

An Order-Invariant Decomposition of Input Accumulation in DEA–Malmquist Productivity Analysis

Kozo Kunimune

kunimune@kwansei.ac.jp

Kwansei Gakuin University

Research Article

Keywords: Data envelopment analysis (DEA), Malmquist productivity index, Input accumulation, Order invariance, Frontier decomposition, Nonparametric productivity measurement

Posted Date: October 14th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9620367/v1>

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

[Read Full License](#)

Additional Declarations: No competing interests reported.

An Order-Invariant Decomposition of Input Accumulation in DEA–Malmquist Productivity Analysis

Kozo Kunimune

School of International Studies, Kwansei Gakuin University,
Nishinomiya, Japan.

Corresponding author(s). E-mail(s): kunimune@kwansei.ac.jp;

Abstract

This paper resolves an order-dependence problem that arises when decomposing the input-accumulation component of the DEA–Malmquist index in multi-input settings. While accumulation is commonly interpreted as movement along the productivity frontier, attributing this frontier movement to individual inputs can depend on the sequence in which input changes are evaluated. We propose a path-symmetrized decomposition that averages accumulation effects over all admissible input-change orderings. The resulting input-specific components are uniquely defined, order-invariant, and multiplicatively consistent with the aggregate accumulation term. We establish these properties formally and provide implementation guidance based on standard DEA projection mappings. As an empirical illustration, we apply the method to Penn World Table data and decompose labor productivity growth into efficiency change, technological change, and order-invariant physical- and human-capital accumulation components. The approach is compatible with common DEA orientations and returns-to-scale assumptions as long as the required projection mapping is well defined. Overall, it offers a transparent way to examine the internal structure of input accumulation in DEA–Malmquist productivity analysis with multiple inputs.

Keywords: Data envelopment analysis (DEA); Malmquist productivity index; Input accumulation; Order invariance; Frontier decomposition; Nonparametric productivity measurement

JEL Classification: C14; C43; C61; D24; O47

Acknowledgments

The author received no specific funding for this work. Any remaining errors are the author's own.

1 Introduction

DEA–Malmquist analysis provides a nonparametric growth-accounting view of productivity change as a combination of efficiency change (catch-up), technological change (frontier shifts), and input accumulation interpreted as movement along the frontier. Following the benchmark contribution of [Kumar and Russell \(2002\)](#), this decomposition has become a standard tool in empirical productivity analysis.

A practical difficulty arises when the input vector is multidimensional and one seeks to attribute the accumulation component to individual inputs (e.g., physical vs. human capital). In multi-input settings, such input-specific accumulation measures are generally order dependent: different sequences of input changes can yield different decompositions, which complicates interpretation and comparability.

Relatively little of the empirical DEA–Malmquist literature explicitly addresses this ordering/permutation issue for input-specific accumulation in genuinely multi-input settings. A prominent line of work instead incorporates human capital in a labor-augmenting form (effective labor $\hat{L} = h \cdot L$) and yields a quadripartite (four-part) decomposition separating physical- and human-capital accumulation alongside efficiency change and technological change ([Henderson and Russell, 2005](#)). This formulation has become a mainstream template and has been widely adopted and extended in subsequent empirical work (hereafter “HR-series”) ([Badunenko and Romero-Ávila, 2013](#); [Badunenko, Henderson, and Russell, 2013](#); [Färe, Grosskopf, and Margaritis, 2006](#); [Margaritis, Färe, and Grosskopf, 2007](#); [Meng, Parmeter, and Zelenyuk, 2023](#)). This paper addresses a source of non-uniqueness that is conceptually distinct from the well-known evaluation-point dependence in DEA–Malmquist decompositions. Even after adopting Fisher-ideal symmetrization for the aggregate technological-change and accumulation terms ([Badunenko, Henderson, and Houssa, 2014](#); [Margaritis et al., 2007](#)), an additional ambiguity remains when the accumulation component is further attributed to individual inputs in genuinely multi-input technologies: the resulting input-specific attribution can depend on the ordering (permutation) of sequential input changes along the frontier. We resolve this ordering/permutation dependence by a path-symmetrized (permutation-averaged) construction that yields unique, order-invariant, and multiplicatively consistent input-specific accumulation components.

Adjacent to the HR-series, the DEA literature has developed methods for input-specific productivity change that can uncover factor-level drivers hidden by aggregate measures; [Aparicio, Kapelko, Mahlberg, and Sainz-Pardo \(2017\)](#) provide a recent methodological treatment. More generally, when an aggregate outcome depends non-linearly on multiple factors, attributions based on “introducing” or “zeroing out” factors can depend on the processing order. In such cases, Shapley-type symmetrization yields an order-invariant attribution by averaging marginal contributions over all orderings ([Audoly, McGee, Ocampo, and Paz-Pardo, 2025](#)).

Motivated by these insights, this paper proposes an order-invariant, input-specific decomposition of the Malmquist accumulation component for multi-input DEA settings. We symmetrize accumulation effects over all admissible sequences of input changes, producing input-specific accumulation components that are invariant to ordering while remaining multiplicatively consistent with the aggregate accumulation effect. We establish theoretical properties (order invariance and consistency), provide implementation guidance, and illustrate the method empirically using international productivity data from the Penn World Table.

Appendix A summarizes the distinct sources of non-uniqueness in DEA–Malmquist decompositions and clarifies which one is resolved by the proposed order-invariant accumulation attribution.

The remainder of the paper is organized as follows. Section 2 reviews the DEA–Malmquist framework and clarifies the source of order dependence in multi-input accumulation measures. Section 3 introduces the path-symmetrized decomposition and its extension to multiple inputs. Section 4 presents theoretical properties. Section 5 discusses implementation details, including the DEA projection mapping and cross-period feasibility issues. Section 6 provides an empirical illustration, beginning with brief numerical checks and then discussing interpretation. Section 7 concludes.

2 Background

2.1 A Conceptual Reference Production Function

As a conceptual starting point for productivity frontier analysis, this paper considers a macro production function normalized by hours worked. Specifically, the following relationship serves as a reference framework:

$$y = Ak^\alpha h^{1-\alpha}, \quad (1)$$

where y denotes real output per hour worked, k is physical capital per hour worked, and h denotes a human-capital term per hour (often interpreted as a labor-augmenting human capital in efficiency units in development/growth accounting). This reference production function is standard in development accounting and growth accounting as a convenient benchmark to separate the roles of physical capital, human capital, and efficiency in explaining differences in output per hour worked (Caselli, 2005; Jones, 2015, 2016). In the empirical illustration, we proxy the conceptual human-capital term h using the Penn World Table human capital index hc .¹

It should be emphasized, however, that the DEA approach employed in this paper is non-parametric and does not impose any specific functional form in estimation. Accordingly, the production function above is not used as an estimating equation, but rather as a conceptual reference to motivate the choice of output and input variables. In particular, it provides a theoretical rationale for treating observed counterparts of k and h as input variables, and observed output per hour worked as the output variable.

In this respect, the paper shares the general motivation of earlier DEA-based productivity studies, while differing from them in several aspects of input and output

¹Because DEA is nonparametric and requires only an ordinal (monotonic) relationship between inputs and output, treating hc as an empirical proxy for h is sufficient for our purposes.

definitions. Early studies, such as [Färe, Grosskopf, Norris, and Zhang \(1994\)](#), typically used level variables for output and inputs, including aggregate production, capital, and labor. [Kumar and Russell \(2002\)](#) adopted labor productivity measures (output per worker) in the later stages of their analysis and effectively reduced inputs to capital per worker. [Henderson and Russell \(2005\)](#), in contrast, incorporated human capital indirectly through efficiency-unit transformations.

By contrast, this paper explicitly introduces human capital as an independent input variable alongside physical capital, and normalizes both output and capital by hours worked. This design allows the comparison of productivity frontiers constructed under a one-dimensional input specification (physical capital only) and a two-dimensional specification (physical and human capital) within a unified analytical framework. Moreover, this setup is convenient for the subsequent decomposition of input accumulation effects into contributions from physical capital and human capital.

2.2 Empirical Implementation Framework

The empirical illustration uses an output-oriented variable returns to scale (VRS) model with contemporaneous (year-by-year) frontiers. The proposed path-symmetrized decomposition itself does not depend on the choice of orientation, returns to scale, and frontier construction, as long as a consistent projection mapping $F(\cdot, t)$ is available. The following sections develop the decomposition and examine its properties and implementation.

2.3 Malmquist Index Decomposition

This subsection presents the formal Malmquist index decomposition using a DEA projection mapping, denoted by $F(\cdot, t)$, which assigns a frontier output value to any given input vector. The formal definition of F is provided in [Section 5](#).

Let y_t denote output per hour worked at time t , and let x_t denote the corresponding input vector. For notational convenience, country indices (j) are suppressed.

Productivity change between t and $t + 1$ is expressed as:

$$\frac{y_{t+1}}{y_t} = \text{EFF} \times \text{TECH} \times \text{ACCUM},$$

where EFF denotes efficiency change, TECH denotes technological change, and ACCUM denotes the input accumulation effect.

Efficiency change is defined as the change in relative efficiency with respect to contemporaneous frontiers:

$$\text{EFF} = \frac{y_{t+1}/F(x_{t+1}, t+1)}{y_t/F(x_t, t)}.$$

Technological change captures shifts in the frontier over time and is defined as the geometric mean of frontier movements evaluated at both period inputs:

$$\text{TECH} = \sqrt{\frac{F(x_t, t+1)}{F(x_t, t)} \cdot \frac{F(x_{t+1}, t+1)}{F(x_{t+1}, t)}}.$$

Input accumulation captures movements along the frontier induced by changes in inputs and is defined as:

$$\text{ACCUM} = \sqrt{\frac{F(x_{t+1}, t)}{F(x_t, t)} \cdot \frac{F(x_{t+1}, t+1)}{F(x_t, t+1)}}.$$

These measures can be computed directly once the DEA projection mapping F is available for each period. Within this framework, $\text{EFF} \times \text{TECH}$ corresponds to the Malmquist productivity index (MPI) and is commonly interpreted as total factor productivity (TFP) change. Accordingly, overall labor productivity growth decomposes into TFP change (MPI) and accumulation (ACCUM), and Section 3 further decomposes ACCUM into input-specific components using the same mapping F .

3 Method

This section develops a path-symmetrized decomposition that resolves the order-dependence problem arising in the accumulation component of the DEA–Malmquist index when the input vector is multidimensional. In such settings, the standard accumulation measure can be factored into input-specific contributions in multiple, equally admissible ways (i.e., across different permutations of componentwise input changes from x_t to x_{t+1}), depending on the sequence in which individual input changes are evaluated along the productivity frontier.

3.1 Decomposing the Accumulation Effect with Two Inputs

This subsection presents the decomposition for the two-input case. The key difficulty is order dependence: when both inputs change simultaneously, the contribution attributed to each input depends on the sequence in which input changes are evaluated. To address this issue, we introduce a path-symmetrization procedure that averages accumulation effects over the two possible input-change orders, yielding input-specific contributions that are invariant to ordering.

Setup and counterfactual input vectors

To reduce notational clutter, define four input vectors that serve as reference points in the computation. Let the k -th component of the input vector be denoted by a superscript (k) . Then the observed input vectors at times t and $t+1$ can be written as

$$x_t = (x_t^{(1)}, x_t^{(2)}), \quad x_{t+1} = (x_{t+1}^{(1)}, x_{t+1}^{(2)}).$$

Following this notation, define the four two-dimensional vectors:

$$p_{00} = (x_t^{(1)}, x_t^{(2)}), \quad p_{10} = (x_{t+1}^{(1)}, x_t^{(2)}), \quad p_{01} = (x_t^{(1)}, x_{t+1}^{(2)}), \quad p_{11} = (x_{t+1}^{(1)}, x_{t+1}^{(2)}).$$

Note that p_{00} and p_{11} are simply alternative notations for x_t and x_{t+1} . In contrast, p_{10} and p_{01} are constructed by partially combining components from x_t and x_{t+1} ; they are "counterfactual" input vectors for which no observed output value exists².

With these definitions, ACCUM can be written as

$$\text{ACCUM} = \sqrt{\frac{F(p_{11}, t)}{F(p_{00}, t)} \cdot \frac{F(p_{11}, t+1)}{F(p_{00}, t+1)}}.$$

The goal is to decompose this effect into the contribution of changes in the first input component and the contribution of changes in the second input component, while removing order dependence.

Path symmetrization

There are two natural input-change orders (paths) from p_{00} to p_{11} : one that changes the first component first (via p_{10}), and one that changes the second component first (via p_{01}). For each path, we factor the frontier-output ratio into two step ratios on each period's frontier.

(i) Path A (via p_{10} ; change input 1 first).

By multiplying and dividing by $F(p_{10}, t)$, we obtain,

$$\frac{F(p_{11}, t)}{F(p_{00}, t)} = \frac{F(p_{10}, t)}{F(p_{00}, t)} \cdot \frac{F(p_{11}, t)}{F(p_{10}, t)}, \quad \frac{F(p_{11}, t+1)}{F(p_{00}, t+1)} = \frac{F(p_{10}, t+1)}{F(p_{00}, t+1)} \cdot \frac{F(p_{11}, t+1)}{F(p_{10}, t+1)}.$$

Define

$$A_1 = \frac{F(p_{10}, t)}{F(p_{00}, t)}, \quad A_2 = \frac{F(p_{11}, t)}{F(p_{10}, t)}, \quad B_1 = \frac{F(p_{10}, t+1)}{F(p_{00}, t+1)}, \quad B_2 = \frac{F(p_{11}, t+1)}{F(p_{10}, t+1)}.$$

Then

$$\text{ACCUM} = \sqrt{A_1 A_2 B_1 B_2}.$$

(ii) Path B (via p_{01} ; change input 2 first).

Similarly,

$$\frac{F(p_{11}, t)}{F(p_{00}, t)} = \frac{F(p_{01}, t)}{F(p_{00}, t)} \cdot \frac{F(p_{11}, t)}{F(p_{01}, t)}, \quad \frac{F(p_{11}, t+1)}{F(p_{00}, t+1)} = \frac{F(p_{01}, t+1)}{F(p_{00}, t+1)} \cdot \frac{F(p_{11}, t+1)}{F(p_{01}, t+1)}.$$

Define

$$C_1 = \frac{F(p_{01}, t)}{F(p_{00}, t)}, \quad C_2 = \frac{F(p_{11}, t)}{F(p_{01}, t)}, \quad D_1 = \frac{F(p_{01}, t+1)}{F(p_{00}, t+1)}, \quad D_2 = \frac{F(p_{11}, t+1)}{F(p_{01}, t+1)}.$$

²For counterfactual input vectors, $F(\cdot, t)$ is defined when the corresponding DEA linear program is feasible. In practice, feasibility may fail near endpoints; see Section 5.2 for implementation rules.

Then

$$\text{ACCUM} = \sqrt{C_1 C_2 D_1 D_2}.$$

(iii) Symmetrized expression.

Order dependence is removed by taking the geometric mean over the two paths. Multiplying the two expressions implies

$$\text{ACCUM}^2 = \sqrt{A_1 A_2 B_1 B_2 C_1 C_2 D_1 D_2},$$

so that

$$\text{ACCUM} = (A_1 A_2 B_1 B_2 C_1 C_2 D_1 D_2)^{1/4}.$$

Input-specific contributions

Rearrange the symmetrized expression by grouping the step ratios associated with changes in the first input component versus the second input component. Define

$$G_1 = (A_1 B_1 C_2 D_2)^{1/4}, \quad G_2 = (A_2 B_2 C_1 D_1)^{1/4}.$$

Then the accumulation effect decomposes multiplicatively as

$$\text{ACCUM} = G_1 \cdot G_2.$$

Here, G_1 can be interpreted as the contribution of changes in the first input component, and G_2 as the contribution of changes in the second component, with order dependence removed by construction. In the empirical analysis, when the first input corresponds to physical capital per hour worked and the second input to human capital, we denote these input-specific accumulation components as KACCUM and HACCUM, respectively.

3.2 Generalization to Multiple Inputs

This subsection generalizes the decomposition method in the previous subsection from the two-dimensional input case to the case in which the input vector has m dimensions.

Consider a path that moves from x_t to x_{t+1} by gradually replacing components of x_t with those of x_{t+1} , one component at a time. Specifically, at each step, we replace exactly one component of x from its time- t value by its time- $t+1$ value.

Let σ denote a particular path, and let S_m denote the set of all such paths. Here S_m is the set of all permutations of $\{1, \dots, m\}$, and thus $\sigma \in S_m$.

Along any path, there are m steps. Let s denote the step index, where s takes values from 1 to m . Note that the input components themselves are indexed by k , which also runs from 1 to m .

For a given path σ , define the following ratio for each step s on the frontier at time t :

$$r_{\sigma,s}^{(t)},$$

as the ratio of the projected frontier output after the step to that before the step, evaluated on the frontier at time t , along the path σ .³

Equivalently, each r represents "the projected output ratio on a given frontier when only the input component associated with that step is changed." (In the previous subsection, these correspond to quantities such as A_1 .)

Along any given path σ , multiplying all step ratios yields the frontier-output ratio on a given frontier; i.e.,

$$\frac{F(x_{t+1}, t)}{F(x_t, t)} = \prod_{s=1}^m r_{\sigma,s}^{(t)}, \quad \frac{F(x_{t+1}, t+1)}{F(x_t, t+1)} = \prod_{s=1}^m r_{\sigma,s}^{(t+1)}.$$

Therefore,

$$\text{ACCUM} = \left(\prod_{s=1}^m r_{\sigma,s}^{(t)} r_{\sigma,s}^{(t+1)} \right)^{1/2}.$$

This equality holds for every path σ , and there are $m!$ such paths. Multiplying the above expression across all $\sigma \in S_m$ yields

$$\text{ACCUM}^{m!} = \left(\prod_{\sigma \in S_m} \prod_{s=1}^m r_{\sigma,s}^{(t)} r_{\sigma,s}^{(t+1)} \right)^{1/2}.$$

Taking the $m!$ -th root of both sides yields

$$\text{ACCUM} = \left(\prod_{\sigma \in S_m} \prod_{s=1}^m r_{\sigma,s}^{(t)} r_{\sigma,s}^{(t+1)} \right)^{1/(2m!)} \quad (2)$$

Now define the position mapping $\text{pos}_\sigma(k)$ as "the position (step number) at which component k appears along path σ "; that is,

$$\text{pos}_\sigma(k) = s \iff \sigma(s) = k.$$

Using this mapping, define the symmetrized contribution of component k as the geometric mean of all ratios corresponding to the step in which component k changes, averaged over all paths and both periods:

$$G_k = \left(\prod_{\sigma \in S_m} r_{\sigma, \text{pos}_\sigma(k)}^{(t)} r_{\sigma, \text{pos}_\sigma(k)}^{(t+1)} \right)^{1/(2m!)}.$$

This G_k represents the contribution of component k within ACCUM, because it collects and symmetrically averages the projected output ratios associated with changes in component k .⁴

Taking the product over all components k gives

³Formally, if $p_{\sigma,s}$ denotes the intermediate input vector after step s along path σ (with $p_{\sigma,0} = x_t$ and $p_{\sigma,m} = x_{t+1}$), then $r_{\sigma,s}^{(t)} = F(p_{\sigma,s}, t) / F(p_{\sigma,s-1}, t)$.

⁴Since each component appears exactly once along each permutation σ , collecting across all σ yields $2m!$ ratios for component k when both periods t and $t+1$ are included. This is why the geometric mean takes the $2m!$ -th root.

$$\prod_{k=1}^m G_k.$$

This product equals the right-hand side of equation (2), because equation (2) multiplies all r terms across all σ and all s , while $\prod_{k=1}^m G_k$ multiplies exactly the same set of terms, merely grouped by component k rather than by step s .⁵ Since the left-hand side of equation (2) is ACCUM, we obtain

$$\text{ACCUM} = \prod_{k=1}^m G_k. \quad (3)$$

While full permutation averaging is straightforward for small m (including our two-input illustration), enumerating all $m!$ paths can be computationally burdensome for large m . Appendix B provides a simple Monte Carlo approximation and a stability diagnostic for such cases.

4 Theoretical Properties

This section summarizes the key theoretical properties of the proposed path-symmetrized decomposition of the Malmquist accumulation component. These properties clarify why the decomposition provides a coherent and interpretable measure of input-specific accumulation in multi-input DEA settings.

4.1 Order Invariance

A central property of the proposed decomposition is order invariance. When multiple inputs change simultaneously, sequential (order-specific) attributions of the accumulation effect to individual inputs can become path dependent: the measured input-specific contributions may vary with the sequence in which input adjustments are evaluated along the frontier. Different sequences correspond to different but equally admissible paths between the same observed input vectors, and thus can yield different input-specific accumulation measures even when the aggregate change in inputs is the same.

By contrast, the proposed approach eliminates this dependence by symmetrizing accumulation effects over all admissible input-change paths. In the two-input case, this is achieved by taking the geometric mean of accumulation effects evaluated along the two possible paths. In the general multi-input case, symmetrization is implemented by averaging over all permutations of input changes. As a result, the contribution attributed to each input is invariant to the ordering of input adjustments.

Order invariance ensures that the decomposition is uniquely defined given the observed data and frontier specification. This property restores comparability across empirical applications and prevents arbitrary modeling choices from affecting input-specific accumulation measures.

⁵The right-hand side of (2) multiplies all r across all σ and all s . The product $\prod_{k=1}^m G_k$ multiplies the same terms, with multiplication reordered by component. By commutativity, the two expressions coincide.

4.2 Multiplicative Consistency

The proposed decomposition also satisfies multiplicative consistency. Specifically, the product of input-specific accumulation components equals the aggregate accumulation effect obtained from the standard Malmquist decomposition.

In the two-input case, the accumulation effect can be written as the product of the physical-capital and human-capital accumulation components. In the general multi-input case, the accumulation effect equals the product of the symmetrized contributions of all input components. This property follows directly from the construction of the decomposition, which redistributes the same set of frontier-output ratios across inputs rather than introducing additional factors.

Multiplicative consistency guarantees internal coherence in the sense that the product of the input-specific accumulation components exactly reproduces the aggregate accumulation term. Thus, the proposed decomposition preserves the standard Malmquist structure while making the contribution of each input explicit.

4.3 Reduction to the Standard Case

When the input vector is one-dimensional, order dependence is absent by construction, because there is no ambiguity in the sequence of input changes. In this case, the proposed construction coincides with the standard Malmquist accumulation measure.

Similarly, when multiple inputs are present but only one input changes between periods, the accumulation effect attributed to that input coincides with the aggregate accumulation component, while the contributions of all other inputs equal unity.⁶ In such cases, different input-change orders yield identical accumulation measures, and symmetrization has no effect on the resulting decomposition.

These reduction results serve as a methodological consistency check, ensuring that the proposed decomposition does not alter accumulation measures in settings where order dependence does not arise.

4.4 Interpretation as Symmetric Marginal Contributions

The permutation-average (geometric-mean) construction is motivated by three desirable properties that are particularly salient in non-linear multi-input decompositions: (i) *symmetry* across inputs (no input should be privileged by an arbitrary ordering), (ii) *path invariance* (input-specific attribution should not depend on the sequence in which inputs are adjusted), and (iii) *multiplicative consistency* (the product of input-specific components should recover the aggregate accumulation effect). Averaging marginal effects over all orderings provides a direct symmetrization device that meets these desirable properties simultaneously, closely mirroring the logic of Shapley-type attribution for order-sensitive decompositions.

The proposed decomposition can be interpreted as assigning symmetric marginal contributions of inputs to accumulation along the productivity frontier. By averaging over all admissible input-change orders, the contribution of each input reflects its

⁶In the empirical illustration, several countries exhibit this collapse directly in the path-specific diagnostics: one component is exactly unity and both paths deliver identical attributions (Appendix E, Table 3; e.g., Canada, Switzerland, United States; Cyprus and Iceland).

average effect on frontier movement when evaluated in all possible positions along an input-change sequence.

This interpretation is formally analogous to the Shapley value concept in cooperative game theory, in which contributions are defined as averages of marginal effects across all orderings (Audoly et al., 2025). However, the present decomposition is distinctive in that it is derived within a DEA-based frontier framework.

Appendix C (Lemma 1) formalizes this Shapley-type interpretation by showing that $\log G_k$ equals the arithmetic average of marginal log-contributions of input k over all admissible orderings and both periods.

4.5 Scope and Applicability

The theoretical properties discussed above do not depend on specific functional-form assumptions or parametric restrictions. They apply to output-oriented DEA models with variable returns to scale and extend naturally to alternative DEA specifications, provided that the relevant projection mappings are well defined. Practical feasibility of the required mappings in cross-period evaluations is discussed in Section 5.2.

These properties establish the proposed path-symmetrized decomposition as a general and robust tool for analyzing the internal structure of input accumulation in multi-input productivity analysis.

5 Implementation Details

This section describes implementation issues relevant for applying the proposed path-symmetrized decomposition in empirical settings. The focus is on the construction of DEA projection mappings and on numerical feasibility issues that arise when evaluating counterfactual input combinations.

5.1 DEA Projection Mapping

Year-by-year DEA frontier analysis requires a systematic way to evaluate the position of each observation relative to the corresponding period’s efficiency frontier. For this purpose, the analysis relies on a DEA projection mapping.

Given an input vector x and evaluation time t , the DEA projection mapping $F(x, t)$ is defined as the solution to the following linear programming problem:

$$F(x, t) = \max_{\{\lambda_{jt}\}} \left\{ \sum_{j=1}^N \lambda_{jt} y_{jt} \mid \sum_{j=1}^N \lambda_{jt} x_{jt} \leq x, \sum_{j=1}^N \lambda_{jt} = 1, \lambda_{jt} \geq 0 \right\}. \quad (4)$$

Equation (4) presents one convenient specification (output-oriented VRS under year-by-year frontiers). More generally, the proposed decomposition applies unchanged as long as the frontier projection mapping $F(\cdot, t)$ is well defined: alternative DEA orientations, returns-to-scale assumptions, or technology constructions can be accommodated by redefining F accordingly.

This mapping represents the maximum attainable output under the DEA technology set constructed from observations in year t . It is not an analytically specified

production function, but a correspondence defined by the optimal value of a linear program. Its role is to assign a frontier output value to any given input vector.

An important feature of the DEA projection mapping is that it can be applied not only to observed input vectors, but also to counterfactual input vectors that do not directly correspond to observed data points. This property allows the accumulation decomposition to evaluate movements along the frontier induced by changes in individual input components.

5.2 Infeasibility under Cross-Period Projections and Endpoint Treatment

Infeasibilities may arise in mixed-period (cross-period) evaluations, that is, when an input vector is evaluated on the other period’s technology, such as $F(x_{t+1}, t)$ or $F(x_t, t + 1)$. These cross-period projections are required not only for the accumulation term (and its input-specific decomposition) but also for the technological-change term TECH, which is defined using frontier values evaluated at both period inputs. Accordingly, cases in which the accumulation decomposition cannot be computed often overlap with cases in which TECH cannot be computed symmetrically (and would otherwise require a one-sided or ad hoc workaround).

In the output-oriented VRS setting, a cross-period projection $F(x, t)$ is feasible only if the evaluation input vector x is weakly spanned by the period- t technology (i.e., some convex combination of observed activities uses inputs no greater than x). When this condition fails—most commonly because x is too small in at least one dimension—the linear program (LP) is infeasible and $F(x, t)$ is undefined.

Myanmar illustrates cross-period infeasibility under contemporaneous technologies: in the two-input setting, $F(x_t, t + 1)$ is undefined. Section 6.1 reports how feasibility and sample selection are handled in the 2000–2010 analysis, and Appendix D details the diagnostic protocol.

6 Empirical Illustration and Numerical Checks

This section provides an empirical illustration of the proposed path-symmetrized decomposition, together with brief numerical checks, using international productivity data; the exercise is descriptive and not intended to support causal claims.

6.1 Data and Empirical Setup

We use Penn World Table version 11.0 (PWT 11.0).⁷

Output per hour is defined as

$$y = \frac{\text{rgdpo}}{\text{emp} \cdot \text{avh}}.$$

⁷Penn World Table version 11.0 (PWT 11.0) is available via Dataverse at DOI: 10.34894/FABVLR.

Physical capital per hour is

$$k = \frac{K}{\text{emp} \cdot \text{avh}}.$$

Since PWT does not provide a capital stock series directly at chained PPPs, we construct a consistent series by scaling chained-PPP output with the domestic-price capital-output ratio:

$$K = \left(\frac{\text{rnna}}{\text{rgdpna}} \right) \cdot \text{rgdp},$$

where rnna and rgdpna are measured at constant 2021 national prices. Human capital is the PWT index hc. Using these definitions, we analyze (i) a one-input specification with k only and (ii) a two-input specification with k and hc , which allows a direct comparison of frontier shapes and accumulation decompositions across specifications.

Sample selection and feasibility

Our starting point is the set of countries in PWT 11.0 with complete data for the variables used in both 2000 and 2010, which yields 68 countries. We then apply a resource-dependence screen to avoid undue influence of hydrocarbon rents on measured productivity components. In the 2000–2010 window, this screen removes four otherwise-eligible countries—Myanmar, Ecuador, Trinidad and Tobago, and Venezuela (Bolivarian Republic of)—so that Tables 1–2 report results for the remaining 64 countries.

Regarding cross-period feasibility, all 64 countries reported in Tables 1–2 are fully decomposable under contemporaneous (year-by-year) technologies: none of the projection values required for EFF, TECH, ACCUM, or the two-input accumulation decomposition is infeasible. When the screen is not applied and all 68 eligible countries are included, only Myanmar becomes infeasible in the two-input setting because the mixed-period projection $F(x_t, t + 1)$ cannot be computed; the other 67 countries remain feasible.⁸

Table 1: Decomposition of labor productivity growth (one-input case; 2000–2010)

Country	TOTAL	TFP	EFF	TECH	ACCUM
Argentina	1.183	1.146	1.140	1.005	1.033
Australia	1.263	1.089	1.083	1.005	1.160
Austria	1.196	1.088	1.009	1.078	1.100
Belgium	1.227	1.093	0.920	1.188	1.123
Bangladesh	1.748	1.025	0.416	2.465	1.706
Bulgaria	1.441	0.905	0.902	1.003	1.592
Brazil	1.470	1.245	1.238	1.005	1.181
Canada	1.110	0.917	0.953	0.963	1.210
Switzerland	1.397	1.131	1.075	1.052	1.234
Chile	1.661	1.155	1.149	1.005	1.439
China	2.298	1.299	1.015	1.279	1.770
Colombia	1.270	1.188	1.185	1.003	1.068
Costa Rica	1.252	1.066	1.063	1.003	1.175
Cyprus	1.136	1.069	0.973	1.098	1.062

⁸Replication code to reproduce the tabulated outputs underlying Tables 1–2 is available at the author’s GitHub repository. The underlying PWT 11.0 data are publicly available.

Country	TOTAL	TFP	EFF	TECH	ACCUM
Czechia	1.334	1.224	1.273	0.961	1.090
Germany	1.222	1.016	1.042	0.975	1.203
Denmark	1.290	1.087	0.981	1.108	1.187
Dominican Republic	1.440	1.144	1.138	1.006	1.258
Spain	1.222	0.995	0.961	1.036	1.228
Estonia	1.788	1.121	1.138	0.986	1.595
Finland	1.168	1.070	0.999	1.070	1.092
France	1.150	1.056	0.895	1.180	1.089
United Kingdom	1.003	0.969	0.988	0.981	1.035
Greece	1.168	1.048	1.070	0.979	1.115
China, Hong Kong SAR	1.236	1.180	1.191	0.991	1.048
Croatia	1.321	1.006	1.025	0.982	1.314
Hungary	1.511	1.074	1.110	0.968	1.407
Indonesia	2.030	1.480	1.475	1.003	1.372
India	2.129	1.167	0.740	1.576	1.825
Ireland	1.672	0.979	0.914	1.070	1.709
Iceland	1.300	1.195	0.881	1.356	1.088
Israel	0.941	0.967	0.988	0.979	0.973
Italy	1.067	0.964	0.868	1.112	1.106
Jamaica	1.161	1.020	1.015	1.005	1.138
Japan	1.141	0.987	1.023	0.965	1.156
Cambodia	1.864	0.905	0.199	4.559	2.059
Republic of Korea	1.583	1.074	1.092	0.983	1.474
Sri Lanka	1.839	1.286	1.141	1.127	1.430
Lithuania	1.559	1.251	1.275	0.981	1.246
Luxembourg	0.940	0.931	0.843	1.105	1.010
Latvia	1.654	1.374	1.353	1.016	1.204
Mexico	1.113	0.955	0.957	0.998	1.166
Malta	1.072	0.884	0.895	0.987	1.213
Malaysia	1.437	1.175	1.170	1.005	1.222
Netherlands	1.182	1.074	1.016	1.057	1.101
Norway	1.476	1.189	1.000	1.189	1.242
New Zealand	1.158	0.998	1.012	0.986	1.161
Pakistan	1.165	0.969	0.577	1.680	1.202
Peru	1.665	1.401	1.398	1.002	1.189
Philippines	1.244	1.178	1.178	1.001	1.055
Poland	1.420	1.116	1.109	1.006	1.273
Portugal	1.278	0.931	0.944	0.987	1.372
Romania	2.474	1.494	1.488	1.004	1.656
Russian Federation	2.278	1.712	1.755	0.975	1.331
Singapore	1.571	1.197	1.229	0.974	1.312
Slovakia	1.534	1.328	1.370	0.969	1.156
Slovenia	1.212	1.074	1.060	1.013	1.129
Sweden	1.130	1.102	0.964	1.143	1.026
Thailand	1.771	1.485	1.481	1.003	1.193
Türkiye	1.473	1.023	1.024	0.999	1.440
Taiwan	1.141	1.132	1.177	0.962	1.008
Uruguay	1.362	1.172	1.175	0.997	1.162
United States	1.233	0.995	1.032	0.964	1.240
Viet Nam	2.393	1.116	0.457	2.443	2.144
Mean	1.425	1.120	1.050	1.149	1.270

Notes: TOTAL denotes the change in output per hour worked, which is decomposed into TFP ($\text{EFF} \times \text{TECH}$) and ACCUM. The country sample follows Section 6.1 (resource-dependent economies are excluded; see Section 5.2 for feasibility considerations).

Table 2: Decomposition of labor productivity growth (two-input case; 2000–2010)

Country	TOTAL	TFP	EFF	TECH	ACCUM	KACCUM	HACCUM
Argentina	1.183	1.118	1.148	0.973	1.058	1.034	1.024
Australia	1.263	1.089	1.084	1.005	1.160	1.160	1.000
Austria	1.196	1.057	0.961	1.100	1.132	1.104	1.025
Belgium	1.227	1.129	0.971	1.163	1.087	1.056	1.029
Bangladesh	1.748	0.928	0.594	1.562	1.885	1.616	1.166
Bulgaria	1.441	0.879	0.916	0.960	1.638	1.627	1.007
Brazil	1.470	1.102	1.146	0.961	1.334	1.200	1.111
Canada	1.110	0.917	0.953	0.963	1.210	1.210	1.000
Switzerland	1.397	1.131	1.075	1.052	1.234	1.234	1.000
Chile	1.661	1.137	1.154	0.985	1.462	1.430	1.022
China	2.298	1.385	1.139	1.216	1.660	1.597	1.039
Colombia	1.270	1.086	1.097	0.990	1.169	1.082	1.081
Costa Rica	1.252	1.027	1.026	1.000	1.219	1.189	1.026
Cyprus	1.136	0.998	0.847	1.179	1.138	1.000	1.138
Czechia	1.334	1.224	1.273	0.961	1.090	1.090	1.000
Germany	1.222	1.016	1.042	0.975	1.203	1.203	1.000
Denmark	1.290	1.085	0.982	1.104	1.190	1.183	1.005
Dominican Republic	1.440	1.061	1.105	0.960	1.357	1.277	1.062
Spain	1.222	0.985	0.880	1.119	1.242	1.145	1.084
Estonia	1.788	1.121	1.138	0.986	1.595	1.594	1.000
Finland	1.168	1.037	0.953	1.088	1.127	1.096	1.028
France	1.150	1.072	0.914	1.172	1.073	1.030	1.042
United Kingdom	1.003	0.969	0.988	0.981	1.035	1.035	1.000
Greece	1.168	1.007	0.933	1.080	1.160	1.104	1.051
China, Hong Kong SAR	1.236	1.058	0.979	1.081	1.168	1.051	1.112
Croatia	1.321	0.982	1.015	0.967	1.346	1.313	1.025
Hungary	1.511	1.054	1.101	0.958	1.434	1.384	1.036
Indonesia	2.030	1.310	1.309	1.001	1.550	1.439	1.077
India	2.129	1.242	0.991	1.253	1.714	1.531	1.119
Ireland	1.672	0.986	0.919	1.074	1.696	1.631	1.040
Iceland	1.300	1.183	1.002	1.180	1.099	1.000	1.099
Israel	0.941	0.967	0.988	0.979	0.973	0.973	1.000
Italy	1.067	0.932	0.804	1.160	1.144	1.056	1.083
Jamaica	1.161	1.023	1.040	0.983	1.135	1.139	0.996
Japan	1.141	0.986	1.023	0.964	1.157	1.156	1.001
Cambodia	1.864	0.795	0.344	2.314	2.345	1.520	1.543
Republic of Korea	1.583	1.070	1.092	0.980	1.479	1.473	1.004
Sri Lanka	1.839	1.283	1.203	1.066	1.434	1.432	1.001
Lithuania	1.559	1.225	1.270	0.964	1.273	1.246	1.022
Luxembourg	0.940	0.895	0.791	1.132	1.051	1.005	1.046
Latvia	1.654	1.325	1.244	1.065	1.248	1.197	1.043

Country	TOTAL	TFP	EFF	TECH	ACCUM	KACCUM	HACCUM
Mexico	1.113	0.922	0.957	0.963	1.207	1.175	1.028
Malta	1.072	0.864	0.894	0.966	1.241	1.214	1.022
Malaysia	1.437	1.126	1.152	0.977	1.276	1.220	1.046
Netherlands	1.182	1.045	0.966	1.082	1.131	1.110	1.019
Norway	1.476	1.186	1.000	1.186	1.245	1.213	1.026
New Zealand	1.158	1.002	1.020	0.983	1.156	1.160	0.997
Pakistan	1.165	0.851	0.757	1.124	1.369	1.100	1.244
Peru	1.665	1.364	1.349	1.011	1.221	1.202	1.016
Philippines	1.244	1.116	1.077	1.036	1.114	1.063	1.049
Poland	1.420	1.104	1.109	0.995	1.287	1.278	1.007
Portugal	1.278	1.106	0.892	1.240	1.155	1.055	1.095
Romania	2.474	1.464	1.511	0.969	1.690	1.677	1.008
Russian Federation	2.278	1.705	1.760	0.968	1.336	1.328	1.006
Singapore	1.571	1.161	1.211	0.959	1.353	1.295	1.045
Slovakia	1.534	1.324	1.370	0.967	1.159	1.155	1.003
Slovenia	1.212	1.070	1.060	1.009	1.133	1.125	1.007
Sweden	1.130	1.083	0.946	1.145	1.043	1.025	1.018
Thailand	1.771	1.314	1.304	1.008	1.348	1.222	1.103
Türkiye	1.473	0.961	1.000	0.961	1.533	1.407	1.090
Taiwan	1.141	1.047	1.038	1.009	1.090	1.006	1.083
Uruguay	1.362	1.158	1.200	0.966	1.175	1.169	1.005
United States	1.233	0.995	1.032	0.964	1.240	1.240	1.000
Viet Nam	2.393	0.989	0.498	1.988	2.419	2.046	1.182
Mean	1.425	1.091	1.040	1.080	1.307	1.239	1.051

Notes: TOTAL denotes the change in output per hour worked. ACCUM is further decomposed into KACCUM (physical-capital accumulation) and HACCUM (human-capital accumulation), such that $ACCUM = KACCUM \times HACCUM$. The country sample follows Section 6.1 (resource-dependent economies are excluded; see Section 5.2 for feasibility considerations).

6.2 Numerical Checks

Before turning to interpretation, we use the one-input specification (physical capital per hour only) as a numerical benchmark. With a single input, the accumulation term is uniquely defined, so this benchmark serves to verify that the multiplicative identities reproduce total productivity change without numerical instability (Table 1).

We then introduce human capital as a second input and compute the proposed path-symmetrized decomposition. In this two-input setting, the constructed input-specific accumulation factors are order invariant by design and satisfy the multiplicative consistency condition that their product reproduces the aggregate accumulation term (Table 2). The remainder of this section focuses on how the resulting components can be interpreted in an applied productivity analysis.

6.3 Comparison and Interpretation of Decomposition Results

Table 1 reports the Malmquist decomposition results for the one-input specification, in which physical capital per hour worked is the only input. This specification serves as a benchmark case in which accumulation is uniquely defined, so the ordering issue is moot. Table 2 reports the corresponding results for the two-input specification, in which both physical capital and human capital are included as inputs.

Comparing Tables 1 and 2 highlights how the interpretation of productivity components can change when a key input is modeled explicitly. On average, moving from the one-input to the two-input specification reduces measured technological change (Mean TECH: $1.149 \rightarrow 1.080$) while increasing measured accumulation (Mean ACCUM: $1.270 \rightarrow 1.307$). This pattern is consistent with the idea that part of what appears as "frontier shift" under a restricted input space can be captured as movement along an expanded frontier once human capital is included as an additional input.

Country patterns illustrate the same point. For Cambodia, for example, the one-input decomposition assigns a very large component to TECH (4.559), whereas under the two-input specification TECH is smaller (2.314) and the movement-along-the-frontier component (ACCUM) is larger ($2.059 \rightarrow 2.345$). Similar shifts from TECH toward ACCUM are observed for Bangladesh (TECH: $2.465 \rightarrow 1.562$; ACCUM: $1.706 \rightarrow 1.885$) and Viet Nam (TECH: $2.443 \rightarrow 1.988$; ACCUM: $2.144 \rightarrow 2.419$). These comparisons underscore that the distinction between "frontier shift" and "movement along the frontier" may be sensitive to whether relevant inputs are modeled explicitly.

The two-input results also reveal heterogeneity in the internal structure of accumulation. In many advanced economies, HACCUM is close to unity (e.g., Australia, Canada, and the United States), suggesting that most of the measured accumulation operates through physical-capital deepening. In contrast, several emerging economies exhibit sizeable human-capital accumulation components (e.g., Pakistan and Cambodia), indicating that the expansion of effective human capital can account for a nontrivial part of frontier movement.

For completeness, Appendix E (Table 3) reports the corresponding path-dependent decompositions via p_{10} and p_{01} and quantifies the resulting order dependence in input-level attributions.

6.4 Practical Remarks

The empirical illustration confirms that the proposed path-symmetrized decomposition can be implemented using standard DEA projection mappings in a straightforward manner. For the reported country sample underlying Tables 1–2, the required counterfactual input combinations are feasible, and the accumulation component can be decomposed into order-invariant physical-capital and human-capital contributions without additional computational complications.

More generally, as discussed in Section 5.2, mixed-period (cross-period) infeasibilities may arise when a counterfactual endpoint lies outside a contemporaneous technology set, so that $F(\cdot, t)$ is not well-defined for at least one period. For transparency and replicability, Appendix D provides a recommended reporting protocol

for diagnosing infeasibility and documenting how it is handled. In the empirical illustration reported here, we follow the conservative option in Appendix D (Step 3) and apply an explicit sample restriction so that all projection values needed for EFF, TECH, ACCUM, and the input-specific accumulation components are well-defined.

7 Conclusion

This paper proposes an order-invariant decomposition of the accumulation component in DEA–Malmquist productivity analysis with multiple inputs. By symmetrizing accumulation effects over all admissible input-change paths, the proposed method yields input-specific accumulation measures that are independent of ordering and multiplicatively consistent with aggregate accumulation.

The framework clarifies the internal structure of "movement along the frontier" and allows researchers to distinguish the contributions of different inputs—such as physical and human capital—within a nonparametric productivity analysis framework. Brief numerical checks confirm that the proposed decomposition eliminates the instability that arises under order-specific (path-dependent) attributions while remaining straightforward to implement.

Although the empirical illustration focuses on international productivity data, the method is general and can be applied to firm-, industry-, or regional-level analyses with multiple inputs. Future research may combine the proposed decomposition with statistical inference procedures or extend it to alternative frontier specifications.

Statements and Declarations

Author contributions

Kozo Kunimune: Conceptualization, Methodology, Data curation, Software, Formal analysis, Visualization, Writing – original draft, Writing–review and editing.

Competing interests

The author declares that they have no competing interests.

Data availability

The data used in this study are from the Penn World Table (PWT) version 11.0, publicly available via Dataverse (DOI: 10.34894/FABVLR).

Code availability

Replication code to reproduce the tabulated outputs underlying Tables 1–2 is publicly available at: <https://github.com/Kandkide/An-Order-Invariant-Decomposition-of-Input-Accumulation-in-DEA-Malmquist-Productivity-Analysis>

Materials availability

Not applicable.

Funding

The author received no specific funding for this work.

References

- Aparicio J, Kapelko M, Mahlberg B, et al (2017) Measuring input-specific productivity change based on the principle of least action. *Journal of Productivity Analysis* 47:17–31. 10.1007/s11123-016-0488-9
- Audoly R, McGee R, Ocampo S, et al (2025) A practitioner’s note on the shapley-owen-shorrocks decomposition. Staff Report 1163, Federal Reserve Bank of New York, 10.59576/sr.1163
- Badunenko O, Romero-Ávila D (2013) Financial development and the sources of growth and convergence. *International Economic Review* 54(2):629–663. 10.1111/iere.12009
- Badunenko O, Henderson DJ, Russell RR (2013) Polarization of the worldwide distribution of productivity. *Journal of Productivity Analysis* 40(2):153–171. 10.1007/s11123-012-0328-5
- Badunenko O, Henderson DJ, Houssa R (2014) Significant drivers of growth in Africa. *Journal of Productivity Analysis* 42(3):339–354. 10.1007/s11123-014-0400-4
- Caselli F (2005) Accounting for cross-country income differences. In: Aghion P, Durlauf SN (eds) *Handbook of Economic Growth*, vol 1. Elsevier, p 679–741, 10.1016/S1574-0684(05)01009-9
- Färe R, Grosskopf S, Norris M, et al (1994) Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries. *The American Economic Review* 84(1):66–83
- Färe R, Grosskopf S, Margaritis D (2006) Productivity Growth and Convergence in the European Union. *Journal of Productivity Analysis* 25:111–141. 10.1007/s11123-006-7134-x
- Henderson DJ, Russell RR (2005) Human capital and convergence: A production-frontier approach. *International Economic Review* 46(4):1167–1205. 10.1111/j.1468-2354.2005.00364.x
- Jones CI (2015) The facts of economic growth. NBER Working Paper 21142, National Bureau of Economic Research, 10.3386/w21142
- Jones CI (2016) The facts of economic growth. In: Taylor JB, Uhlig H (eds) *Handbook of Macroeconomics*, vol 2. Elsevier, p 3–69, 10.1016/bs.hesmac.2016.03.002
- Kumar S, Russell RR (2002) Technological change, technological catch-up, and capital deepening: Relative contributions to growth and convergence. *American Economic Review* 92(3):527–548. 10.1257/00028280260136381
- Margaritis D, Färe R, Grosskopf S (2007) Productivity, convergence and policy: A study of OECD countries and industries. *Journal of Productivity Analysis* 28:87–105. 10.1007/s11123-007-0044-8
- Meng Y, Parmeter CF, Zelenyuk V (2023) Is newer always better? A reinvestigation of productivity dynamics using updated PWT data. *Journal of Productivity Analysis* 59(1):1–13. 10.1007/s11123-022-00649-w

A What non-uniqueness is resolved in this paper (and what is not)

Several sources of non-uniqueness (or "path dependence") arise in DEA–Malmquist decompositions, and they are often conflated.

(1) Evaluation-point dependence (base vs. current period) in cross-period distances.

A well-known issue is that cross-period distances may be evaluated at different reference input vectors (e.g., base-period vs. current-period inputs), which can yield different measures of frontier shift and frontier movement. A common remedy is a Fisher-ideal (geometric-average) construction that symmetrizes across evaluation points (Badunenko et al., 2014; Margaritis et al., 2007).

(2) Ordering/permutation dependence in input-specific attribution within ACCUM (this paper).

Even after adopting a Fisher-ideal definition of aggregate ACCUM, an additional non-uniqueness remains when one attempts to split ACCUM into input-specific components in a genuinely multi-input setting. Specifically, the resulting input-specific attribution can depend on the sequence in which input changes are evaluated along the frontier (e.g., changing input 1 first vs. input 2 first).

In the HR-series (e.g., Henderson and Russell (2005)), human capital is incorporated as labor-augmenting (effective labor $\hat{L} = h \cdot L$), so the ordering/permutation dependence that arises when k and h are treated as separate frontier inputs is largely sidestepped rather than symmetrized. Moreover, under constant returns to scale (CRS) the two-input technology in $(K, \hat{L})$ can be normalized to a one-dimensional representation in terms of capital intensity $K/\hat{L}$, so the specific ordering issue we address—allocating ACCUM across separate frontier inputs k and h —does not arise in the same way in the HR formulation.

Resolution (this paper). We resolve source (2) by a path-symmetrized (permutation-averaged) attribution of ACCUM into input-specific components that are order-invariant and multiplicatively consistent. See Sections 3–4 for the formal construction and properties.

B Monte Carlo approximation for large m (optional implementation note)

The multi-input definition averages over all $m!$ permutations of input changes to achieve order invariance. While this is straightforward for $m = 2$ (our empirical illustration) and still feasible for small m , enumerating all $m!$ permutations becomes computationally burdensome as m grows. In such cases, a natural approximation is obtained by replacing the full permutation set S_m in the definition of G_k by a randomly sampled subset. Let $S_N \subset S_m$ denote a set of N permutations sampled uniformly at random (with or without replacement). Define the Monte Carlo analog of the component- k factor by

$$\widehat{G}_k(\mathcal{S}_N) = \prod_{\sigma \in \mathcal{S}_N} \left(r_{\sigma, \text{pos}_\sigma(k)}^{(t)} r_{\sigma, \text{pos}_\sigma(k)}^{(t+1)} \right)^{1/(2N)},$$

which is obtained from the definition of G_k by restricting the product to $\sigma \in \mathcal{S}_N$ and replacing $2m!$ by $2N$ in the geometric-mean exponent. Similarly, an approximation to ACCUM can be formed by replacing S_m with $\mathcal{S}_N$ in equation (2).

In practice, N can be increased until the estimates stabilize. A simple diagnostic is to use nested samples (e.g., $\mathcal{S}_{N/2} \subset \mathcal{S}_N$) and check that the log-components change little:

$$\max_k |\log \widehat{G}_k(\mathcal{S}_N) - \log \widehat{G}_k(\mathcal{S}_{N/2})| < \varepsilon,$$

for a small tolerance ε (e.g., 10^{-3}). Reporting N , ε , and the random seed (or the sampled set $\mathcal{S}_N$) facilitates exact replication. As N increases, this sampling-based procedure provides an increasingly accurate approximation to the full permutation average under standard Monte Carlo logic. Since the purpose of this paper is to illustrate the proposed decomposition in a two-input setting, we leave a systematic investigation of large- m approximation performance to future work.

C Lemma 1: Log-additive interpretation and Shapley-type symmetrization

Consider the multi-input construction in Section 3.2. For any permutation (path) $\sigma \in S_m$ and step s , let $r_{\sigma,s}^{(t)}$ denote the projected frontier-output ratio on the period- t frontier associated with step s along path σ , and define $r_{\sigma,s}^{(t+1)}$ analogously. Let G_k be the proposed order-invariant input-specific accumulation component for input k :

$$G_k = \prod_{\sigma \in S_m} \left(r_{\sigma, \text{pos}_\sigma(k)}^{(t)} r_{\sigma, \text{pos}_\sigma(k)}^{(t+1)} \right)^{1/(2m!)}.$$

Then $\log G_k$ equals the simple arithmetic average of the marginal log-contributions of input k over all admissible orderings and both periods:

$$\log G_k = \frac{1}{2m!} \sum_{\sigma \in S_m} \left(\log r_{\sigma, \text{pos}_\sigma(k)}^{(t)} + \log r_{\sigma, \text{pos}_\sigma(k)}^{(t+1)} \right).$$

In particular, the proposed geometric-mean (multiplicative) attribution is equivalent to a Shapley-type symmetrization in log space: it averages the marginal impact of input k across all positions it can take in the input-change sequence.

Proof

Proof. By definition,

$$G_k = \prod_{\sigma \in S_m} \left(r_{\sigma, \text{pos}_\sigma(k)}^{(t)} r_{\sigma, \text{pos}_\sigma(k)}^{(t+1)} \right)^{1/(2m!)}.$$

Taking logs and using $\log \prod_i a_i = \sum_i \log a_i$ and $\log(a^c) = c \log a$, we obtain

$$\log G_k = \sum_{\sigma \in S_m} \frac{1}{2m!} \log \left(r_{\sigma, \text{pos}_\sigma(k)}^{(t)} r_{\sigma, \text{pos}_\sigma(k)}^{(t+1)} \right) = \frac{1}{2m!} \sum_{\sigma \in S_m} \left(\log r_{\sigma, \text{pos}_\sigma(k)}^{(t)} + \log r_{\sigma, \text{pos}_\sigma(k)}^{(t+1)} \right).$$

This is exactly the arithmetic average of the marginal log-contributions of input k across all admissible orderings and both periods. $\square$

Remark

Lemma 1 shows that the multiplicative component G_k is equivalent to a symmetric (permutation-average) attribution in log space: $\log G_k$ is the arithmetic mean of marginal log-ratios for input k over all admissible orderings and both periods.

D Practical protocol for handling cross-period infeasibility (recommended reporting standard)

Cross-period infeasibility arises when a mixed-period projection (e.g., $F(x_{t+1}, t)$ or $F(x_t, t+1)$) is undefined because the evaluation input vector lies outside the feasible input support of the reference technology. This issue can affect TECH as well as ACCUM and its input-specific factors because both rely on cross-period evaluations.

To ensure transparency and replicability, we recommend the following protocol.

Step 1 (Diagnose)

For each unit and period transition, attempt to compute all projection values required for EFF, TECH, ACCUM, and (when applicable) the counterfactual projections needed for the input-specific decomposition (e.g., $F(p_{10}, t)$, $F(p_{01}, t)$ in the two-input case). Record which specific projection(s) fail.

Step 2 (Preferred remedy: technology enlargement, when conceptually acceptable)

If the empirical design permits, adopt a technology construction that enlarges the reference set and mitigates infeasibility, such as a sequential (cumulative) technology or a global technology. This remedy preserves the definition of projection mappings while reducing endpoint failures.

Step 3 (Conservative remedy: transparent sample restriction)

If the technology construction is deliberately kept contemporaneous (year-by-year) and changing it would alter the intended interpretation, apply a clearly stated sample restriction: exclude units/periods for which the required projection values are undefined, and report (i) the list of excluded units, (ii) which projection(s) were infeasible, and (iii) the resulting sample size.

Step 4 (Endpoint adjustment, as a last resort)

If neither Step 2 nor Step 3 is desirable (e.g., infeasibility is rare but affects substantively important units), replace the infeasible endpoint by a nearest feasible

input vector under a documented rule (e.g., minimal proportional expansion of the infeasibility-causing input dimension(s) to reach feasibility). Any such adjustment should be fully documented and subjected to a sensitivity check.

Reporting principle

Regardless of the remedy, the paper should report the incidence of infeasibility and the handling choice in a way that allows exact replication. In the empirical illustration in Section 6, we follow Step 3 and apply an explicit sample restriction so that all required projection values are well-defined.

E Magnitude of ordering dependence in path-specific input accumulation (illustrative cases)

This appendix reports the magnitude of order (path) dependence in input-level attributions of the accumulation component when inputs are changed sequentially along alternative intermediate points. The main text reports the order-invariant (path-symmetrized) decomposition; here we provide the corresponding path-dependent components for the two admissible orders in the two-input setting.

Table 3 lists the path-specific components computed via p_{10} (changing input 1 first) and via p_{01} (changing input 2 first), together with gap, defined as the maximum absolute deviation from the symmetrized components reported in the main table. Larger values indicate stronger sensitivity of the input-level attribution to the chosen input-change order.

Table 3 shows that, although ACCUM itself is invariant to the ordering of input changes, the input-level attribution can be materially order-dependent when a single path is used (via p_{10} versus via p_{01}). In many countries the gap is close to zero, indicating that the two admissible paths assign almost the same contributions to k and h . However, there are also cases where the attribution changes noticeably with the chosen path: for example, Cambodia exhibits a very large gap (≈ 0.52), and Pakistan and several European economies show gaps on the order of a few hundredths (e.g., Spain, Italy, Portugal; similarly for Türkiye). These magnitudes are non-negligible relative to the reported KACCUM/HACCUM components and imply that a path-specific decomposition may deliver substantially different narratives about whether capital accumulation or human-capital accumulation "drives" ACCUM, depending solely on the input-change order. In some cases, the input-change order even flips the relative magnitudes: for Cambodia, HACCUM exceeds KACCUM via p_{10} (2.0599 vs. 1.1385), whereas KACCUM exceeds HACCUM via p_{01} (2.0293 vs. 1.1557). Similar rank reversals occur for Spain, Italy, and Portugal.

At the other extreme, Table 3 also contains cases in which the two admissible paths coincide exactly and one input-specific component equals unity (e.g., Canada, Switzerland, and the United States with HACCUM = 1; Cyprus and Iceland with KACCUM = 1). These rows illustrate the reduction property in Section 4.3: when the decomposition collapses to a single active input, the ordering becomes immaterial and symmetrization leaves the attribution unchanged.

Accordingly, Table 3 is best read as a diagnostic of potential path-induced instability—motivating the symmetrized, order-invariant decomposition in the main text—rather than as an alternative set of preferred estimates.

Table 3: Path-dependent decompositions via p_{10} and p_{01} and the resulting order-dependence gap (two-input case; 2000–2010)

Country	KACCUM(p_{10})	HACCUM(p_{10})	KACCUM(p_{01})	HACCUM(p_{01})	gap
Cambodia	1.1385	2.0599	2.0293	1.1557	0.5170
Pakistan	1.0463	1.3081	1.1565	1.1834	0.0639
Türkiye	1.3648	1.1236	1.4515	1.0565	0.0440
Spain	1.1031	1.1255	1.1894	1.0438	0.0440
Ireland	1.5921	1.0652	1.6702	1.0154	0.0395
Italy	1.0246	1.1168	1.0887	1.0510	0.0334
Norway	1.1839	1.0513	1.2422	1.0019	0.0295
Portugal	1.0273	1.1246	1.0828	1.0670	0.0292
Indonesia	1.4671	1.0564	1.4117	1.0979	0.0280
Viet Nam	2.0195	1.1978	2.0724	1.1672	0.0266
Hungary	1.3646	1.0507	1.4036	1.0216	0.0196
France	1.0116	1.0603	1.0481	1.0233	0.0187
Thailand	1.2383	1.0884	1.2054	1.1182	0.0166
Brazil	1.2166	1.0966	1.1842	1.1265	0.0163
Singapore	1.2810	1.0559	1.3088	1.0335	0.0139
Romania	1.6633	1.0161	1.6900	1.0000	0.0134
Belgium	1.0427	1.0421	1.0692	1.0163	0.0133
China	1.5845	1.0475	1.6101	1.0308	0.0129
Bulgaria	1.6149	1.0146	1.6385	1.0000	0.0118
Dominican Republic	1.2885	1.0528	1.2660	1.0715	0.0113
Bangladesh	1.6062	1.1733	1.6268	1.1584	0.0103
China, Hong Kong SAR	1.0412	1.1218	1.0602	1.1017	0.0101
Poland	1.2684	1.0147	1.2870	1.0000	0.0094
Chile	1.4228	1.0273	1.4375	1.0167	0.0074
Austria	1.1099	1.0196	1.0988	1.0299	0.0056
Netherlands	1.1151	1.0146	1.1045	1.0244	0.0053
Finland	1.1014	1.0231	1.0910	1.0328	0.0052
Malaysia	1.2247	1.0416	1.2145	1.0504	0.0051
Latvia	1.2014	1.0388	1.1924	1.0467	0.0045
Luxembourg	1.0008	1.0502	1.0094	1.0413	0.0045
Costa Rica	1.1929	1.0223	1.1841	1.0299	0.0044
Colombia	1.0854	1.0773	1.0785	1.0842	0.0035
India	1.5279	1.1218	1.5347	1.1168	0.0034
Malta	1.2170	1.0199	1.2119	1.0243	0.0026
Mexico	1.1773	1.0254	1.1722	1.0299	0.0026
Slovenia	1.1222	1.0095	1.1273	1.0050	0.0025
Denmark	1.1815	1.0070	1.1854	1.0037	0.0020

Country	KACCUM(p_{10})	HACCUM(p_{10})	KACCUM(p_{01})	HACCUM(p_{01})	gap
Croatia	1.3149	1.0239	1.3116	1.0265	0.0017
Sri Lanka	1.4304	1.0023	1.4337	1.0000	0.0016
Philippines	1.0642	1.0471	1.0614	1.0499	0.0014
Republic of Korea	1.4714	1.0052	1.4739	1.0035	0.0012
Lithuania	1.2471	1.0210	1.2455	1.0223	0.0008
Japan	1.1551	1.0019	1.1564	1.0008	0.0006
Estonia	1.5934	1.0007	1.5945	1.0000	0.0006
Greece	1.1047	1.0500	1.1036	1.0510	0.0005
Russian Federation	1.3278	1.0063	1.3288	1.0055	0.0005
Slovakia	1.1546	1.0034	1.1556	1.0025	0.0005
Uruguay	1.1700	1.0047	1.1690	1.0055	0.0005
Argentina	1.0339	1.0235	1.0331	1.0243	0.0004
Jamaica	1.1392	0.9966	1.1397	0.9962	0.0003
Taiwan	1.0064	1.0829	1.0061	1.0833	0.0002
Australia	1.1599	0.9997	1.1596	1.0000	0.0002
New Zealand	1.1599	0.9967	1.1597	0.9969	0.0001
Sweden	1.0252	1.0178	1.0254	1.0176	0.0001
Peru	1.2016	1.0161	1.2016	1.0161	0.0000
Canada	1.2100	1.0000	1.2100	1.0000	0.0000
Switzerland	1.2344	1.0000	1.2344	1.0000	0.0000
Cyprus	1.0000	1.1377	1.0000	1.1377	0.0000
Czechia	1.0898	1.0000	1.0898	1.0000	0.0000
Germany	1.2028	1.0000	1.2028	1.0000	0.0000
United Kingdom	1.0352	1.0000	1.0352	1.0000	0.0000
Iceland	1.0000	1.0992	1.0000	1.0992	0.0000
Israel	0.9729	1.0000	0.9729	1.0000	0.0000
United States	1.2397	1.0000	1.2397	1.0000	0.0000

Country	KACCUM(p_{10})	HACCUM(p_{10})	KACCUM(p_{01})	HACCUM(p_{01})	gap
---------	--------------------	--------------------	--------------------	--------------------	-----

Notes: $p_{00} = (x_t^{(1)}, x_t^{(2)})$, $p_{10} = (x_{t+1}^{(1)}, x_t^{(2)})$, $p_{01} = (x_t^{(1)}, x_{t+1}^{(2)})$, and $p_{11} = (x_{t+1}^{(1)}, x_{t+1}^{(2)})$. Let $F(p, t)$ denote the DEA projection mapping defined in Section 5.1 (output-oriented VRS). The path-specific components are:

$$\text{KACCUM}(p_{10}) = \left[\frac{F(p_{10}, t)}{F(p_{00}, t)} \cdot \frac{F(p_{10}, t+1)}{F(p_{00}, t+1)} \right]^{1/2},$$

$$\text{HACCUM}(p_{10}) = \left[\frac{F(p_{11}, t)}{F(p_{10}, t)} \cdot \frac{F(p_{11}, t+1)}{F(p_{10}, t+1)} \right]^{1/2}.$$

$$\text{KACCUM}(p_{01}) = \left[\frac{F(p_{11}, t)}{F(p_{01}, t)} \cdot \frac{F(p_{11}, t+1)}{F(p_{01}, t+1)} \right]^{1/2},$$

$$\text{HACCUM}(p_{01}) = \left[\frac{F(p_{01}, t)}{F(p_{00}, t)} \cdot \frac{F(p_{01}, t+1)}{F(p_{00}, t+1)} \right]^{1/2}.$$

The order-invariant components KACCUM and HACCUM are reported in Table 2. The gap is

$$\text{gap} = \max\{|\text{KACCUM} - \text{KACCUM}(p_{10})|, |\text{HACCUM} - \text{HACCUM}(p_{10})|, |\text{KACCUM} - \text{KACCUM}(p_{01})|, |\text{HACCUM} - \text{HACCUM}(p_{01})|\}.$$