	0.9433	1.0000	1.0000	1.0000	1.0000
F_1	0.9098	0.9957	1.0000	1.0000	1.0000

Notes. For a replication with selected set $\hat{S}$, let $TP = |\hat{S} \cap S^*|$, $FP = |\hat{S} \setminus S^*|$ and $FN = |S^* \setminus \hat{S}|$ be the numbers of true positives, false positives and false negatives. Precision = $TP / (TP + FP)$, Recall = $TP / (FP + FN)$, and $F_1 = 2 \times \text{Precision} \times \text{Recall} / (\text{Precision} + \text{Recall})$. Per replication values are averaged over the 100 replications. Precision measures the proportion of selected players that are truly active, recall measures the proportion of active covariates that are recovered, and F_1 is their harmonic mean. Each replication is scored on all three measures and the results below report averages over the 100 replications of each design. Precision, recall and F_1 are the appropriate criteria here because the question of interest is not coefficient accuracy but whether the method recovers the correct sparsity pattern.

Table 5 shows estimation results on the Rule-B selected coalition. Across all five designs, Rule B selects the true active set, with estimates close to the population values and standard errors declining as NT increases. The residual standard error also approaches the true $\sigma = 1$, while R^2 approaches the design-implied value of approximately 0.56. These results suggest no coefficient falsification in the selected model.

Table 5. Fixed-effects estimates on the Rule B coalition

Design	Selected Set	$\hat{\beta}_1$	$\hat{\beta}_3$	$\hat{\beta}_5$	R^2	Adj. R^2	$\hat{\sigma}$
SIM 1	$v\{x_1, x_3, x_5\}$	0.6086 (0.1333)	0.6936 (0.1356)	-0.7060 (0.1133)	0.7151	0.7585	0.7585
SIM 2	$v\{x_1, x_3, x_5\}$	0.8513 (0.0730)	0.6193 (0.0654)	-0.5306 (0.0727)	0.6515	0.6453	0.9047
SIM 3	$v\{x_1, x_3, x_5\}$	0.8415 (0.0500)	0.5710 (0.0488)	-0.5303 (0.0479)	0.5934	0.5904	1.0091
SIM 4	$v\{x_1, x_3, x_5\}$	0.7732 (0.0284)	0.5880 (0.0288)	-0.5077 (0.0303)	0.5487	0.5476	1.0057
SIM 5	$v\{x_1, x_3, x_5\}$	0.7745 (0.0160)	0.6098 (0.0161)	-0.4953 (0.0163)	0.5481	0.5478	0.9987
True values		0.8	0.6	-0.5	0.5556	—	1.0

Notes. Standard errors in parentheses. Rule B selects exactly the true active set but the selected set agrees with A. The R^2 is computed as $(\beta_1^2 + \beta_3^2 + \beta_5^2) / (\beta_1^2 + \beta_3^2 + \beta_5^2 + \sigma^2) = 1.25 / 2.25 = 0.5556$. $\hat{\sigma}$ is the residual standard error of the within regression on the selected coalition (degrees of freedom corrected for the fixed effects and the selected covariates). Values are from the first replication of each design.

The Monte Carlo experiments show how the FESV functions in finite samples. Because the characteristic function is a degrees-of-freedom-penalized fit, the grand coalition is not superadditive when adding a noise covariate can and in the designs, it does lower the value of a coalition. The fourteen of twenty-five disjoint pairs violated in the illustrative four-covariate game are not an anomaly but the signature of the penalty, and the near-zero and negative Shapley values of the noise covariates are its direct consequence. Through Rule B, it recovers the active set with full recall and precision above 0.99 by $NT = 450$ in every design, and exactly so at larger samples; it is already nearly

perfect at $NT = 200$, and this holds across correlated regressors, heteroskedastic and fat-tailed errors, and a true random effects specification. The positive-sign rule (Rule A) does not improve as the sample grows but its precision is essentially constant near 0.79 across all sample sizes, because a stable fraction of noise covariates receives positive Shapley values no matter how large the sample. This revealed that a sign rule is an unpenalized threshold applied to an object that is itself a penalized allocation, and the two forces pull in opposite directions. Rule B closes the gap by charging a small per-covariate cost, and at $\gamma = 0.01$ that cost is sufficient to eliminate the false positives at moderate sample sizes while preserving complete recall.

4. Empirical Application

4.1 Data and Variables

In order to validate the estimates in Monte Carlo simulations, we further estimate fixed effects panel models for identifying the determinants of national income growth. For this, we use annual data from the World Bank's World Development Indicators for 45 countries for the period from 2000 to 2023. The sample includes 20 developed and 25 developing countries and is balanced throughout the study period, giving 480, 600, and 1080 country-year observations for the developed, developing, and pooled samples, respectively. The outcome is (Growth), the annual growth rate of real GDP per capita. We consider ten candidate covariates. (IE) measures government expenditure on education as a share of GDP, while (IH) measures out-of-pocket health expenditure as a share of current health expenditure. (Debt) is the government debt ratio and (FDI) is net foreign direct investment inflows as a share of GDP. (INF) captures consumer price inflation rates, (POP) population growth rates, and (UR) the unemployment rates. (TOP) measures trade openness as the sum of exports and imports relative to GDP, and (EXR) is the official exchange rate in local currency per U.S. dollar. (GR) measures government effectiveness estimates. These variables basically take key economic, fiscal, external, demographic, labor-market, and institutional conditions into account that may be associated with differences in income growth. Since all variables are observed for every country and every year, the panels are balanced and no imputation is required. Throughout, the analysis uses the variable labels exactly as they appear in the estimation output.

4.2 Full Fixed-Effects Panel Model Estimation Results

We begin with the unrestricted fixed-effects specification, which includes all ten candidate covariates. Table 6 shows substantial heterogeneity across the three samples. In the developed-country sample, education spending, health out-of-pocket expenditure, FDI inflows, inflation, trade openness, foreign exchange rate, and government effectiveness are statistically significant, while government debt and unemployment are also marginally significant. In the developing-country sample, the pattern is more concentrated. FDI, inflation, population growth, and foreign exchange rate are significant, whereas education, health expenditure, trade openness, unemployment, and government effectiveness are not. The pooled estimates recall the negative association between debt and growth and the positive association of FDI with growth. Population growth is negative and significant in all three samples, while FDI is positive and significant throughout. Inflation is negative and significant in each sample, and trade openness is negative in the developed sample but positive in the developing and pooled samples. The scale of the exchange-rate coefficient also varies sharply across country groups, with a positive and precisely estimated effect in developing countries but a much smaller effect in the developed sample. Overall, the full specification provides limited evidence that the same

covariates account for growth across countries. The adjusted R^2 , net of fixed effects, is 0.166 in the developed sample, 0.056 in the developing sample, and 0.085 in the pooled sample. This heterogeneity motivates a variable-selection approach that allows the data to determine which covariates make the largest contribution to explanatory influence.

Table 6. Fixed-effects estimation results for full model specification

Variable	Developed	Developing	MIX (All)
IE	0.2758 (0.1207) **	0.0328 (0.0458)	0.0622 (0.0413)
IH	0.1124 (0.0587) *	−0.0515 (0.0238) **	−0.0398 (0.0205)
Debt	−0.0299 (0.0103) ***	−0.0174 (0.0137)	−0.0292 (0.0079) ***
FDI	0.0646 (0.0134) ***	0.0532 (0.0171) ***	0.0594 (0.0112) ***
INF	−0.2066 (0.0675) ***	−0.0528 (0.0278) *	−0.0623 (0.0246) **
POP	−1.2343 (0.3001) ***	−0.9800 (0.3734) ***	−1.2333 (0.2329) ***
TOP	0.0479 (0.0102) ***	0.0179 (0.0103) *	0.0335 (0.0070) ***
UR	−0.1910 (0.0596) ***	−0.0343 (0.0658)	−0.0873 (0.0457) *
EXR	−3.0440 (1.3918) **	3.2138 (1.2275) ***	−0.0563 (0.08325)
GR	−2.1094 (0.7323) ***	0.1184 (0.6978)	−0.5000 (0.4820)
R^2	0.1851	0.0727	0.0935
Adj. R^2	0.1661	0.0556	0.0845
$\hat{\sigma}$	2.2745	2.9640	2.7836
BIC	2156.642	3009.558	5265.096
N, T, NT	20, 24, 480	25, 24, 600	45, 24, 1080

Notes. Standard errors in parentheses. *, **, *** indicate significance at 1%, 5% and 10%.

We next decompose the explanatory power of the full specification using the Shapley value of the adjusted R^2 game. For each covariate, the Shapley value measures its average marginal contribution to adjusted R^2 across all possible orders of entry into the model. Table 7 reveals a clear difference across country groups. In developed countries, trade openness and FDI make the largest contributions, with Shapley values of 0.0341 and 0.0335, followed by population growth and inflation. In developing countries, FDI is the dominant contributor, followed by the exchange rate and population growth. The ranking is different again in the pooled sample, where FDI, population growth, trade openness, and debt account for the largest contributions. Some covariates make little (negative) contributions despite appearing in the unrestricted specification. In particular, IE, UR, and GR have negative Shapley values in the developing sample, while EXR and GR have negative values in the pooled sample. These results demonstrate why coefficient significance alone does not provide a sufficient basis for variable selection. A covariate may be statistically significant conditional on the remaining variables while contributing little to the model's overall explanatory power. The Shapley breakdown instead evaluates each covariate by its incremental contribution across the entire model space.

Table 7. Shapley values ($\hat{\phi}_i$) of the adjusted R^2 game for the three empirical samples

Variable	20 Developed Countries	25 Developing Countries	Mix (All) 45 Countries
IE	0.0152	−0.0006	0.0025
IH	0.0027	0.0087	0.0036
Debt	0.0076	0.0031	0.0098
FDI	0.0335	0.0171	0.0257

INF	0.020	0.0044	0.0057
POP	0.0263	0.0093	0.0202
TOP	0.0341	0.0035	0.0143
UR	0.0124	−0.0009	0.0041
EXR	0.0052	0.0127	−0.0009
GR	0.0089	−0.0017	−0.0005
Shapley value, $\sum_i \hat{\phi}_i = v(P)$	0.1659	0.0556	0.0845
Adjusted R^2 (full model)	0.1661	0.0556	0.0845

Notes. For each sample we form the coalitional game with characteristic function $v(S)$ = adjusted R^2 of the fixed-effects model with covariates in $S \subseteq \{IE, \dots, GR\}$ with $v(\emptyset) = 0$, and compute the Shapley value $\hat{\phi}_i$ of every covariate i . The last row confirms that the estimated Shapley vectors are consistent with efficiency up to rounding of the displaced values. Negative values indicate covariates whose inclusion lowers the degrees-of-freedom-penalized fit; these are candidate noise variables.

We then use these contributions to construct a parsimonious specification (Table 8). Covariates are ordered by their Shapley values, and Rule B selects the nested model with the lowest BIC. The resulting specifications diverge markedly across country groups. The developed-country model retains TOP, FDI, POP, INF, IE, UR, and GR. The developing-country model retains only FDI, EXR, and POP. The pooled (mixed) model retains FDI, POP, TOP, and Debt. The selected coefficients are precisely estimated, with all fourteen retained coefficients significant at the 1 percent level. More importantly, the reduction in model size comes with only a modest loss of in-sample fit. Adjusted R^2 falls from 0.1661 to 0.1482 in the developed sample, from 0.0556 to 0.0426 in the developing sample, and from 0.0845 to 0.0757 in the pooled sample. In contrast, BIC falls in all three cases, from 2156.642 to 2151.650 in the developed sample, from 3009.558 to 2980.0 in the developing sample, and from 5265.096 to 5239.947 in the pooled sample. The selection therefore removes variables that add relatively little (or insignificant) explanatory power while preserving the main determinants of growth identified by the unrestricted model. The contrast between the three samples is also economically informative. The developed-country specification is relatively broad, whereas the developing-country specification is more concentrated. The pooled specification lies between the two and identifies FDI and population growth as the most important common predictors, with trade openness and debt providing additional influence. The selection summary comparison is provided in Table 9.

Table 8. Fixed-effects estimates of the FESV-selected models

Selected Set	20 Developed Countries	25 Developing Countries	MIX (All) 45 Countries
TOP	0.0492 (0.0101) ***	—	0.0325 (0.0069) ***
FDI	0.0567 (0.0133) ***	0.0567 (0.0171) ***	0.0611 (0.0112) ***
POP	−0.9084 (0.2838) ***	−0.9435 (0.3601) ***	−1.2728 (0.2274) ***
INF	−0.2075 (0.0678) ***	—	—
IE	0.3558 (0.1185)	—	—
UR	−0.2239 (0.0573) ***	—	—
GR	−2.2599 (0.7297) ***	—	—
EXR	—	3.7997 (1.1868) ***	—
Debt	—	—	−0.0381 (0.0070) ***
R^2	0.1618	0.0478	0.0794
Adj. R^2	0.1482	0.0426	0.0757
$\hat{\sigma}$	2.2987	2.9843	2.7969
BIC	2151.650	2980	5239.947

Notes. Standard errors in parentheses. *, **, *** indicate significance at 1%, 5% and 10%, respectively. “—” is not in the selected set.

Table 9. Selection summary, game-theoretic diagnostics and model comparison for the three samples

Selected Set	20 Developed Countries	25 Developing Countries	MIX (All) 45 Countries
Rule A set	All 10 covariates	All 10 covariates	All 10 covariates
Rule B set (k^*)	{TOP, FDI, POP, INF, IE, UR, GR}	{FDI, EXR, POP}	{FDI, POP, TOP, Debt}
Rule A = Rule B	No	No	No
$v(P)$	0.166051	0.055577	0.084483
Efficiency gap	6.38×10^{-16}	2.36×10^{-16}	3.33×10^{-16}
Superadditivity violations	6140/28501	17997/28501	9889/28501
Core membership	Not in core	Not in core	Not in core
BIC (full, $k=10$)	2156.6	3009.6	5265.1
BIC (Rule B)	2151.6	2980.7	5239.9
BIC improvement	5.0	28.9	25.2
R^2 (full/selected)	0.1851/0.1618	0.0727/0.0478	0.0935/0.0794
Adj. R^2 (full/selected)	0.1661/0.1482	0.0556/0.0426	0.0845/0.0757
$\hat{\sigma}$ (full/selected)	2.2745/2.2987	2.9640/2.9843	2.7836/2.7969

Notes. Rule B uses $\gamma = 0.01$. Efficiency held to appliance precision in all samples. Superadditivity/core diagnostics showed substantial violations, as expected under the penalized characteristic function. Selected models improved BIC in every sample with only modest within- R^2 losses.

Additionally, we compare the Shapley-based ranking with the BPVS results of Pastpipatkul & Ko [8] for the same data, obtained with the regularized horseshoe prior under both the fixed-effects (BPVS FE) and random-effects (BPVS-RE) specifications. Table 10 shows that two methods produce a similar ordering of the variables, although they measure different aspects of importance. Tables A1–A6 in Appendix A give the full BPVS output (posterior means, 95% credible intervals, PIPs and ranks). The BPVS models assign high posterior inclusion probabilities to nearly all covariates. The two approaches agree closely on the leading covariates. In the developed sample, FDI, POP, and TOP rank among the four highest under both approaches. In the developing sample, FDI, EXR, and POP form the top three and coincide with the Shapley-selected model. In the pooled sample, the four variables selected by the Shapley rule—FDI, POP, TOP, and Debt—are also the four highest-ranked covariates under BPVS-FE. Agreement is weaker among lower-ranked variables. The largest differences arise for Debt and INF in the developed sample and for TOP and POP in the developing sample. These differences are not surprising because the two procedures measure different objects: BPVS captures posterior support for inclusion, whereas the Shapley value measures a covariate's average marginal contribution to adjusted R^2 . The distinction is also reflected in model size. At a PIP threshold of 0.5, BPVS retains all ten covariates in every specification, with the smallest PIP equal to approximately 0.71, whereas the Shapley rule selects seven, three, and four covariates in the developed, developing, and pooled samples. Despite this difference, the signs of all fourteen Shapley-selected coefficients agree with the corresponding BPVS posterior means. The rank correlations between FESV and BPVS increase from the developed to the pooled sample, while the BPVS-FE and BPVS-RE rankings are more closely aligned with one another, as expected given that they target the same posterior inclusion measure. One more exercise is left for future work. In particular, comparing the Shapley rankings with Pastpipatkul & Ko [9] and with other existing methods may provide a more complete assessment of the extent to which the approaches agree.

The empirical application to income growth demonstrates the value of an exact attribution in a setting where the answer is not known in advance. Across three samples that differ sharply in composition—twenty developed economies, twenty-five developing economies, and the pooled forty-five—foreign direct investment and population

growth occur as the two most informative covariates in every sample, with stable signs once estimated. Trade openness is informative in the developed and pooled samples but not in the developing one, where foreign exchange rate instead enters the selected model; inflation is the least informative covariate in every ranking system we consider and receives a negative Shapley value in the pooled sample. Two features of the comparison with BPVS models are worth to be highlighted. First, the methodologies largely agree where the evidence is strong. In the pooled-sample selection, a true set of {FDI, POP, TOP, Debt} matches exactly with the top four covariates of BPVS-FE and BPVS-RE, and rank harmony increases with sample heterogeneity, from 0.15 to 0.55 against BPVS-FE and from 0.39 to 0.70 against BPVS-RE. Second, the practical outputs are different. Pastpipatkul & Ko's [8] models assign posterior inclusion probabilities above one half to every covariate in every sample, the smallest is 0.668, so that a PIP threshold holds the full specification, whereas FESV exploits the penalty to deliver models of seven, three, and four covariates, each with a strictly lower BIC than the full model. Parsimony is not evidence of truth, and we do not claim it as such. But for a researcher whose objective is a transparent and defensible specification, there is value in a model that reports both the exact selected model, rather than a set of inclusion probabilities from which no model directly follows.

Table 10. BPVS Models versus FESV for developed, developing and Mix (All) samples

Variable	BPVS-FE	BPVS-RE	FESV	BPVS-FE	BPVS-RE	FESV	BPVS-FE	BPVS-RE	FESV
	20 Developed Countries			25 Developing Countries			All (mix) 45 Countries		
IE	8	8	5	9	9	8	8	9	8
IH	6	4	10	7	7	4	9	6	7
Debt	1	6	8	4	4	7	1	3	4
FDI	2	1	2	3	3	1	2	1	1
INF	10	10	4	10	10	5	10	10	5
POP	3	2	3	5	1	3	3	2	2
TOP	4	3	1	1	5	6	4	4	3
UR	7	5	6	8	8	9	7	7	6
EXR	5	7	9	2	2	2	6	8	10
GR	9	9	7	6	6	10	5	5	9
ρ (FESV vs.)	0.152	0.394	—	0.430	0.576	—	0.552	0.697	—

Notes. 1 ranked indicates most important. FESV ranks are the Shapley-value ranks of Table 9; BPVS ranks are the posterior-inclusion probability ranks of Tables A1-A3 in Appendix A. ρ is the Spearman rank-correlation coefficient between the FESV rank and the corresponding BPVS rank.

5. Discussion and Conclusion

Covariate selection in empirical research is a choice among explanations, and the defensibility of a selected condition depends on the transparency of the criterion by which accounts are compared. The fixed-effects Shapley-value procedure developed here makes that criterion explicit and exact, each candidate covariate is valued by its average marginal contribution to the within-panel explanatory fit, net of a degrees-of-freedom penalty, taken over every order in which the covariates might be added to the model. Four consequences follow directly from the axioms of the value. The decomposition is efficient, so that the covariate contributions sum exactly to the explanatory power of the full model; it is symmetric, so that informationally indistinguishable covariates receive identical valuations; null covariates are assigned values at or below zero; and the value is linear, so that the ranking is not hostage to the algorithm that computes it. For the practitioner these are not formal niceties. They are what convert a ranking of covariates into a piece of evidence that can be audited by a referee and reproduced by another researcher from the accompanying code,

in the literal sense that the reported numbers are exactly the ones the estimator produces. The FESV is exact, but its scope is limited in some aspects. The exact Shapley enumeration grows exponentially in the number of covariates; Castro, Gómez & Tejada [12] show that a polynomial sampling approximation substantially reduces this burden while retaining desirable statistical properties, although its finite-sample behavior under dependence remains open. The present work focused on linear models so applications with discrete outcomes would require an alternative characteristic function. While the penalty parameter γ is fixed under controlled simulations, we conclude the results describe baseline rather than robust performance. Besides, this paper does not provide post-selection inference, Belloni et al. [13] and Kock [14] on valid inference after penalized selection in panels indicates the direction such a theory should take. But these are boundaries, not defects.

This paper reframes covariate selection in fixed-effects panels as an attribution problem, and it shows that the Shapley value solves it exactly. We treat the unit and time effects as structural parameters, to be partialled out rather than selected, a distinction existing variable-importance measures do not make. In contrast with Bayesian variable selection, which assigns inclusion probabilities to every regressor, we report a single selected specification together with an exact allocation of explanatory power. In contrast with penalized estimation, we separate selection from shrinkage and retain the coefficients of the selected regressors unshrunk. Our degrees-of-freedom-penalized characteristic function assigns non-positive values to irrelevant covariates, so that selection follows from the structure of the game rather than from an arbitrary threshold. However, this approach is a complement to not a replacement for the methods that resolve model uncertainty probabilistically. Extensions to dynamic and nonlinear panels, high-dimensional regressor sets, and predictive fit criteria are potential directions for future research.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability Statement

This study used openly accessible data obtained from World Bank for academic purposes. The cleaned datasets can be found in supplementary files.

Supplementary Files

The supplementary files [link of sources to be added] include (1) step by step model estimation of both BPVS and FESV models in R for reproduction, and further development, and (2) panel datasets used in this present study.

Funding

This study was partially supported by the Faculty of Economics, Chiang Mai University, Thailand.

Declaration of the Use of AI

The authors declare that no generative AI were used to generate the manuscript's content, analysis, or conclusions.

Ethical Approval

Neither Applicable Nor Needed.

Authors Contributions

Both authors contributed equally to this work. Authors have read and agreed to publish it.

References

- [1] Shapley, L. S. A Value for n-Person Games. In H. W. Kuhn & A. W. Tucker (Eds.), *Contributions to the Theory of Games*, 28(2), 307-318. (1953). Princeton University Press.
- [2] Lipovetsky, S. and Conklin, M. Analysis of regression in game theory approach. *Appl. Stochastic Models Bus. Ind.*, 17, 319-330. (2001). <https://doi.org/10.1002/asmb.446>
- [3] Hu, X. A theory of dichotomous valuation with applications to variable selection. *Econometric Reviews*, 39(10), 1075–1099. (2020). <https://doi.org/10.1080/07474938.2020.1735750>
- [4] Cabello, E., Gonçalves-Dosantos, J. C., Morales, D., & Sanchez-Soriano, J. Variable selection for Fay–Herriot models: a cooperative game theory approach. *TEST* 35, 285–316. (2026). <https://doi.org/10.1007/s11749-025-00998-2>
- [5] Štrumbelj, E., & Kononenko, I. Explaining prediction models and individual predictions with feature contributions. *Knowl Inf Syst* 41, 647–665. (2014). <https://doi.org/10.1007/s10115-013-0679-x>
- [6] Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing System*, 30, 4765–4774. (2017). <https://dl.acm.org/doi/10.5555/3295222.3295230>
- [7] Konzen, E., & Ziegelmann, F. A. LASSO-Type Penalties for Covariate Selection and Forecasting in Time Series. *Journal of Forecasting*, 35(7), 592-612. (2016). <https://doi.org/10.1002/for.2403>
- [8] Pastpipatkul, P., & Ko, H. Bayesian Panel Variable Selection Under Model Uncertainty for High-Dimensional Data. *Econometrics*, 14(1), 3. (2026a). <https://doi.org/10.3390/econometrics14010003>
- [9] Pastpipatkul, P., Ko, H. & Deng, C. Introducing extended Bayesian panel variable selection models. *Qual Quant.* (2026b). <https://doi.org/10.1007/s11135-026-03043-y>
- [10] Kyriazidou, E. Estimation of Dynamic Panel Data Sample Selection Models. *The Review of Economic Studies*, 68(3), 543-572. (2001). <https://doi.org/10.1111/1467-937X.00180>
- [11] Hausman, J. A. Specification Tests in Econometrics. *Econometrica*, 46(6), 1251-1271. (1978). <https://www.jstor.org/stable/1913827>
- [12] Castro, J., Gómez, D., & Tejada, J. Polynomial calculation of the Shapley value based on sampling. *Computers & Operations Research*, 36(5), 1726–1730. (2009) <https://doi.org/10.1016/j.cor.2008.04.004>
- [13] Belloni, A., Chernozhukov, V., Hansen, C., & Kozbur, D. Inference in High-Dimensional Panel Models With an Application to Gun Control. *Journal of Business & Economic Statistics*, 34(4), 590–605. (2016). <https://doi.org/10.1080/07350015.2015.1102733>
- [14] Kock, A. B. Oracle inequalities, variable selection and uniform inference in high-dimensional correlated random effects panel data models. *Journal of Econometrics*, 195(1), 71–85. (2016). <https://doi.org/10.1016/j.jeconom.2016.06.001>

Table A1. BPVS-FE and BPVS-RE estimation results for the 20 developed countries sample

Variable	Estimate	CI Lower	CI Upper	PIP	Rank	Selection	Variable	Estimate	CI Lower	CI Upper	PIP	Rank	Selection
Debt	-1.85884	-2.59668	-1.10342	1	1	Selected	FDI	0.962984	0.626755	1.288355	1	1	Selected
FDI	1.030005	0.694627	1.352523	1	2	Selected	POP	-1.18902	-1.91174	-0.36335	0.999	2	Selected
POP	-1.80141	-2.42489	-1.17657	1	3	Selected	TOP	1.1255	0.349677	2.122802	0.99875	3	Selected
TOP	2.28407	1.428904	3.132849	1	4	Selected	IH	1.069083	0.290914	1.984163	0.998125	4	Selected
EXR	-2.40197	-3.59736	-1.18904	1	5	Selected	UR	-0.85312	-1.4125	-0.26173	0.996625	5	Selected
IH	1.591137	0.621459	2.58707	0.99975	6	Selected	Debt	-0.83166	-1.73658	-0.06091	0.987375	6	Selected
UR	-0.91705	-1.4671	-0.36123	0.99925	7	Selected	EXR	-1.01274	-2.27851	-0.00886	0.982	7	Selected
IE	0.824113	0.120531	1.528727	0.992625	8	Selected	IE	0.56164	-0.04369	1.239799	0.9725	8	Selected
GR	-0.63632	-1.61003	0.261685	0.966625	9	Selected	GR	-0.46871	-1.22803	0.19253	0.950625	9	Selected
INF	-0.01299	-0.35718	0.316519	0.789625	10	Selected	INF	0.001383	-0.32507	0.320566	0.774875	10	Selected

Notes. All ten covariates satisfy $PIP > 0.5$; ranks are by PIP. For each specification, we report the posterior means of the (regularized) coefficient, the 95 percent credible interval, the posterior inclusion probabilities (PIP), and the PIP rank. BPVS results are computed following Pastpipatkul & Ko [8].

Table A2. BPVS-FE and BPVS-RE estimation results for the 25 developing countries sample

Variable	Estimate	CI Lower	CI Upper	PIP	Rank	Selection	Variable	Estimate	CI Lower	CI Upper	PIP	Rank	Selection
TOP	0.952116	0.052315	1.747921	0.98525	1	Selected	POP	-0.7513	-1.2765	-0.10011	0.9875	1	Selected
EXR	0.686799	-0.01547	1.42911	0.96275	2	Selected	EXR	0.577072	0.058335	1.068686	0.983125	2	Selected
FDI	0.288761	-0.00992	0.60038	0.942	3	Selected	FDI	0.303941	0.002131	0.603705	0.95875	3	Selected
Debt	-0.32385	-0.77754	0.049764	0.91975	4	Selected	Debt	-0.29686	-0.6791	0.033425	0.915	4	Selected
POP	-0.48075	-1.66992	0.182536	0.89975	5	Selected	TOP	0.321834	-0.07477	0.934782	0.89625	5	Selected
GR	-0.2232	-0.97651	0.220902	0.85525	6	Selected	GR	-0.16202	-0.74013	0.159568	0.8205	6	Selected
IH	-0.11476	-0.97919	0.366781	0.82525	7	Selected	IH	0.120373	-0.23382	0.650731	0.809	7	Selected
UR	-0.02547	-0.65493	0.50136	0.7975	8	Selected	UR	-0.02184	-0.42776	0.292168	0.72475	8	Selected
IE	0.063003	-0.26752	0.531376	0.773875	9	Selected	IE	-0.00452	-0.34276	0.311988	0.696125	9	Selected
INF	-0.01331	-0.33692	0.274909	0.708875	10	Selected	INF	0.002748	-0.25156	0.266836	0.66825	10	Selected

Notes. Same as described in A1.

Table A3. BPVS-FE and BPVS-RE estimation results for the 45 mixed (all) countries sample

Variable	Estimate	CI Lower	CI Upper	PIP	Rank	Selection	Variable	Estimate	CI Lower	CI Upper	PIP	Rank	Selection
Debt	-1.02413	-1.46021	-0.58897	1	1	Selected	FDI	0.617614	0.395668	0.838827	1	1	Selected
FDI	0.638779	0.412248	0.860453	1	2	Selected	POP	-0.88629	-1.27092	-0.51954	1	2	Selected
POP	-1.45234	-2.08123	-0.81544	1	3	Selected	Debt	-0.5867	-0.89057	-0.28718	0.999625	3	Selected
TOP	1.557391	0.953352	2.154842	1	4	Selected	TOP	0.561946	0.202832	0.954686	0.998625	4	Selected
GR	-0.82269	-1.84778	0.101287	0.97125	5	Selected	GR	-0.56795	-1.08035	-0.04021	0.98225	5	Selected
EXR	-0.64752	-1.57947	0.166345	0.96125	6	Selected	IH	0.395679	-0.05372	0.883739	0.9435	6	Selected
UR	-0.41708	-0.91451	0.044193	0.9575	7	Selected	UR	-0.14714	-0.49788	0.111126	0.828875	7	Selected
IE	0.216143	-0.14871	0.63658	0.888	8	Selected	EXR	-0.06525	-0.45979	0.229287	0.769375	8	Selected
IH	0.093175	-0.56156	0.851044	0.881375	9	Selected	IE	0.034874	-0.21301	0.348483	0.7195	9	Selected
INF	0.017194	-0.24183	0.305004	0.735375	10	Selected	INF	0.059105	-0.13961	0.317059	0.71	10	Selected

Notes. Same as described in A1 and A2.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryData.xlsx](#)
- [SupplementaryFESVandBPVSEstimationinR.pdf](#)