Two-Step Nonlinear ARDL Estimation of the Relationship between R&D Intensity and Investment*

JIN SEO CHO
School of Economics, Yonsei University, Seoul, South Korea
Email: jinseocho@yonsei.ac.kr

MATTHEW GREENWOOD-NIMMO
Faculty of Business and Economics, University of Melbourne, Carlton, VIC 3053, Australia
Melbourne Centre for Data Science, University of Melbourne, Parkville, VIC 3010, Australia
Centre for Applied Macroeconomic Analysis, Australian National University, Canberra, ACT 2600, Australia
Codera Analytics, Johannesburg, GP 2193, South Africa
Email: matthew.greenwood@unimelb.edu.au

YONGCHEOL SHIN
Department of Economics and Related Studies, University of York, York, U.K.
Email: yongcheol.shin@york.ac.uk

This version: December 2023

Abstract

We investigate the relationship between R&D expenditure and investment using a nonlinear autoregressive distributed lag (NARDL) model. We propose a two-step NARDL (2SNARDL) estimation framework, that is free from the asymptotic singular matrix issue. We first estimate the parameters of the long-run relationship by the fully-modified least squares estimator and estimate the dynamic parameters in the next step. Under mild regularity conditions, we establish consistency of the 2SNARDL estimator and derive the relevant limit distributions. We also develop Wald test statistics for the hypotheses of short-run and long-run parameter asymmetry. By applying the 2SNARDL estimation procedure we find that R&D expenditure is asymmetrically associated with investment in the U.S.

Key Words: Asymmetric Relationship between R&D Expenditure and Investment; Asymptotic Singularity; Two-step Estimation of Nonlinear Autoregressive Distributed Lag Model.

JEL Classifications: C22, E22, O32.

---

*This is a substantially revised version of an earlier manuscript circulated under the title "Two-Step Estimation of the Nonlinear Autoregressive Distributed Lag Model." We are grateful for the insightful comments of In Choi, Tae-Hwan Kim, Rui Lin, Viet Nguyen, Barbara Rossi and Michael Thornton, as well as seminar participants at the Universities of Melbourne, Yonsei, and York. Cho is grateful for financial support from the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (Grant number NRF-2019S1A5A2A01035568). Greenwood-Nimmo and Shin acknowledge financial support from the Economic and Social Research Council (Grant number ES/T01573X/1). The views expressed herein are those of the authors and should not be reported as the views of Codera Analytics. The usual disclaimer applies.

1 Introduction

The potentially asymmetric relationship between research and development (R&D) expenditure and physical investment has received little attention, despite the growing literature on innovation and growth. R&D expenditure is typically characterized as an investment in knowledge creation that may lead to an innovation. Such innovation requires physical investment for its implementation (see Lach and Rob, 1996). But, this simple mechanism fails to distinguish among different types of R&D activities. For example, the product life cycle literature distinguishes between early-stage R&D that is often associated with new product development, and later-stage R&D that tends to focus on scaling production and achieving efficiency gains (e.g. Gort and Wall, 1986; Audretsch, 1987).

In this paper, we aim to make three contributions. First, we develop a theory relating early-stage and later-stage R&D expenditures to physical investment. Henceforth, we refer to early-stage and later-stage R&D as *innovative* and *managerial R&D*. Innovative R&D expenditure determines the scope of production as it creates a new product or technology through the discovery of a novel (knowledge) production function. The more innovative R&D activity, the larger the scale of production, implying that the production activity is mainly determined by the size of innovative R&D activity. On the contrary, managerial R&D activity does not necessarily create a new product, but produces an existing product more efficiently, implying that less physical capital is required per unit of output. We derive a theoretical prediction that innovative R&D expenditure is a complement to physical investment while managerial R&D expenditure is a substitute and develop the testable hypothesis that innovative R&D expenditure is positively related with investment by virtue of complementarity while the relationship between managerial R&D expenditure and physical investment may be negative due to their nature as substitutes.

Second, we propose a corresponding empirical specification that takes the form of a NARDL model[1] in which real investment is regressed on the positive and negative partial sums of R&D in-

[1]The NARDL model proposed by Shin, Yu, and Greenwood-Nimmo (2014, hereafter SYG) is an asymmetric generalization of the ARDL model of Pesaran and Shin (1998) and Pesaran, Shin, and Smith (2001). It is a single-equation error-correction model that has been widely applied to accommodate asymmetry in the long-run equilibrium relationship and/or the short-run dynamic coefficients via the use of partial sum decomposition of the independent variable(s). The NARDL approach has grown in popularity, with applications in fields including criminology (Box, Gratzer, and Lin, 2019), economic growth (Eberhardt and Presbitero, 2015), energy economics (Hammoudeh, Lahiani, Nguyen, and Sousa, 2015), exchange rates and trade (Verheyen, 2013; Brun-Aguerre, Fuertes, and Greenwood-Nimmo, 2017), financial economics (He and Zhou, 2018), health economics (Barati and Fariditavana, 2020), the tourism (Süssmuth and Woitek, 2013)

tensity (defined as the ratio of R&D expenditure to GDP). Noting that innovative (managerial) R&D expenditures are larger (smaller) than the output that they generate, the sign of the growth rate of R&D intensity (denoted Δr_t) can dictate the relative prevalence of innovative and managerial R&D activities. If $\Delta r_t \geq 0$, then R&D expenditure grows as fast as output, indicating the prevalence of innovative R&D activity. If $\Delta r_t < 0$, then managerial R&D activity prevails. Furthermore, the NARDL model has two main attributes. First, the NARDL model can provide a natural vehicle to test for sign/magnitude asymmetry in the relationship between R&D intensity and investment. Next, importantly, we show that the NARDL model can analyze an asymmetric cointegrating relationship between two integrated variables with different time drifts without the need to include a deterministic time trend in the model.[2]

Third, we note in passing that theoretical foundations for estimation of and inference on the NARDL model have yet to be fully developed. SYG estimate the NARDL parameters in a single step by ordinary least squares (OLS), though the positive and negative partial sums of the regressors are dominated by deterministic trends that are asymptotically perfectly collinear. This introduces an asymptotic singularity issue that frustrates efforts to derive the limit distribution of the single-step NARDL estimator. We propose an analytically tractable two-step NARDL estimation procedure that overcomes this issue. For a bivariate model, the asymmetric long-run relationship is expressed among the level of the dependent variable and the positive and negative cumulative partial sums of the explanatory variable, leading to a singular matrix. We resolve this issue by replacing one of the partial sum processes with an original regressor and show that the long-run parameters from this re-parameterized model can be consistently estimated. By applying the fully-modified OLS (FM-OLS) estimator of Phillips and Hansen (1990) in the first step, we obtain a super-consistent estimator that asymptotically follows a mixed normal distribution. Given the super-consistency of the long-run parameter estimator, the error-correction term can be treated as known in the second step such that OLS yields a consistent and asymptotically normal estimator of the short-run parameters. We also develop standard Wald tests for inference on the short- and long-run parameters. A suite of Monte Carlo

and political science (Ferris, Winer, and Olmstead, 2020), to list only a few. See Cho, Greenwood-Nimmo, and Shin (2023b) for an extensive survey.

[2]In empirical application, we find that physical investment is a unit root process with a upward trend while R&D intensity is a trendless unit root process. In a linear model, to capture the cointegrating relationship between these variables, we may include a deterministic time trend as an additional regressor. This is a practical but inefficient solution.

simulations indicates that our asymptotic results offer good approximations in finite samples.

We demonstrate the utility of our proposed approach with an application to the asymmetric relationship between R&D intensity and physical investment using quarterly data in the U.S. covering the period from 1960q1 to 2019q4. Our overall finding suggests that, in the long-run, physical investment responds positively to R&D expenditures when their growth rate exceeds the growth rate of GDP, but negatively when they grow more slowly than GDP. This supports our theoretical predictions regarding the nature of innovative (managerial) R&D expenditure as a complement to (substitute for) physical investment. Furthermore, we find that physical investment is more sensitive to changes in R&D intensity when managerial R&D activity prevails.

This paper proceeds in 7 sections. In Section 2, we review the literature on R&D expenditure and investment and propose a novel theoretical model of their relationship. In Sections 3 and 4, we introduce the NARDL model and explore the asymptotic singularity problem. We then introduce our two-step estimation framework and develop Wald tests for the null hypotheses of short-run and long-run symmetry. In Section 5, we scrutinize the finite sample properties of our estimators and test statistics by simulation. Section 6 is devoted to our empirical analysis of the relationship between R&D intensity and physical investment. Section 7 concludes. Proofs and additional simulation and estimation results are relegated to the Online Supplement.

2 Literature Review and Stylized Model

In this section, we review the literature and develop a theory that predicts an asymmetric relationship between R&D intensity and investment.

2.1 The Relationship between R&D Expenditure and Investment

Following Schumpeter's seminal 1942 work on creative destruction, a large literature has emerged on aspects of R&D activities. Important contributions include Utterback and Abernathy (1975), who find that R&D activity is conducted differently across the different stages of product life, the product life-cycle theory developed by Gort and Wall (1986) and Audretsch (1987), and the game-theoretic approach to R&D activity associated with Kamien and Schwartz (1972), Reinganum (1982), Fudenberg, Gilbert, Stiglitz, and Tirole (1983), Grossman and Shapiro (1987) and Harris and Vickers (1987).

It is common to distinguish between two different stages of R&D activity. Early-stage (innovative) R&D expenditure focuses on the development of a new product or technology, leading to a subsequent large-scale investment. By contrast, later-stage (managerial) R&D expenditure focuses on improvements to production efficiency. Consequently, managerial R&D expenditure should not exceed the expected increase in output, which results in a smaller-scale investment than innovative R&D. Overall, R&D expenditure tends to increase sharply in the early stage before leveling off or decreasing at the later stage.

A number of theoretical studies differentiate between innovative and managerial R&D activities and their effects on other economic variables, including Klepper (1996, 1997) and Agarwal and Audretsch (2001). Comin and Philippon (2005) and Aghion, Blundell, Griffth, Howitt, and Prantl (2009) empirically examine the relationship between the entry and/or exit rate of firms and innovative R&D expenditure. Similar theories have been developed in other fields including engineering and management—for example, Zif and McCarthy (1997), who classify R&D activity into multiple stages following the product life-cycle theory (see also Chung and Shin, 2020).

However, one area in which the distinction between innovative and managerial R&D activity is yet to be fully investigated is the relationship between R&D expenditure and investment. Early studies (e.g. Schmookler, 1966) focus on the causal relationship between R&D expenditure and investment. Applying vector autoregressions (VARs) to firm- and industry-level data, Lach and Schankerman (1989) and Lach and Rob (1996) find that R&D expenditure Granger causes investment but not *vice versa*. However, using longer time series, Chiao (2001) documents a two-way causal relationship between the growth rates of the R&D expenditure and investment. Employing a vector error-correction (VEC) model, Baussola (2000) documents evidence in favor of unidirectional Granger causality from R&D expenditure to investment. These results should be treated with care because the failure to distinguish between innovative and managerial R&D activity undermines efforts to accurately capture the potentially asymmetric relationship between R&D expenditure and investment.

2.2 The Asymmetric Relationship between R&D Intensity and Investment

We propose a theoretical model relating innovative and managerial R&D expenditures to investment. Innovative R&D expenditure determines the scope of production, as it describes a research activity that creates a new product or technology through the discovery of a novel production function. The more innovative R&D activity, the larger the scale of production, suggesting that the limit of production activity is determined by the amount of innovative R&D activity. By contrast, managerial R&D activity does not create a new product, but instead produces an existing product more efficiently, implying that less physical capital is required per unit of output.

Let k and y be the levels of physical capital and output, while we let c and s be the capital levels converted from innovative and managerial R&D expenditures, respectively. We assume that c is complementary to production activity conducted using physical capital while s is a substitute. Consider a production function embodying this mechanism as follows:

$$y = \min[c, k + s]. \tag{1}$$

The complementary relationship with c limits production activity, as output cannot be produced in excess of the level of innovative R&D activity. On the other hand, managerial R&D activity can produce capital s that substitutes for k.[3]

We use a dynamic optimization approach and apply the q-theory of investment to examine how physical investment responds to external shocks to R&D expenditure. First, physical capital, k_t, is formed by accumulating physical investment, i_t, through $\dot{k}_t = i_t - \delta k_t$, where δ is the depreciation rate of the physical capital. Similarly, c_t and s_t, are accumulated through $\dot{c}_t = r_t - \tau c_t$ and $\dot{s}_t = d_t - \gamma s_t$, where τ and γ denote depreciation rates of c_t and s_t, respectively. Next, consider the cost functions associated with converting R&D expenditures and physical investment into capital. Let $\kappa(r_t)$, $\xi(d_t)$ and $\phi(i_t)$ be the cost levels from innovative and managerial R&D expenditure and physical investment, respectively. We assume that the cost functions are convex with respect to R&D expenditures and physical investment, such that $\kappa'(\cdot) > 0$, $\xi'(\cdot) > 0$, $\phi'(\cdot) > 0$, $\kappa''(\cdot) > 0$, $\xi''(\cdot) > 0$, and $\phi''(\cdot) > 0$.[4]

[3] In the Online Supplement, we present a comparative analysis to show how the three types of capital interact in response to external shocks.

[4] For example, incidental and/or additional costs may be incurred to convert R&D expenditures into capital for produc-

To generalize the production function (1) into a differentiable function, we assume that output is given by $y_t = f(c_t, k_t + s_t)$, where $f(\cdot,\cdot)$ is a differentiable function with $f_c(c,a) > 0$, $f_a(c,a) > 0$, $f_{cc}(c,a) < 0$, $f_{aa}(c,a) < 0$, and $f_{ca}(c,a) > 0$ uniformly on the space of (c,a). Notice that k_t and s_t are strictly substitutes, while c_t and $a_t := k_t + s_t$ are weakly complements.[5]

The representative firm determines the optimal path of capital, investment, and R&D expenditures by maximizing discounted aggregate profit:

$$\max_{\{c_t,k_t,s_t,r_t,i_t,d_t\}} \int_0^\infty \{f(c_t, k_t + s_t) - \kappa(r_t) - \phi(i_t) - \xi(d_t)\} e^{-\rho t} dt$$
$$\text{subject to } \dot{c}_t = r_t - \tau c_t, \quad \dot{k}_t = i_t - \delta k_t \text{ and } \dot{s}_t = d_t - \gamma s_t,$$

where ρ is the discount rate. This extends the standard q-theory of investment by considering the role of capital converted from R&D expenditures as well as physical capital, with the three different accumulation rules as constraints.[6]

To analyze the long-run relationship between physical investment and R&D expenditures, we set up the dynamic optimization problem. Let $(c_*, k_*, s_*, r_*, i_*, d_*)$ be the steady-state equilibrium, which must satisfy the steady-state conditions given by $f_a(c_*, k_* + s_*) = (\delta + \rho)\phi'(i_*)$, $f_a(c_*, k_* + s_*) = (\gamma + \rho)\xi'(d_*)$, $f_c(c_*, k_* + s_*) = (\tau + \rho)\kappa'(r_*)$, $i_* = \delta k_*$, $d_* = \gamma s_*$, and $r_* = \tau c_*$. To display the steady-state equilibrium, we plot a set of phase diagrams and marginal cost functions in Figure 1. The panels on the left show the phase diagrams of (r_t, c_t), (i_t, k_t), and (d_t, s_t), while those on the right display the marginal cost functions of innovative R&D expenditure, physical investment, and managerial R&D expenditure, respectively. The phase diagrams also indicate the steady-state equilibrium levels of the variables along with the stable arms denoted by the dotted lines, such that the steady-state equilibrium can be reached by moving toward the equilibrium following the arms.

---

tion. These include patent fees, monetary or non-monetary incentives for researchers, safety management fees, training costs, and so on. These costs and fees are assumed to form the convex functions.

[5]It is possible to define alternative production functions that exhibit a weakly substitutionary relationship between s_t and k_t. Suppose that the production function (1) can be generalized by the constant elasticity of substitution (CES) function, $\ell(x, y; \beta) := (x^{\frac{\beta-1}{\beta}} + y^{\frac{\beta-1}{\beta}})^{\frac{\beta}{\beta-1}}$. Then, we can derive the generalized twofold CES production function: $f(c_t, k_t, s_t; \beta, \sigma) := \ell(c_t, \ell(k_t, s_t; \sigma); \beta)$, from which it follows that $\lim_{\sigma\to\infty} \lim_{\beta\to 0} f(c_t, k_t, s_t; \beta, \sigma) = \min[c_t, k_t + s_t]$. Provided that σ, $\beta > 0$, we can apply optimization theory as the production function is differentiable.

[6]The representative firm is assumed to choose the optimal time paths of r_t and d_t simultaneously, ignoring the fact that innovative R&D activity is conducted earlier than managerial R&D activity. This is not restrictive, as the firm represents a multitude of firms in the economy, where innovative and managerial R&D activities are conducted simultaneously.

The equilibrium cannot be reached unless the initial levels of (r_0, c_0), (i_0, k_0), and (d_0, s_0) are on the stable arms, simultaneously, as it would violate the transversality conditions.

— Insert Figure 1 Here —

To examine how the steady-state equilibrium responds to external shocks, we conduct two experiments by changing the marginal cost functions of each type of R&D expenditure. In the first experiment, we let the marginal cost function of innovative R&D expenditure decrease from $\kappa_0'(\cdot)$ to $\kappa_1'(\cdot)$. Denote $(c_*,\ k_*,\ s_*,\ r_*,\ i_*,\ d_*)$ and $(c_{**},\ k_{**},\ s_{**},\ r_{**},\ i_{**},\ d_{**})$ as the initial and new steady-state equilibria. The adjustment processes are displayed in the left panel of Figure 2. This decline in the marginal cost function shifts the locus of $\dot{c}_t = 0$ to the locus of $\dot{c}_t' = 0$, denoted by the dashed line in the first phase diagram. The steady-state equilibrium (r_{**}, c_{**}) is reached at a level greater than (r_*, c_*). To attain the new steady-state equilibrium, r_* jumps to the new stable arm, denoted by the dotted line. By contrast, c_t is a stock, so it cannot jump to the new stable arm. Thus, (r_t, c_t) gradually tends to (r_{**}, c_{**}). Then, the loci of $\dot{i}_t = 0$ and $\dot{d}_t = 0$ move to $\dot{i}_t' = 0$ and $\dot{d}_t' = 0$, which are denoted by the dashed lines in the second and third diagrams. The steady-state equilibrium is determined at (i_{**}, k_{**}) and (d_{**}, s_{**}), which are greater than (i_*, k_*) and (d_*, s_*). Both i_t and d_t jump to the new stable arms denoted by the dotted lines, and (i_t, k_t) and (d_t, s_t) tend to the new steady-state equilibrium. This reveals that physical investment and innovative R&D expenditure move in the same direction following the change in the cost function of innovative R&D expenditure, implying that they are complements. The economic intuition is straightforward—as innovative R&D activities become relatively cheaper, the firm tends to accumulate more capital from innovative R&D activities, enhancing productivity. This allows the firm to invest more, thereby accumulating more physical capital.

— Insert Figure 2 Here —

In the second experiment, the marginal cost function of managerial R&D expenditure decreases from $\xi_0'(\cdot)$ to $\xi_1'(\cdot)$. The right panel in Figure 2 displays the adjustment process to the new equilibrium. The decrease in the marginal cost function shifts the locus of $\dot{d}_t = 0$ to the locus of $\dot{d}_t' = 0$, denoted by the dashed line in the third phase diagram. The new steady-state equilibrium (d_{**}, s_{**}) is reached at a level greater than (d_*, s_*). To attain the new steady-state equilibrium, d_* jumps to the new stable arm, denoted by the dotted line. However, s_t, as a stock, cannot jump to the new stable arm. Consequently, (d_t, s_t) tends to (d_{**}, s_{**}) gradually and the locus of $\dot{i}_t = 0$ shifts to the locus of $\dot{i}'_t = 0$. The steady-state equilibrium level is determined at (i_{**}, k_{**}), where k_t, as a stock, cannot jump to a new level while i_t jumps to the stable arm denoted as the dotted line in the second phase diagram. Overall, following the decrease in the marginal cost of managerial R&D expenditure, d_* rises to d_{**} but i_* falls to i_{**}, revealing a substitutionary relationship between i_t and d_t. This implies that physical capital decreases from k_* to k_{**}, while s increases to s_* from s_{**}.

As s_t and k_t move in opposite directions, $a_{**} := k_{**} + s_{**}$ can be greater or less than $a_* := k_* + s_*$. The sign of the change depends upon the functional shapes of $f_a(\cdot, \cdot)$, $\xi'(\cdot)$, $\phi'(\cdot)$, and the depreciation rates δ and γ. In Figure 2 under the assumption that $a_{**} > a_*$, the locus of $\dot{c}_t = 0$ is shifted to $\dot{c}'_t = 0$, and a new equilibrium is achieved at (r_{**}, c_{**}), as indicated in the first phase diagram. Then, r_t jumps to the new stable arm denoted as the dotted line, and (r_t, c_t) approaches (r_{**}, c_{**}) gradually. In this case, $r_{**} > r_*$ and $c_{**} > c_*$, which is achieved mainly by virtue of the complementary relationship between c_t and a_t. On the other hand, consider the case with a_{**} less than a_* in which case the locus of $\dot{c}'_t = 0$ is shifted to the left of $\dot{c}_t = 0$. Then, we obtain $r_{**} < r_*$ and $c_{**} < c_*$. The economic intuition of the substitutionary relationship between i_t and d_t is also straightforward. As managerial R&D activity becomes relatively cheaper, the firm tends to accumulate capital by converting managerial R&D expenditures, thereby substituting physical investment, implying that both physical investment and capital will decrease.

These experiments yield important testable implications—the relationship between physical investment and innovative R&D expenditure is expected to be positive by virtue of their complementarity, while the relationship between managerial R&D expenditure and investment is more likely to be negative due to their nature as substitutes.

3 A NARDL Model of R&D Expenditure and Investment

In this section, we develop a NARDL model to characterize the asymmetric relationship between R&D expenditure and physical investment and propose a tractable two-step estimation procedure.

3.1 NARDL Specification

Our empirical specification is grounded in two stylized features of R&D expenditure highlighted in Section 2.1 and the theory developed in Section 2.2. First, as innovative R&D expenditures tend to focus on product innovation, their scale is often large relative to output. This suggests that R&D expenditure is expected to grow faster than output in the early stage, where start-up costs are large and the scale of production typically small. Second, as managerial R&D expenditures focus on enhancing production efficiency, their scale is typically smaller than output.

Let r_t denote aggregate R&D intensity in the t-th period, defined as a ratio of aggregate R&D expenditure to GDP. Noting that aggregate R&D expenditure incorporates the spectrum of R&D activities conducted throughout the economy, the sign of Δr_t determines the relative prevalence of innovative and managerial R&D activities. If $\Delta r_t \geq 0$, then R&D expenditure grows as fast as output, indicating a prevalence of innovative R&D activity. By contrast, if $\Delta r_t < 0$, then output grows faster than R&D expenditure, indicating a prevalence of managerial R&D activity. Given the different characteristics of innovative and managerial R&D, it is reasonable to expect that the relationship between R&D intensity and physical investment may be asymmetric.

To analyze the potential asymmetric impacts of r_t on the log of investment, denoted i_t, in the short-run and the long-run, we consider the following asymmetric error-correction model:

$$\Delta i_t = \gamma_* + \rho_* u_{t-1} + \sum_{j=1}^{p-1} \varphi_{j*} \Delta i_{t-j} + \sum_{j=0}^{q-1} \pi_{j*}^{+} \Delta r_{t-j}^{+} + \sum_{j=0}^{q-1} \pi_{j*}^{-} \Delta r_{t-j}^{-} + e_t, \tag{2}$$

where $u_{t-1} = i_{t-1} - \beta_*^{+} r_{t-1}^{+} - \beta_*^{-} r_{t-1}^{-}$ is the asymmetric error correction term and e_t is a serially uncorrelated error term, given sufficiently large lag orders, p and q. Here, $\Delta r_t^{+} := \Delta r_t \mathbb{1}_{\{\Delta r_t \geq 0\}}$ and $\Delta r_t^{-} := \Delta r_t \mathbb{1}_{\{\Delta r_t < 0\}}$, where $\mathbb{1}_{\{\cdot\}}$ is an indicator function taking unity if the condition in brace is satisfied, and zero otherwise.

The process in (2) is equivalent to the NARDL(p, q) process advanced by SYG, $i_t = \gamma_* + \sum_{j=1}^{p} \phi_{j*} i_{t-j} + \sum_{j=0}^{q} (\theta_{j*}^{+\prime} r_{t-j}^{+} + \theta_{j*}^{-\prime} r_{t-j}^{-}) + e_t$. The NARDL process allows for both the long-run parameters, β_*^{+} and β_*^{-}, and the short-run parameters, π_{j*}^{+} and π_{j*}^{-}, to differ, enabling us to jointly analyze long- and short-run asymmetric relationships between R&D intensity and investment. Furthermore, it is important to notice that the NARDL process can accommodate a cointegrating relationship between integrated time series with mismatched time drifts. As $\Delta r_t^+ \geq 0$ and $\Delta r_t^- \leq 0$ with probability one even if $E(\Delta r_t) = 0$, the partial sum processes, r_t^+ and r_t^-, will be integrated series with positive and negative time drifts, respectively. Thus, if there exists an asymmetric cointegrating relationship between i_t and r_t, then the dependent variable, i_t should be an integrated series with a drift. This has the important implication that the NARDL model can analyze an asymmetric cointegrating relationship between two integrated variables with different drifts without the need to include a deterministic time trend in the model.[7]

3.2 Singularity Issue in the NARDL Model

We review the asymptotic theory for the NARDL (p, q) model advanced by SYG:

$$y_t = \gamma_* + \sum_{j=1}^{p} \phi_{j*} y_{t-j} + \sum_{j=0}^{q} \left(\boldsymbol{\theta}_{j*}^{+\prime} \boldsymbol{x}_{t-j}^{+} + \boldsymbol{\theta}_{j*}^{-\prime} \boldsymbol{x}_{t-j}^{-} \right) + e_t,$$

where $\boldsymbol{x}_t \in \mathbb{R}^k$, $\boldsymbol{x}_t^+ := \sum_{j=1}^{t} \Delta \boldsymbol{x}_j^+$, $\boldsymbol{x}_t^- := \sum_{j=1}^{t} \Delta \boldsymbol{x}_j^-$, $\Delta \boldsymbol{x}_t^+ := \max[\mathbf{0}, \Delta \boldsymbol{x}_t]$, and $\Delta \boldsymbol{x}_t^- := \min[\mathbf{0}, \Delta \boldsymbol{x}_t]$. The corresponding error-correction model is given by

$$\Delta y_t = \rho_* y_{t-1} + \boldsymbol{\theta}_*^{+\prime} \boldsymbol{x}_{t-1}^{+} + \boldsymbol{\theta}_*^{-\prime} \boldsymbol{x}_{t-1}^{-} + \gamma_* + \sum_{j=1}^{p-1} \varphi_{j*} \Delta y_{t-j} + \sum_{j=0}^{q-1} \left(\boldsymbol{\pi}_{j*}^{+\prime} \Delta \boldsymbol{x}_{t-j}^{+} + \boldsymbol{\pi}_{j*}^{-\prime} \Delta \boldsymbol{x}_{t-j}^{-} \right) + e_t, \quad (3)$$

for some ρ_*, $\boldsymbol{\theta}_*^+$, $\boldsymbol{\theta}_*^-$, γ_*, φ_{j*} $(j = 1, 2, \ldots, p-1)$, $\boldsymbol{\pi}_{j*}^+$, and $\boldsymbol{\pi}_{j*}^-$ $(j = 0, 1, \ldots, q-1)$, where $\{e_t, \mathcal{F}_t\}$ is a martingale difference sequence and $\mathcal{F}_t$ is the smallest σ-algebra driven by $\{y_{t-1}, \boldsymbol{x}_t^+, \boldsymbol{x}_t^-, y_{t-2}, \boldsymbol{x}_{t-1}^+, \boldsymbol{x}_{t-1}^-, \ldots\}$. If y_t is cointegrated with $(\boldsymbol{x}_t^{+\prime}, \boldsymbol{x}_t^{-\prime})'$, then we rewrite (3) as $\Delta y_t = \rho_* u_{t-1} + \gamma_* + \sum_{j=1}^{p-1} \varphi_{j*} \Delta y_{t-j} + \sum_{j=0}^{q-1} \left(\boldsymbol{\pi}_{j*}^{+\prime} \Delta \boldsymbol{x}_{t-j}^{+} + \boldsymbol{\pi}_{j*}^{-\prime} \Delta \boldsymbol{x}_{t-j}^{-} \right) + e_t$, where $u_{t-1} := y_{t-1} - \boldsymbol{\beta}_*^{+\prime} \boldsymbol{x}_{t-1}^+ - \boldsymbol{\beta}_*^{-\prime} \boldsymbol{x}_{t-1}^-$ is the error correction term, $\boldsymbol{\beta}_*^+ := -(\boldsymbol{\theta}_*^+ / \rho_*)$ and $\boldsymbol{\beta}_*^- := -(\boldsymbol{\theta}_*^- / \rho_*)$. Notice that u_t can be correlated with $\Delta \boldsymbol{x}_t$.

We show that the NARDL process can capture a cointegrating relationship between a unit-root process with a deterministic time drift and another unit-root process without a drift. Suppose that $\mathbb{E}[\Delta \boldsymbol{x}_t] \equiv \mathbf{0}$. Let $\boldsymbol{\mu}_*^+ := \mathbb{E}[\Delta \boldsymbol{x}_t^+]$ and $\boldsymbol{\mu}_*^- := \mathbb{E}[\Delta \boldsymbol{x}_t^-]$. Then, $\boldsymbol{\mu}_*^+ + \boldsymbol{\mu}_*^- \equiv \mathbf{0}$ by construction. Define $\boldsymbol{s}_t^+ := \Delta \boldsymbol{x}_t^+ - \boldsymbol{\mu}_*^+$ and $\boldsymbol{s}_t^- := \Delta \boldsymbol{x}_t^- - \boldsymbol{\mu}_*^-$ such that $\boldsymbol{x}_t^+ = \boldsymbol{\mu}_*^+ t + \sum_{j=1}^{t} \boldsymbol{s}_j^+$ and $\boldsymbol{x}_t^- = \boldsymbol{\mu}_*^- t +$

[7] In Section 6, we find that r_t is a unit-root process without a drift while i_t is a unit-root process with a drift.

$\sum_{j=1}^{t} \boldsymbol{s}_j^-$, showing that $\boldsymbol{x}_t^+$ and $\boldsymbol{x}_t^-$ are unit-root processes with time drifts. Notice, however, that Δy_t is not necessarily distributed around zero even if $\boldsymbol{x}_t$ is a unit-root process without a deterministic trend. From the NARDL(p, q) process, it is easily seen that $\delta_* := \mathbb{E}[\Delta y_t] = -\frac{1}{\rho_*}[\sum_{j=0}^{q}(\boldsymbol{\theta}_{j*}^+)'\boldsymbol{\mu}_*^+ + \sum_{j=0}^{q}(\boldsymbol{\theta}_{j*}^-)'\boldsymbol{\mu}_*^-]$ by $\mathbb{E}[u_t] = 0$, where $\rho_* := 1 - \sum_{j=1}^{p}\phi_{j*}$. Let $d_t := \Delta y_t - \delta_*$. Then, $y_t = \delta_* t + \sum_{j=1}^{t} d_j$, establishing that y_t can become a unit-root process with a drift unless $\delta_* = 0$.

SYG propose to estimate the parameters of the NARDL model in a single step by OLS. Nevertheless, as shown in Lemma 1, the OLS estimation suffers from the asymptotic singularity of the inverse matrix associated with the one-step NARDL estimator. We make the following regular assumptions:

Assumption 1. *(i)* $\{(\Delta\boldsymbol{x}_t', u_t)'\}$ *is a globally covariance stationary mixing process of* $(k+1)\times 1$ *vectors of* ϕ *of size* $-r/(2(r-1))$ *or* α *of size* $-r/(r-2)$ *and* $r > 2$*; (ii)* $\mathbb{E}[\Delta\boldsymbol{x}_t] = \boldsymbol{0}$, $\mathbb{E}[|\Delta\boldsymbol{x}_{ti}|^r] < \infty$ $(i = 1, 2, \ldots, k)$, $\mathbb{E}[|u_t|^r] < \infty$, *and* $\mathbb{E}[|e_t|^2] < \infty$*; (iii)* $\lim_{T\to\infty}$ *Var*$[T^{-1/2}\sum_{t=1}^{T}(\Delta\boldsymbol{x}_t', u_t)']$ *exists and is positive definite (PD); and (iv) for some* $(\rho_*, \boldsymbol{\theta}_*^{+\prime}, \boldsymbol{\theta}_*^{-\prime}, \gamma_*, \varphi_{1*}, \ldots, \varphi_{p-1*}, \boldsymbol{\pi}_{0*}^{+\prime}, \ldots, \boldsymbol{\pi}_{q-1*}^{+\prime}, \boldsymbol{\pi}_{0*}^{-\prime}, \ldots, \boldsymbol{\pi}_{q-1*}^{-\prime})'$, Δy_t *is generated by* (3)*, such that* $\{e_t, \mathcal{F}_t\}$ *is a martingale difference sequence and* $\mathcal{F}_t$ *is the smallest* σ*-algebra driven by* $\{y_{t-1}, \boldsymbol{x}_t^+, \boldsymbol{x}_t^-, y_{t-2}, \boldsymbol{x}_{t-1}^+, \boldsymbol{x}_{t-1}^-, \ldots\}$. □

For notational simplicity, we define $\boldsymbol{z}_t := [\boldsymbol{z}_{1t}' \vdots \boldsymbol{z}_{2t}']' := [y_{t-1}, \boldsymbol{x}_{t-1}^{+\prime}, \boldsymbol{x}_{t-1}^{-\prime} \vdots 1, \Delta\boldsymbol{y}_{t-1}', \Delta\boldsymbol{x}_t^{+\prime}, \ldots, \Delta\boldsymbol{x}_{t-q+1}^{+\prime}, \Delta\boldsymbol{x}_t^{-\prime}, \ldots, \Delta\boldsymbol{x}_{t-q+1}^{-\prime}]'$, where $\Delta\boldsymbol{y}_{t-1} := [\Delta y_{t-1}, \Delta y_{t-2}, \ldots, \Delta y_{t-p+1}]'$, and $\boldsymbol{z}_t$ is partitioned into nonstationary and stationary variables. Next, we partition $\boldsymbol{z}_{2t}$ as $\boldsymbol{z}_{2t} := [1 \vdots \boldsymbol{w}_t']' := [1 \vdots \boldsymbol{w}_{1t}' \vdots \boldsymbol{w}_{2t}' \vdots \boldsymbol{w}_{3t}']' := [1 \vdots \Delta\boldsymbol{y}_{t-1}' \vdots \Delta\boldsymbol{x}_t^{+\prime}, \ldots, \Delta\boldsymbol{x}_{t-q+1}^{+\prime} \vdots \Delta\boldsymbol{x}_t^{-\prime}, \ldots, \Delta\boldsymbol{x}_{t-q+1}^{-\prime}]'$. Define $\boldsymbol{\alpha}_* := [\boldsymbol{\alpha}_{1*}' \vdots \boldsymbol{\alpha}_{2*}']' := [\rho_*, \boldsymbol{\theta}_*^{+\prime}, \boldsymbol{\theta}_*^{-\prime} \vdots \gamma_*, \boldsymbol{\varphi}_*', \boldsymbol{\pi}_{0*}^{+\prime}, \ldots, \boldsymbol{\pi}_{q-1*}^{+\prime}, \boldsymbol{\pi}_{0*}^{-\prime}, \ldots, \boldsymbol{\pi}_{q-1*}^{-\prime}]'$, where $\boldsymbol{\varphi}_* := [\varphi_{1*}, \varphi_{2*}, \ldots, \varphi_{p-1*}]'$. Then, the single-step NARDL estimator is given by $\widehat{\boldsymbol{\alpha}}_T := (\sum_{t=1}^{T}\boldsymbol{z}_t\boldsymbol{z}_t')^{-1}\sum_{t=1}^{T}\boldsymbol{z}_t\Delta y_t = \boldsymbol{\alpha}_* + (\sum_{t=1}^{T}\boldsymbol{z}_t\boldsymbol{z}_t')^{-1}\sum_{t=1}^{T}\boldsymbol{z}_t e_t$.

In Lemma 1 we show that inference using $\widehat{\boldsymbol{\alpha}}_T$ is challenging because $\sum_{t=1}^{T}\boldsymbol{z}_t\boldsymbol{z}_t'$ is asymptotically singular.

Lemma 1. *Under Assumption 1, (i)* $T^{-3}\sum_{t=1}^{T}\boldsymbol{z}_{t1}\boldsymbol{z}_{t1}' \xrightarrow{\mathbb{P}} \mathrm{M}_{11} := \frac{1}{3}\boldsymbol{n}_1\boldsymbol{n}_1'$*; (ii)* $T^{-2}\sum_{t=1}^{T}\boldsymbol{z}_{1t}\,\boldsymbol{z}_{2t}' \xrightarrow{\mathbb{P}} \mathrm{M}_{12} := \frac{1}{2}\boldsymbol{n}_1\boldsymbol{n}_2'$*; and (iii)*

$$T^{-1}\sum_{t=1}^{T}\boldsymbol{z}_{2t}\boldsymbol{z}_{2t}' \xrightarrow{\mathbb{P}} \mathrm{M}_{22} := \begin{bmatrix} 1 & \boldsymbol{n}_2^{(1)\prime} \\ \boldsymbol{n}_2^{(1)} & \mathbb{E}[\boldsymbol{w}_t\boldsymbol{w}_t'] \end{bmatrix},$$

where $\boldsymbol{n}_1 := [\delta_*, \boldsymbol{\mu}_*^{+\prime}, \boldsymbol{\mu}_*^{-\prime}]'$, $\boldsymbol{n}_2 := [1, \boldsymbol{n}_2^{(1)}]$ *and* $\boldsymbol{n}_2^{(1)} := [\delta_* \boldsymbol{\iota}_{p-1}', \boldsymbol{\iota}_q' \otimes \boldsymbol{\mu}_*^{+\prime}, \boldsymbol{\iota}_q' \otimes \boldsymbol{\mu}_*^{-\prime}]'$ *with* $\boldsymbol{\iota}_a$ *being an* $a \times 1$ *vector of ones.* □

Lemma 1 implies that $\mathbf{D}_T^{-1}(\sum_{t=1}^T \boldsymbol{z}_t \boldsymbol{z}_t')\, \mathbf{D}_T^{-1} \xrightarrow{\mathbb{P}} \mathbf{M}_*$, where $\mathbf{D}_T := \text{diag}[T^{3/2}\mathbf{I}_{2+2k}, T^{1/2}\mathbf{I}_{p+2qk}]$, and $\mathbf{M}_*$ is a 2×2 block matrix whose i-th row and j-th column block is $\mathbf{M}_{ij}$, with $\mathbf{M}_{21} = \mathbf{M}_{12}'$. As $\mathbf{M}_*$ is singular, it is difficult to derive the limit distribution of $\widehat{\boldsymbol{\alpha}}_T$ directly, unless we can derive the limit distribution of the determinant of $\sum_{t=1}^T \boldsymbol{z}_t \boldsymbol{z}_t'$, which is analytically challenging.

3.3 Two-step NARDL Estimation

We develop an estimation methodology for the model, (2) with $k = 1$. We provide an extension to the model with the multiple regressors in the Online Supplement.

First-Step Estimation by OLS Following the seminal two-step estimation framework of Engle and Granger (1987), we first estimate the parameters of the long-run relationship, $y_t = \alpha_* + \beta_*^+ x_t^+ + \beta_*^- x_t^- + u_t$ by OLS. Let $\bar{\mathbf{D}}_T := \text{diag}[T^{1/2}, T^{3/2}\mathbf{I}_2]$ and $\boldsymbol{v}_t := (1, x_t^+, x_t^-)'$. Then, it follows that $\bar{\mathbf{D}}_T^{-1}\left(\sum_{t=1}^T \boldsymbol{v}_t \boldsymbol{v}_t'\right)\bar{\mathbf{D}}_T^{-1} \xrightarrow{\mathbb{P}} \mathbf{M}_{11}$. By Lemma 1(*i*), this is a singular matrix due to the collinearity of the trends in x_t^+ and x_t^-.

We propose to reparameterize the long-run relationship as $y_t = \alpha_* + \lambda_* x_t^+ + \eta_* x_t + u_t$, where $x_t \equiv x_t^+ + x_t^-$, $\lambda_* = \beta_*^+ - \beta_*^-$ and $\eta_* = \beta_*^-$. The long-run parameters can be estimated as $\beta_*^+ = \lambda_* + \eta_*$ and $\beta_*^- = \eta_*$. The OLS estimator of $\boldsymbol{\varrho}_* := (\alpha_*, \lambda_*, \eta_*)'$ is obtained by $\widehat{\boldsymbol{\varrho}}_T := (\widehat{\alpha}_T, \widehat{\lambda}_T, \widehat{\eta}_T)' := \arg\min_{\alpha,\, \lambda,\, \eta} \sum_{t=1}^T (y_t - \alpha - \lambda x_t^+ - \eta x_t)^2$, where $\widehat{\beta}_T^+ := \widehat{\lambda}_T + \widehat{\eta}_T$ and $\widehat{\beta}_T^- = \widehat{\eta}_T$. Letting $\boldsymbol{q}_t := (1, x_t^+, x_t)'$, then $\widehat{\boldsymbol{\varrho}}_T = \boldsymbol{\varrho}_* + (\sum_{t=1}^T \boldsymbol{q}_t \boldsymbol{q}_t')^{-1}(\sum_{t=1}^T \boldsymbol{q}_t u_t)$. Define $\boldsymbol{\Sigma}_* := \lim_{T\to\infty} T^{-1} \sum_{t=1}^T \sum_{s=1}^T \mathbb{E}[\boldsymbol{g}_t \boldsymbol{g}_s]$ and $[\mathcal{B}_x(\cdot), \mathcal{B}_u(\cdot)]' := \boldsymbol{\Sigma}_*^{1/2}\,[\mathcal{W}_x(\cdot), \mathcal{W}_u(\cdot)]'$, where $\boldsymbol{g}_t := [\Delta x_t, u_t]'$ and $[\mathcal{W}_x(\cdot), \mathcal{W}_u(\cdot)]'$ is a 2×1 vector of independent Wiener processes. Let $\sigma_*^{(i,j)}$ be the (i, j)-th element of $\boldsymbol{\Sigma}_*$. If u_t is serially uncorrelated and independent of Δx_t, then $\boldsymbol{\Sigma}_*$ and $[\mathcal{B}_x(\cdot), \mathcal{B}_u(\cdot)]'$ are reduced to diag $[\sigma_x^2, \sigma_u^2]$ and $[\sigma_x \mathcal{W}_x(\cdot), \sigma_u \mathcal{W}_u(\cdot)]'$, where $\sigma_x^2 := \lim_{T\to\infty} T^{-1} \sum_{t=1}^T \sum_{s=1}^T \mathbb{E}[\Delta x_t \Delta x_s]$ and $\sigma_u^2 := \mathbb{E}[u_t^2]$.

Lemma 2 provides the limit behaviors of the components constituting the OLS estimator:

Lemma 2. *Under Assumption 1, if $k = 1$ and $\mathbf{\Sigma}_*$ is PD, then (i)*

$$\widehat{\mathbf{Q}}_T := \widetilde{\mathbf{D}}_T^{-1}\left(\sum_{t=1}^{T} \boldsymbol{q}_t \boldsymbol{q}_t'\right)\widetilde{\mathbf{D}}_T^{-1} \Rightarrow \boldsymbol{\mathcal{Q}} := \begin{bmatrix} 1 & \frac{1}{2}\mu_*^+ & \int \mathcal{B}_x \\ \frac{1}{2}\mu_*^+ & \frac{1}{3}\mu_*^+\mu_*^+ & \mu_*^+ \int r\mathcal{B}_x \\ \int \mathcal{B}_x & \int r\mathcal{B}_x \mu_*^+ & \int \mathcal{B}_x\mathcal{B}_x \end{bmatrix},$$

where $\widetilde{\mathbf{D}}_T := diag[T^{1/2}, T^{3/2}, T]$; and (ii) if $\upsilon_* := \lim_{T\to\infty} T^{-1}\sum_{t=1}^{T}\sum_{i=1}^{t} \mathbb{E}[\Delta x_i u_t]$ is finite, then $\widehat{\mathbf{U}}_T := \widetilde{\mathbf{D}}_T^{-1}(\sum_{t=1}^{T} \boldsymbol{q}_t u_t) \Rightarrow \boldsymbol{\mathcal{U}} := [\int d\mathcal{B}_u, \mu_*^+ \int r d\mathcal{B}_u, \int \mathcal{B}_x d\mathcal{B}_u + \upsilon_*]$.[8] □

$\boldsymbol{\mathcal{Q}}$ is nonsingular with probability 1, and the limit distribution of $\widehat{\boldsymbol{\varrho}}_T$ is obtained as a product of $\boldsymbol{\mathcal{Q}}^{-1}$ and $\boldsymbol{\mathcal{U}}$, as stated in corollary 1:

Corollary 1. *Under Assumption 1, if $k = 1$ and $\mathbf{\Sigma}_*$ is PD, then $\widetilde{\mathbf{D}}_T(\widehat{\boldsymbol{\varrho}}_T - \boldsymbol{\varrho}_*) \Rightarrow \boldsymbol{\mathcal{Q}}^{-1}\boldsymbol{\mathcal{U}}$.* □

Corollary 1 has important implications. First, by the reparameterization, the collinearity between x_t^+ and x_t^- can be avoided because $\sum_{t=1}^{T} x_t = O_{\mathbb{P}}(T^{3/2})$ and $\sum_{t=1}^{T} x_t^+ = O_{\mathbb{P}}(T^2)$ lead to different convergence rates for $\widehat{\lambda}_T$ and $\widehat{\eta}_T$, viz., $\widehat{\lambda}_T - \lambda_* = O_{\mathbb{P}}(T^{-3/2})$ and $\widehat{\eta}_T - \eta_* = O_{\mathbb{P}}(T^{-1})$. Next, to derive the limit distribution of $\widehat{\boldsymbol{\varrho}}_T$, we apply the FCLT only to $\sum_{t=1}^{T(\cdot)} \boldsymbol{g}_t$ where $\boldsymbol{g}_t := [\Delta x_t, u_t]'$, not to $\sum_{t=1}^{[T(\cdot)]}(x_t^+ - \mu_*^+)$. Third, using the definition of $\widehat{\lambda}_T$, we have: $T\{(\widehat{\beta}_T^+ - \widehat{\beta}_T^-) - (\beta_*^+ - \beta_*^-)\} = O_{\mathbb{P}}(T^{-1/2})$. Hence, $T(\widehat{\beta}_T^+ - \beta_*^+) = T(\widehat{\beta}_T^- - \beta_*^-) + o_{\mathbb{P}}(1)$, implying that the limit distributions of $T(\widehat{\beta}_T^+ - \beta_*^+)$ and $T(\widehat{\beta}_T^- - \beta_*^-)$ are asymptotically equivalent. Finally, as the long-run parameter estimator is super-consistent, $\widehat{\beta}_T^+$ and $\widehat{\beta}_T^-$ can be treated as known when estimating the short-run dynamic parameters.

Theorem 1. *Under Assumption 1 and for $k = 1$, $T[(\widehat{\beta}_T^+ - \beta_*^+), (\widehat{\beta}_T^- - \beta_*^-)]' \Rightarrow \boldsymbol{\iota}_2 \otimes \mathbf{S}\boldsymbol{\mathcal{Q}}^{-1}\boldsymbol{\mathcal{U}}$, where $\mathbf{S} := [\mathbf{0}_{1\times 2}, 1]$.* □

First-Step Estimation by FM-OLS The limit distribution in Theorem 1 is non-normal due to its dependence on the nuisance parameters $\mathbf{\Sigma}_*$ and υ_*. Furthermore, except in the special case where $\{u_t\}$ is independent of $\{\Delta x_t\}$ and/or serially uncorrelated, the OLS estimator of the long-run parameter exhibits an asymptotic bias determined by υ_*. Phillips and Hansen (1990) propose the FM-OLS estimator, which is shown to be free from asymptotic biases even in the presence of endogenous re-

[8] Here, all integrals are computed with respect to $r \in [0, 1]$. For example, $\int r\mathcal{B}_x$ denotes $\int_0^1 r\mathcal{B}_x(r)dr$.

gressors and/or serial correlation, and which follows an asymptotic mixed normal distribution. Hence, we advocate the use of FM-OLS to estimate the long-run parameters in the first step.

Suppose that $\mathbf{\Sigma}_*$ is consistently estimated by a heteroskedasticity and autocorrelation consistent covariance matrix estimator, $\widetilde{\mathbf{\Sigma}}_T$, whose (i,j)-th element is denoted by $\widetilde{\sigma}_T^{(i,j)}$. For example, following Newey and West (1987), we have $\widetilde{\mathbf{\Sigma}}_T := \frac{1}{T}\sum_{t=1}^T \widehat{\boldsymbol{g}}_t \widehat{\boldsymbol{g}}_t' + \frac{1}{T}\sum_{k=1}^{\ell} \omega_{\ell k} \sum_{t=k+1}^T \{\widehat{\boldsymbol{g}}_{t-k}\widehat{\boldsymbol{g}}_t' + \widehat{\boldsymbol{g}}_t \widehat{\boldsymbol{g}}_{t-k}'\}$, where $\widehat{\boldsymbol{g}}_t := [\Delta x_t, \widehat{u}_t]'$, $\omega_{\ell k} := 1 - k/(1+\ell)$, $\ell = O(T^{1/4})$ and $\widehat{u}_t := y_t - \widehat{\alpha}_T - \widehat{\beta}_T^+ x_t^+ - \widehat{\beta}_T^- x_t^-$. Under mild regularity conditions, it is straightforward to show that the asymptotic bias, υ_* in $\boldsymbol{\mathcal{U}}$, can be consistently estimated by $\widetilde{\mathbf{\Pi}}_T := \frac{1}{T}\sum_{k=0}^{\ell}\sum_{t=k+1}^T \widehat{\boldsymbol{g}}_{t-k}\widehat{\boldsymbol{g}}_t'$. Let $\widetilde{\pi}_T^{(i,j)}$ be the (i,j)-th element of $\widetilde{\mathbf{\Pi}}_T$. Define the long-run parameter estimator: $\widetilde{\boldsymbol{\varrho}}_T := (\widetilde{\alpha}_T, \widetilde{\lambda}_T, \widetilde{\eta}_T)' := (\sum_{t=1}^T \boldsymbol{q}_t \boldsymbol{q}_t')^{-1} (\sum_{t=1}^T \boldsymbol{q}_t \widetilde{y}_t - T\mathbf{S}'\widetilde{\upsilon}_T)$, where $\widetilde{y}_t := y_t - \Delta x_t (\widetilde{\sigma}_T^{(1,1)})^{-1} \widetilde{\sigma}_T^{(1,2)}$ and $\widetilde{\upsilon}_T := \widetilde{\pi}_T^{(1,2)} - \widetilde{\pi}_T^{(1,1)} (\widetilde{\sigma}_T^{(1,1)})^{-1}\, \widetilde{\sigma}_T^{(1,2)}$. Finally, the FM-OLS estimators of the long-run parameters are obtained as $\widetilde{\beta}_T^+ := \widetilde{\lambda}_T + \widetilde{\eta}_T$ and $\widetilde{\beta}_T^- := \widetilde{\eta}_T$.

To derive the limit distribution of the FM-OLS estimator, we make the following assumptions.

Assumption 2. *(i) $\mathbf{\Sigma}_*$ is finite and PD and $\widetilde{\mathbf{\Sigma}}_T \overset{\mathbb{P}}{\to} \mathbf{\Sigma}_*$; and (ii) $\mathbf{\Pi}_*$ is finite and $\widetilde{\mathbf{\Pi}}_T \overset{\mathbb{P}}{\to} \mathbf{\Pi}_*$, where $\mathbf{\Pi}_* := \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^T \sum_{i=1}^t \mathbb{E}[\boldsymbol{g}_i \boldsymbol{g}_t']$.*

Let $\pi_*^{(i,j)}$ be the (i,j)-th element of $\mathbf{\Pi}_*$. Then, $\pi_*^{(1,2)}$ is identical to υ_*.

Lemma 3. *Under Assumptions 1 and 2 and for $k = 1$, $\widetilde{\mathbf{U}}_T := \widetilde{\mathbf{D}}_T^{-1}\{\sum_{t=1}^T \boldsymbol{q}_t(u_t - \Delta x_t (\widetilde{\sigma}_T^{(1,1)})^{-1} \widetilde{\sigma}_T^{(1,2)}) - T\mathbf{S}'\widetilde{\upsilon}_T\} \Rightarrow \widetilde{\boldsymbol{\mathcal{U}}} := \tau_*[\int d\mathcal{W}_u, \mu_*^+ \int r d\mathcal{W}_u, \int \mathcal{B}_x d\mathcal{W}_u]'$, where $\tau_*^2 :=$ plim$_{T\to\infty}\widetilde{\tau}_T^2$ and $\widetilde{\tau}_T^2 := \widetilde{\sigma}_T^{(2,2)} - \widetilde{\sigma}_T^{(2,1)} (\widetilde{\sigma}_T^{(1,1)})^{-1} \widetilde{\sigma}_T^{(1,2)}$.* □

By Lemma 2(*i*), $\widehat{\mathbf{Q}}_T \Rightarrow \boldsymbol{\mathcal{Q}}$, which is nonsingular with probability 1. The limit distribution of $\widetilde{\boldsymbol{\varrho}}_T$ can be obtained as the product of $\boldsymbol{\mathcal{Q}}^{-1}$ and $\widetilde{\boldsymbol{\mathcal{U}}}$:

Corollary 2. *Under Assumption 1 and for $k = 1$, $\widetilde{\mathbf{D}}_T(\widetilde{\boldsymbol{\varrho}}_T - \boldsymbol{\varrho}_*) \Rightarrow \boldsymbol{\mathcal{Q}}^{-1}\widetilde{\boldsymbol{\mathcal{U}}}$.* □

Corollary 2 has several implications. First, the limit distribution of the FM-OLS estimator is a mixed normal. Conditional on $\sigma\{\mathcal{B}_x(r), r \in (0,1]\}$, the limit distribution of $\widetilde{\mathbf{D}}_T(\widetilde{\boldsymbol{\varrho}}_T - \boldsymbol{\varrho}_*)$ is $N(\mathbf{0}, \tau_*^2 \boldsymbol{\mathcal{Q}}^{-1})$. Consequently, the null limit distribution of a Wald test statistic constructed using the FM-OLS estimator will be chi-squared. Second, we have: $T(\widetilde{\beta}_T^+ - \beta_*^+) = T(\widetilde{\beta}_T^- - \beta_*^-) + o_{\mathbb{P}}(1)$, implying that the limit distribution of $\widetilde{\beta}_T^+$ is equivalent to that of $\widetilde{\beta}_T^-$, where the limit distribution of $\widetilde{\beta}_T^-$ is given by that of $\widetilde{\eta}_T$. Third, the convergence rates of $\widetilde{\beta}_T^+$ and $\widetilde{\beta}_T^-$ are both T, enabling us to estimate the short-run parameters in the second stage by replacing u_{t-1} with $\widetilde{u}_{t-1} := y_{t-1} - \widetilde{\alpha}_T - \widetilde{\beta}_T^+ x_{t-1}^+ - \widetilde{\beta}_T^- x_{t-1}^-$.

Theorem 2 presents the limit distribution of the FM-OLS estimator:

Theorem 2. *Under Assumptions 1 and 2 and for* $k = 1$, $T[(\widetilde{\beta}_T^+ - \beta_*^+), (\widetilde{\beta}_T^- - \beta_*^-)]' \Rightarrow \boldsymbol{\iota}_2 \otimes \mathbf{S}\boldsymbol{\mathcal{Q}}^{-1}\widetilde{\boldsymbol{\mathcal{U}}}$. □

As the long-run coefficients are super-consistent, we can treat them as known when estimating the short-run parameters. Let $u_{t-1} := y_{t-1} - \beta_*^+ x_{t-1}^+ - \beta_*^- x_{t-1}^-$. Then, we rewrite (3) as:

$$\Delta y_t = \rho_* u_{t-1} + \gamma_* + \sum_{j=1}^{p-1} \varphi_{j*} \Delta y_{t-j} + \sum_{j=0}^{q-1} \left(\pi_{j*}^{+\prime} \Delta x_{t-j}^+ + \pi_{j*}^{-\prime} \Delta x_{t-j}^- \right) + e_t,$$

which can be expressed as $\Delta y_t = \boldsymbol{\zeta}_*' \boldsymbol{h}_t + e_t$, where $\boldsymbol{\zeta}_* := (\rho_*, \boldsymbol{\beta}_{2*}')'$, $\boldsymbol{\beta}_{2*} := (\gamma_*, \varphi_{1*}, \ldots, \varphi_{p-1*}, \pi_{0*}^{+\prime}, \ldots, \pi_{q-1*}^{+\prime}, \pi_{0*}^{-\prime}, \ldots, \pi_{q-1*}^{-\prime},)'$ and $\boldsymbol{h}_t := (u_{t-1}, \boldsymbol{z}_{2t}')'$. Because all variables in this equation are stationary, the dynamic parameters can be estimated by the OLS estimator: $\widehat{\boldsymbol{\zeta}}_T := (\sum_{t=1}^T \boldsymbol{h}_t \boldsymbol{h}_t')^{-1} \sum_{t=1}^T \boldsymbol{h}_t \Delta y_t = \boldsymbol{\zeta}_* + (\sum_{t=1}^T \boldsymbol{h}_t \boldsymbol{h}_t')^{-1} \sum_{t=1}^T \boldsymbol{h}_t e_t$.

Lemma 4. *Under Assumption 1, (i)* $\widehat{\boldsymbol{\Gamma}}_T := T^{-1} \sum_{t=1}^T \boldsymbol{h}_t \boldsymbol{h}_t' \stackrel{\mathbb{P}}{\to} \boldsymbol{\Gamma}_* := \mathbb{E}[\boldsymbol{h}_t \boldsymbol{h}_t']$; *(ii)* $T^{-1/2} \sum_{t=1}^T \boldsymbol{h}_t e_t \stackrel{A}{\sim} N[\mathbf{0}, \boldsymbol{\Omega}_*]$, *where* $\boldsymbol{\Omega}_* := \mathbb{E}[e_t^2 \boldsymbol{h}_t \boldsymbol{h}_t']$; *and (iii) in the special case where* $\mathbb{E}[e_t^2 | \boldsymbol{h}_t] = \sigma_*^2$, $\boldsymbol{\Omega}_*$ *simplifies to* $\sigma_*^2 \boldsymbol{\Gamma}_*$. □

Using Lemma 4, we derive the limit distribution of $\widehat{\boldsymbol{\zeta}}_T$ in Theorem 3:

Theorem 3. *Under Assumption 1, and if* $\boldsymbol{\Gamma}_*$ *and* $\boldsymbol{\Omega}_*$ *are PD, then (i)* $\sqrt{T}(\widehat{\boldsymbol{\zeta}}_T - \boldsymbol{\zeta}_*) \stackrel{A}{\sim} N(\mathbf{0}, \boldsymbol{\Gamma}_*^{-1} \boldsymbol{\Omega}_* \boldsymbol{\Gamma}_*^{-1})$; *and (ii) further if* $\mathbb{E}[e_t^2 | \boldsymbol{h}_t] = \sigma_*^2$, *then* $\sqrt{T}(\widehat{\boldsymbol{\zeta}}_T - \boldsymbol{\zeta}_*) \stackrel{A}{\sim} N(\mathbf{0}, \sigma_*^2 \boldsymbol{\Gamma}_*^{-1})$. □

4 Hypothesis Testing

We develop the testing procedure for the presence of asymmetries in both the long-run and short-run. Consider $H_0' : (\beta_*^+ - \beta_*^-) = r$ vs. $H_1' : (\beta_*^+ - \beta_*^-) \neq r$ for some $r \in \mathbb{R}$. By setting $r = 0$, we can test whether $\beta_*^+ = \beta_*^-$. Recalling that $\lambda_* := \beta_*^+ - \beta_*^-$, we express these hypotheses as $H_0'' : \lambda_* = r$ vs. $H_1'' : \lambda_* \neq r$. It is straightforward to test this restriction if λ_* is estimated by the FM-OLS estimator, which is asymptotically mixed-normal. The corresponding Wald statistic follows a chi-squared distribution under the null asymptotically. This is a significant advantage of the FM-OLS estimator over the OLS estimator.

Corollary 2 provides the limit distribution of $\widetilde{\lambda}_T$. Letting $\mathbf{S}_\ell := [0, 1, 0]$, then $T^{3/2}(\widetilde{\lambda}_T - \lambda_*) = \mathbf{S}_\ell \widetilde{\mathbf{D}}_T(\widetilde{\boldsymbol{\varrho}}_T - \boldsymbol{\varrho}_*) \Rightarrow \mathbf{S}_\ell \boldsymbol{\mathcal{Q}}^{-1}\widetilde{\boldsymbol{\mathcal{U}}}$, implying that $T^{3/2}(\widetilde{\lambda}_T - r) \Rightarrow \mathbf{S}_\ell \boldsymbol{\mathcal{Q}}^{-1}\widetilde{\boldsymbol{\mathcal{U}}}$ under H_0''. The Wald test statistic is constructed as $\mathcal{W}_T^{(\ell)} := T^3(\widetilde{\lambda}_T - r)^2(\widetilde{\tau}_T^2 \mathbf{S}_\ell \widehat{\mathbf{Q}}_T^{-1}\mathbf{S}_\ell')^{-1}$. Notice, however, that this Wald statistic may be inappropriate to test other forms of hypothesis. For example, consider $H_0''' : \mathbf{R}\boldsymbol{\beta}_* = \mathbf{r}$ vs. $H_1''' : \mathbf{R}\boldsymbol{\beta}_* \neq \mathbf{r}$ for some $\mathbf{R} \in \mathbb{R}^{r\times 2}$ and $\mathbf{r} \in \mathbb{R}^r$ ($r \in \{1, 2\}$), where $\boldsymbol{\beta}_* := (\beta_*^+, \beta_*^-)'$. Let:

$$\widetilde{\mathbf{R}}_\ell := \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

These hypotheses can be rewritten as $H_0''' : \widetilde{\mathbf{R}}\boldsymbol{\varrho}_* = \mathbf{r}$ vs. $H_1''' : \widetilde{\mathbf{R}}\boldsymbol{\varrho}_* \neq \mathbf{r}$, where $\widetilde{\mathbf{R}}\boldsymbol{\varrho}_* = \boldsymbol{\beta}_*$ and $\widetilde{\mathbf{R}} := \mathbf{R}\widetilde{\mathbf{R}}_\ell$. Then, we construct the Wald test statistic as

$$\widetilde{\mathcal{W}}_T^{(\ell)} := (\widetilde{\mathbf{R}}\widetilde{\boldsymbol{\varrho}}_T - \mathbf{r})'(\widetilde{\tau}_T^2 \widetilde{\mathbf{R}}\mathbf{Q}_T^{-1}\widetilde{\mathbf{R}}')^{-1}(\widetilde{\mathbf{R}}\widetilde{\boldsymbol{\varrho}}_T - \mathbf{r}),$$

where $\mathbf{Q}_T := \sum_{t=1}^T \boldsymbol{q}_t \boldsymbol{q}_t'$.

Theorem 4. *Under Assumptions 1 and 2, $\mathcal{W}_T^{(\ell)} \overset{A}{\sim} \mathcal{X}_1^2$ under H_0'' and $\widetilde{\mathcal{W}}_T^{(\ell)} \overset{A}{\sim} \mathcal{X}_2^2$ under H_0'''. For any sequence, c_T and $\widetilde{c}_T$ such that $c_T = o(T^3)$ and $\widetilde{c}_T = o(T^2)$, $\mathbb{P}(\mathcal{W}_T^{(\ell)} > c_T) \to 1$ under H_1'' and $\mathbb{P}(\widetilde{\mathcal{W}}_T^{(\ell)} > \widetilde{c}_T) \to 1$ under H_1'''.* □

In the NARDL literature, it is common to test for additive symmetry of the short-run dynamic parameters.[9] Consider $H_0 : \mathbf{R}_s\boldsymbol{\zeta}_* = \mathbf{r}$ vs. $H_1 : \mathbf{R}_s\boldsymbol{\zeta}_* \neq \mathbf{r}$, where $\mathbf{R}_s \in \mathbb{R}^{r\times(1+p+2q)}$ and $\mathbf{r} \in \mathbb{R}^r$ ($r \in \mathbb{N}$) are selection matrices. Define $\mathbf{R}_s := \left[\mathbf{0}_{1+p}', \boldsymbol{\iota}_q', -\boldsymbol{\iota}_q'\right]$ and $\mathbf{r} = 0$. The null hypothesis of additive short-run symmetry can be tested against the alternative hypothesis of asymmetry: $H_0 : \sum_{j=0}^{q-1} \pi_{j*}^+ = \sum_{j=0}^{q-1} \pi_{j*}^-$ vs. $H_1 : \sum_{j=0}^{q-1} \pi_{j*}^+ \neq \sum_{j=0}^{q-1} \pi_{j*}^-$. The Wald test statistic is constructed as

$$\mathcal{W}_T^{(s)} := T(\mathbf{R}_s\widehat{\boldsymbol{\zeta}}_T - \mathbf{r})'(\mathbf{R}_s\widehat{\boldsymbol{\Gamma}}_T^{-1}\widehat{\boldsymbol{\Omega}}_T\widehat{\boldsymbol{\Gamma}}_T^{-1}\mathbf{R}_s')^{-1}(\mathbf{R}_s\widehat{\boldsymbol{\zeta}}_T - \mathbf{r}),$$

[9] Several forms of short-run symmetry restrictions have been considered, including pairwise symmetry of $\boldsymbol{\pi}_{j*}^+$ and $\boldsymbol{\pi}_{j*}^-$ for $j = 0, \ldots, q-1$ (SYG) and impact symmetry defined by $\boldsymbol{\pi}_{0*}^+$ and $\boldsymbol{\pi}_{0*}^-$ (Greenwood-Nimmo and Shin, 2013). It is straightforward to test these alternative restrictions by appropriately specifying selection matrices, $\mathbf{R}_s$ and $\mathbf{r}$.

where $\widehat{\mathbf{\Omega}}_T := T^{-1}\sum_{t=1}^T \widehat{e}_t^2 \boldsymbol{h}_t\boldsymbol{h}_t'$ is a consistent estimator of $\mathbf{\Omega}_*$. Let $\widehat{e}_t := \Delta y_t - \widehat{\boldsymbol{\zeta}}_T'\boldsymbol{h}_t$, where $\widehat{\boldsymbol{\zeta}}_T$ is constructed from the first-step regression using FM-OLS. Further, if the condition in Lemma 4(*iii*) holds, then the Wald statistic simplifies to

$$\mathcal{W}_T^{(s)} := T(\mathbf{R}_s\widehat{\boldsymbol{\zeta}}_T - \mathbf{r})'(\widehat{\sigma}_{e,T}^2\mathbf{R}_s\widehat{\mathbf{\Gamma}}_T^{-1}\mathbf{R}_s')^{-1}(\mathbf{R}_s\widehat{\boldsymbol{\zeta}}_T - \mathbf{r}),$$

where $\widehat{\sigma}_{e,T}^2 := T^{-1}\sum_{t=1}^T \widehat{e}_t^2$.

Theorem 5. *Given Assumption 1, if $\mathbf{\Gamma}_*$ and $\mathbf{\Omega}_*$ are PD, then $\mathcal{W}_T^{(s)} \overset{A}{\sim} \mathcal{X}_r^2$ under H_0 and, for any sequence c_T such that $c_T = o(T)$, $\mathbb{P}(\mathcal{W}_T^{(s)} > c_T) \to 1$ under H_1.* □

5 Monte Carlo Simulations

We conduct stochastic simulations to examine the finite sample properties of the Wald tests described in Section 4.

Consider the NARDL(1,0) data generating process (DGP): $\Delta y_t = \gamma_* + \rho_* u_{t-1} + \varphi_*\Delta y_{t-1} + \pi_*^+\Delta x_t^+ + \pi_*^-\Delta x_t^- + e_t$, where $u_{t-1} := y_{t-1} - \alpha_* - \beta_*^+ x_{t-1}^+ - \beta_*^- x_{t-1}^-$, $\Delta x_t := \kappa_*\Delta x_{t-1} + \sqrt{1-\kappa_*^2}v_t$, and $(e_t, v_t)' \sim \text{IID}N(\mathbf{0}_2, \mathbf{I}_2)$. We focus on the case where the FM-OLS estimator is used in the first step and set $(\alpha_*, \beta_*^+, \beta_*^-, \gamma_*, \rho_*, \varphi_*, \pi_*^+, \pi_*^-, \kappa_*) = (0, 2, 1, 0, -2/3, \varphi_*, 1, 1/2, 1/2)$. We test $H_0^{(\ell)} : \beta_*^+ - \beta_*^- = 0$ vs. $H_1^{(\ell)} : \beta_*^+ - \beta_*^- \neq 0$ by allowing φ_* to vary over $-1/2$, $-1/4$, 0, $1/4$ and $1/2$. The simulation results reported in Table 1 reveal some size distortion in small samples for negative values of φ_*. Nonetheless, as T increases, the distribution of the Wald test statistic becomes well-approximated by the chi-squared distribution with one degree of freedom.

— Insert Table 1 Here —

To examine the power of the Wald test, we generate data with $(\alpha_*, \beta_*^+, \beta_*^-, \gamma_*, \rho_*, \varphi_*, \pi_*^+, \pi_*^-, \kappa_*) = (0, 1.01, 1, 0, -2/3, \varphi_*, 1/3, 1/2, 1/2)$ and allow φ_* to vary as above. The simulation results for $\mathcal{W}_T^{(\ell)}$ reported in Table 2 confirm that the Wald test statistic is consistent under the alternative hypothesis. Furthermore, its power patterns are largely insensitive to the value of φ_*.

— Insert Table 2 Here —

To examine the empirical levels of the Wald test statistic, we set $(\alpha_*, \beta_*^+, \beta_*^-, \gamma_*, \rho_*, \varphi_*, \pi_*^+, \pi_*^-, \kappa_*) = (0, 2, 1, 0, -2/3, \varphi_*, 1/2, 1/2, 1/2)$ and allow φ_* to vary as above. We first estimate the long-run parameters by FM-OLS and compute $\widehat{u}_t$ before estimating the short-run parameters by OLS. We then test $H_0^{(s)} : \pi_*^+ - \pi_*^- = 0$ vs. $H_1^{(s)} : \pi_*^+ - \pi_*^- \neq 0$ using $\mathcal{W}_T^{(s)}$ with the heteroskedasticity consistent covariance estimator, $\widehat{\mathbf{\Omega}}_T$. The simulation results reported in Table 3 reveal that the finite sample distribution of the Wald test statistic is well-approximated by the chi-squared distribution. For each level of significance, its empirical level tends to the nominal level once the number of observations reaches 500. Furthermore, the empirical sizes display little sensitivity to the value of φ_* even for small T.

— Insert Table 3 Here —

Next, we examine the empirical powers of the Wald test statistic. We maintain the same hypotheses but update the parameters to $(\alpha_*, \beta_*^+, \beta_*^-, \gamma_*, \rho_*, \varphi_*, \pi_*^+, \pi_*^-, \kappa_*) = (0, 2, 1, 0, -2/3, \varphi_*, 1, 1/2, 1/2)$. We compute $\mathcal{W}_T^{(s)}$ using the heteroskedasticity consistent covariance matrix estimator. Two points are noteworthy from the simulation results reported in Table 4. First, the empirical power of the Wald test rises with T, indicating that the test is consistent. Second, the power of the Wald test exhibits little sensitivity to the degree of autocorrelation captured by the value of φ_*.

— Insert Table 4 Here —

6 Empirical Application

We examine the asymmetric relationship between R&D intensity and physical investment using quarterly U.S. data covering the period from 1960q1 to 2019q4, which we collect from the Federal Reserve Economic Data service at the Federal Reserve Bank of St. Louis. R&D intensity denoted r_t, is measured as 100 times the ratio of seasonally adjusted nominal R&D expenditure to seasonally adjusted nominal GDP. Investment, denoted i_t, is the log of seasonally adjusted real gross private domestic investment in 2012 prices (GPDI). The National Income and Product Accounts (NIPAs) separately aggregate R&D expenditure and GPDI, so there is no double-counting of R&D expenditure.[10]

[10]The Bureau of Economic Analysis constructs a partial R&D satellite account and revises the NIPAs by treating R&D expenditure as part of investment (see Fraumeni and Okubo, 2005).

The Phillips and Perron (1988) unit root test results indicate that both r_t and i_t are nonstationary. Hence, we report descriptive statistics for the first differences of both series, see the Online Supplement. While the growth rate of R&D intensity is well approximated by a normal distribution centered at zero, the GPDI growth rate exhibits a non-zero mean with negative skewness and excess kurtosis. Furthermore, we find that i_t is a unit-root process with a time drift, but r_t is a driftless unit-root process.[11] As described above, the NARDL model can capture an asymmetric cointegrating relationship between i_t and r_t without the need to include a deterministic time trend.

We estimate the NARDL model using the two-step procedure proposed in Section 3, with the FM-OLS estimator used in the first step. We select the lag orders using the Akaike Information Criterion (AIC), which results in the following NARDL(2,2) error-correction specification:

$$\Delta i_t = \gamma_* + \rho_* u_{t-1} + \varphi_* \Delta i_{t-1} + \pi_{0*}^+ \Delta r_t^+ + \pi_{0*}^- \Delta r_t^- + \pi_{1*}^+ \Delta r_{t-1}^+ + \pi_{1*}^- \Delta r_{t-1}^- + e_t,$$

where $i_t = \beta_*^+ r_t^+ + \beta_*^- r_t^- + u_t$. When applying the Phillips and Perron (1988) unit root test to the residuals, we reject the null hypothesis of no cointegration with a p-value of 0.01, concluding that there exists an asymmetric long-run relationship between i_t, r_t^+ and r_t^-.

— Insert Table 5 Here —

The long-run parameter estimates, reported in Table 5, are not only different in magnitude, but they also display opposite signs: $\tilde{\beta}_*^+$ is positive and $\tilde{\beta}_*^-$ is negative, with both being highly significant. Furthermore, the null hypothesis of long-run symmetry is strongly rejected.[12] Our results indicate that an increase in R&D spending equivalent to 1% of GDP is associated with an increase in real investment of 2.7% in the long run when R&D growth exceeds GDP growth (i.e. when innovative R&D is prevalent). On the other hand, an increase in R&D spending of 1% of GDP reduces real investment by 6.4% when R&D growth is slower than GDP growth (i.e. when managerial R&D is prevalent).[13] This is consistent with our theoretical prediction regarding the asymmetric relationship

[11] We observe that the time trend coefficient is significantly different from zero for i_t, but not for r_t.

[12] Due to the reparamerization applied to the FM-OLS estimator, the Wald test of $H_0 : \beta_*^+ = \beta_*^-$ versus $H_1 : \beta_*^+ \neq \beta_*^-$ is equivalent to a t-test of $H_0 : \lambda_* = 0$ versus $H_1 : \lambda_* \neq 0$, which returns a p-value almost identical to zero.

[13] The average of U.S. R&D intensity over our sample period is 2.69% of GDP with a standard deviation of 0.2%. Hence, a shock equivalent to 1% of GDP is large by historical standards.

between R&D intensity and investment. We expect $\beta_*^+ > 0$ due to the complementarity of innovative R&D expenditure and investment, while $\beta_*^- < 0$ is consistent with the nature of managerial R&D expenditure and investment as substitutes. This suggests that the substitution effect may be much stronger than the complementary effect in the long run.

Next, from the short-run estimation results reported in Table 6, we find no evidence of residual autocorrelation up to order 4, with the p-value of the Breusch-Godfrey Lagrange multiplier (LM) test statistic being 0.39. However, the Breusch-Pagan LM test rejects the conditional homoskedasticity hypothesis with a p-value of 0.035. Thus, we report the robust standard errors obtained using the HAC covariance matrix estimator.

— Insert Table 6 Here —

We observe that disequilibrium errors are corrected at a statistically significant rate of 6.8% per quarter, but we find no evidence of short-run asymmetry. The Wald test of the null hypothesis of impact symmetry, $H_0 : \pi_{0*}^+ = \pi_{0*}^-$, versus the alternative, $H_1 : \pi_{0*}^+ \neq \pi_{0*}^-$, returns a p-value of 0.677. Moreover, the null hypothesis of additive short-run symmetry, $H_0 : \pi_{0*}^+ + \pi_{1*}^+ = \pi_{0*}^- + \pi_{1*}^-$, is not rejected against the alternative, $H_1 : \pi_{0*}^+ + \pi_{1*}^+ \neq \pi_{0*}^- + \pi_{1*}^-$, with a p-value of 0.146.

In panels (a) and (b) of Figure 3, we report the cumulative dynamic multipliers associated with an increase in R&D intensity equal to 1% of GDP in each regime (i.e. when R&D grows faster than GDP and where it grows slower than GDP). In panel (c), we display the difference between these two dynamic multipliers as a measure of asymmetry at each horizon with an empirical 95% confidence interval obtained from 5,000 iterations of a moving block bootstrap procedure with a block length of $T^{1/3}$.

— Insert Figure 3 Here —

In Figure 3(a), when R&D growth exceeds GDP growth (i.e. when innovative R&D predominates), we find that a 1% increase in R&D intensity initially reduces real investment with a peak reduction of 6.7% after one quarter. From the perspective of creative destruction, this initial reduction may reflect that large innovative R&D expenditures focus on new product development, creating a degree of obsolescence in existing technologies and promoting an incentive to reduce investment in those technologies. Then, the dynamic multiplier rises steadily and becomes positive after 8 quarters, as newly developed technologies begin to mature and scale. In the long run, the impact of innovative R&D expenditure on real physical investment is positive, at 2.7%.

Figure 3(b) shows that when managerial R&D predominates, a 1% increase in R&D intensity leads to an immediate reduction in the real investment of -3.6%, reflecting short-term substitution. The dynamic multiplier effect is then indistinguishable from zero until horizon 11, after which it converges to a long-run value of -6.4%. Overall, an economic environment that favors managerial R&D expenditure is conducive to decreased real investment, particularly in the long run, as the efficiency gains derived from managerial R&D raise the return on each dollar invested.

In Figure 3(c), we characterize the asymmetry between the cumulative dynamic multipliers associated with innovative and managerial R&D expenditures. Specifically, we construct the difference between the cumulative dynamic multiplier reported in panel (a) and in panel (b) with 95% bootstrap intervals, that can be used to test for asymmetry at any horizon. In the short run, we observe substantial negative asymmetry that is significant at the 10% level. This suggests that real investment is more responsive to innovative than managerial R&D expenditures in the short run. However, in the long run, we observe positive asymmetry as the substitution effect associated with managerial R&D expenditure is stronger than the complementary effect of innovative R&D expenditure. This pattern is generally in line with the product life cycle, where large innovative R&D expenditure occurs early often exerting a disruptive influence on existing products. Once the new product becomes established, managerial R&D expenditure focuses on scaling-up production and delivering efficiency gains. Collectively, these results have important implications for the endogenous growth literature, where linear functional forms are routinely imposed to characterize the relationship between the stock of knowledge and R&D intensity (e.g. Romer, 1990). Our main findings suggest that the imposition of a linear functional form may produce misleading results.

In the Online Supplement, we present the additional estimation results obtained using the single-step OLS estimation procedure popularized by SYG, which are summarized in Figure A.1. A comparison of Figures 3 and A.1 confirms that both results are very similar.

— Insert Figure A.1 Here —

7 Concluding Remarks

In this paper, we analyze the potentially asymmetric relationship between R&D intensity and physical investment using the quarterly time series data in the U.S. over the period 1960q1–2019q4.

We have developed a theoretical model that exploits documented differences in innovative and managerial R&D characteristics to establish that there exists an asymmetric relationship between physical investment and R&D expenditures. We also develop the testable hypothesis that innovative (managerial) R&D expenditure is a complement to (a substitute for) physical investment.

We have proposed a corresponding empirical specification that takes the form of a NARDL model in which real investment is regressed on the positive and negative partial sums of aggregate R&D intensity. While the NARDL model has been widely applied to the analysis of asymmetric relationships (see Cho et al., 2023b), the asymptotic theory for the single-step NARDL estimator has yet to be fully developed due to an asymptotic singularity issues that frustrate efforts to derive its limit distribution. We develop a tractable two-step estimation procedure. In the first step, we estimate the parameters of a transformed long-run relationship using the FM-OLS estimator that accounts for serial correlation and potential endogeneity of the regressors and that facilitates standard inference. In the second step, the short-run dynamic coefficients can be consistently estimated by OLS, treating the error-correction term from the first step as given. We derive the asymptotic distributions of the estimators and develop standard Wald tests for inference on the short- and long-run parameters. A suite of Monte Carlo simulations indicates that our asymptotic results offer good approximations in finite samples.

Employing the U.S. quarterly data covering the period from 1960q1 to 2019q4, we find comprehensive empirical evidence in favor of the asymmetric relationship between R&D intensity and investment. In the long-run, investment responds positively to R&D expenditures when their growth rate exceeds the growth rate of GDP and negatively when they grow more slowly than GDP. This supports theoretical predictions that the innovative (managerial) R&D expenditure is a complement to (substitute for) physical investment. Furthermore, we find that physical investment is more sensitive to changes in R&D intensity when the growth rate of R&D expenditure is lower than GDP growth.

Our work opens several avenues for continuing research. On the methodological side, there is scope to generalize our approach to accommodate trending regressors or to estimate an unknown threshold parameter in the construction of the partial sum processes. On the empirical side, our results motivate the study of potential asymmetries in other areas of the innovation literature, such as in the knowledge production function that is routinely estimated in linear form in the endogenous growth literature.

References

AGARWAL, R. AND D. B. AUDRETSCH (2001): "Does Entry Size Matter? The Impact of the Life Cycle and Technology on Firm Survival," *Journal of Industrial Economics*, 49, 21–43.

AGHION, P., R. BLUNDELL, R. GRIFFTH, P. HOWITT, AND S. PRANTL (2009): "The Effects of Entry on Incumbent Innovation and Productivity," *Review of Economics and Statistics*, 91, 20–32.

AUDRETSCH, D. B. (1987): "An Empirical Test of the Industry Life Cycle," *Weltwirtschaftliches Archiv*, 123, 297–308.

BARATI, M. AND H. FARIDITAVANA (2020): "Asymmetric Effect of Income on the US Healthcare Expenditure: Evidence from the Nonlinear Autoregressive Distributed Lag (ARDL) Approach," *Empirical Economics*, 58, 1979–2008.

BAUSSOLA, M. (2000): "The Causality Between R&D and Investment," *Economics of Innovation and New Technology*, 9, 385–399.

BOX, M., K. GRATZER, AND X. LIN (2019): "The Asymmetric Effect of Bankruptcy Fraud in Sweden: A Long-Term Perspective," *Journal of Quantitative Criminology*, 35, 287–312.

BRUN-AGUERRE, R. X., A.-M. FUERTES, AND M. J. GREENWOOD-NIMMO (2017): "Heads I Win; Tails You Lose: Asymmetry in Exchange Rate Pass-Through into Import Prices," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 180, 587–612.

CHIAO, C. (2001): "The Relationship between R&D and Physical Investment of Firms in Science-Based Industries," *Applied Economics*, 33, 23–35.

CHO, J., M. GREENWOOD-NIMMO, AND Y. SHIN (2023a): "Nonlinear Autoregressive Distributed Lag Model Estimation Using Data with Drifts," Working paper, School of Economics, Yonsei University, Seoul.

——— (2023b): "Recent Developments of the Autoregressive Distributed Lag Modelling Framework," *Journal of Economic Surveys*, 37, 7–32.

CHUNG, D. AND D. SHIN (2020): "When do Firms Invest in R&D? Two Types of Performance Feedback and Organizational Search in the Korean Shipbuilding Industry," *Asian Business & Management*, 20, 583–617.

COMIN, D. AND T. PHILIPPON (2005): "The Rise in Firm-level Volatility: Causes and Consequences," *NBER/Macroeconomics Annual*, 20, 167–201.

EBERHARDT, M. AND A. F. PRESBITERO (2015): "Public Debt and Growth: Heterogeneity and Non-Linearity," *Journal of International Economics*, 97, 45–58.

ENGLE, R. F. AND C. W. GRANGER (1987): "Co-integration and Error Correction: Representation, Estimation and Testing," *Econometrica*, 55, 251–276.

FERRIS, J. S., S. L. WINER, AND D. E. OLMSTEAD (2020): "A Dynamic Model of Political Party Equilibrium: The Evolution of ENP in Canada, 1870 – 2015," Working Paper 8387, CESifo, Munich.

FRAUMENI, B. M. AND S. OKUBO (2005): "R&D in the National Income and Product Accounts: A First Look at Its Effect on GDP," in *Measuring Capital in the New Economy*, ed. by J. H. C. Corrado and D. Sichel, Chicago (IL): The University of Chicago Press, 275–322.

FUDENBERG, D., R. GILBERT, J. STIGLITZ, AND J. TIROLE (1983): "Preemption, Leapfrogging and Competition in Patent Races," *European Economic Review*, 22, 3–31.

GORT, M. AND R. A. WALL (1986): "The Evolution of Technologies and Investment in Innovations," *Economic Journal*, 96, 741–757.

GREENWOOD-NIMMO, M. J. AND Y. SHIN (2013): "Taxation and the Asymmetric Adjustment of Selected Retail Energy Prices in the UK," *Economics Letters*, 121, 411–416.

GROSSMAN, G. M. AND C. SHAPIRO (1987): "Dynamic R&D Competition," *Economic Journal*, 97, 372–387.

HAMMOUDEH, S., A. LAHIANI, D. K. NGUYEN, AND R. M. SOUSA (2015): "An Empirical Analysis of Energy Cost Pass-Through to CO2 Emission Prices," *Energy Economics*, 49, 149–156.

HARRIS, C. J. AND J. S. VICKERS (1987): "Racing with Uncertainty," *Review of Economic Studies*, 54, 1–21.

HE, Z. AND F. ZHOU (2018): "Time-Varying and Asymmetric Eeffects of the Oil-Specific Demand Shock on Investor Sentiment," *PLOS ONE*, 13, e0200734.

KAMIEN, M. I. AND N. L. SCHWARTZ (1972): "Timing of Innovations under Rivalry," *Econometrica*, 40, 43–60.

KLEPPER, S. (1996): "Entry, Exit, Growth, and Innovation over the Product Life Cycle," *American Economic Review*, 86, 562–583.

——— (1997): "Industry Life Cycles," *Industrial and Corporate Change*, 6, 145–182.

LACH, S. AND R. ROB (1996): "R&D Investment and Industry Dynamics," *Journal of Economics and Management Strategy*, 5, 217–249.

LACH, S. AND M. SCHANKERMAN (1989): "Dynamics of R&D and Physical Investment in the Scientific Sector," *Journal of Political Economy*, 97, 880–904.

NEWEY, W. K. AND K. D. WEST (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

PESARAN, M. H. AND Y. SHIN (1998): "An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis," in *Econometrics and Economic Theory: The Ragnar Frisch Centennial Symposium*, ed. by S. Strom, Cambridge: Cambridge University Press, Econometric Society Monographs, 371–413.

PESARAN, M. H., Y. SHIN, AND R. J. SMITH (2001): "Bounds Testing Approaches to the Analysis of Level Relationships," *Journal of Applied Econometrics*, 16, 289–326.

PHILLIPS, P. AND P. PERRON (1988): "Testing for a Unit Root in Time Series Regression," *Biometrika*, 75, 335–346.

PHILLIPS, P. C. AND B. E. HANSEN (1990): "Statistical Inference in Instrumental Variable Regression with I(1) Processes," *Review of Economic Studies*, 57, 99–125.

REINGANUM, J. F. (1982): "A Dynamic Game of R&D: Patent Protection and Competitive Behavior," *Econometrica*, 50, 671–688.

ROMER, P. M. (1990): "Endogenous Technological Change," *Journal of Political Economy*, 98, 71–102.

SCHMOOKLER, J. (1966): *Innovation and Economic Growth*, Cambridge (MA): Harvard University Press.

SCHUMPETER, J. A. (1942): *Capitalism, Socialism, and Democracy*, New York (NY): Harper & Row.

SHIN, Y., B. YU, AND M. J. GREENWOOD-NIMMO (2014): "Modelling Asymmetric Cointegration and Dynamic Multipliers in a Nonlinear ARDL Framework," in *Festschrift in Honor of Peter Schmidt: Econometric Methods and Applications*, ed. by W. Horrace and R. Sickles, New York (NY): Springer Science & Business Media, 281–314.

SÜSSMUTH, B. AND U. WOITEK (2013): "Estimating Dynamic Asymmetries in Demand at the Munich Oktoberfest," *Tourism Economics*, 19, 653–674.

UTTERBACK, J. AND W. Y. ABERNATHY (1975): "A Dynamic Model of Process and Product Innovation," *Omega*, 3, 639–659.

VERHEYEN, F. (2013): "Exchange Rate Nonlinearities in EMU Exports to the US," *Economic Modelling*, 32, 66–76.

ZIF, J. AND D. J. MCCARTHY (1997): "The R&D Cycle: The Influence of Product and Process R&D on Short-Term ROI," *IEEE Transactions on Engineering Management*, 44, 114–123.

φ_*	sample size	100	250	500	750	1,000
-0.50	1%	12.40	5.06	2.38	1.82	1.44
	5%	23.34	12.88	7.96	7.70	5.90
	10%	31.50	21.04	14.08	13.22	11.46
-0.25	1%	8.74	3.80	2.54	1.74	1.18
	5%	19.06	11.10	8.44	6.62	5.48
	10%	27.22	18.36	14.96	11.68	10.48
0.00	1%	4.96	2.92	1.84	1.72	1.42
	5%	13.86	9.76	7.20	6.60	6.12
	10%	21.40	16.28	13.02	12.20	11.34
0.25	1%	3.32	1.62	1.34	1.22	1.06
	5%	10.38	6.24	5.56	5.66	5.60
	10%	17.28	11.42	10.96	11.22	10.70
0.50	1%	1.70	0.86	0.78	0.66	0.72
	5%	5.82	4.06	4.60	4.30	4.22
	10%	10.74	8.20	9.88	9.08	9.12

Table 1: EMPIRICAL LEVELS OF THE WALD TEST FOR LONG-RUN SYMMETRY. The FM-OLS is used in the first step. The data is generated as follows: $\Delta y_t = -(2/3)u_{t-1} + \varphi_* \Delta y_{t-1} + (1/3)\Delta x_t^+ + (1/2)\Delta x_t^- + e_t$, where $u_t := y_t - x_t^+ - x_t^-$, $\Delta x_t = 0.5\Delta x_{t-1} + \sqrt{1-0.5^2}v_t$, and $(e_t, v_t)' \sim \text{IID}N(\mathbf{0}_2, \mathbf{I}_2)$. $H_0^{(\ell)} : \beta_*^+ - \beta_*^- = 0$ vs. $H_1^{(\ell)} : \beta_*^+ - \beta_*^- \neq 0$. The simulation results are obtained using $R = 5,000$ replications.

φ_*	sample size	100	250	500	750	1,000
-0.50	1%	9.80	20.76	83.66	97.34	99.76
	5%	20.00	35.78	89.36	98.54	99.96
	10%	27.66	44.70	91.98	98.92	99.96
-0.25	1%	7.90	26.04	88.16	98.44	99.82
	5%	17.98	41.74	92.88	99.32	99.92
	10%	25.08	51.30	94.58	99.60	99.94
0.00	1%	5.74	32.24	91.22	99.20	99.86
	5%	14.72	49.46	95.12	99.68	99.96
	10%	22.40	58.54	96.66	99.80	99.98
0.25	1%	4.4	34.46	92.36	99.38	99.96
	5%	12.08	52.82	96.04	99.70	100.0
	10%	19.2	61.96	97.22	99.82	100.0
0.50	1%	2.92	25.40	90.96	99.16	99.98
	5%	9.44	47.06	95.40	99.70	100.0
	10%	16.08	58.68	96.98	99.82	100.0

Table 2: EMPIRICAL POWER OF THE WALD TEST FOR LONG-RUN SYMMETRY (IN PERCENT). The FM-OLS is used in the first step. The data is generated as follows: $\Delta y_t = -(2/3)u_{t-1} + \varphi_* \Delta y_{t-1} + (1/3)\Delta x_t^+ + (1/2)\Delta x_t^- + e_t$, where $u_t := y_t - 1.01x_t^+ - x_t^-$, $\Delta x_t = 0.5\Delta x_{t-1} + \sqrt{1-0.5^2}v_t$, and $(e_t, v_t)' \sim \text{IID}N(\mathbf{0}_2, \mathbf{I}_2)$. $H_0^{(\ell)} : \beta_*^+ - \beta_*^- = 0$ vs. $H_1^{(\ell)} : \beta_*^+ - \beta_*^- \neq 0$. The simulation results are obtained using $R = 5,000$ replications.

φ_*	sample size	100	250	500	750	1,000
-0.50	1%	2.44	1.60	0.98	1.18	1.12
	5%	8.06	6.06	5.42	5.46	6.02
	10%	13.82	11.00	10.90	10.36	10.94
-0.25	1%	2.38	1.74	1.46	1.30	1.14
	5%	7.38	6.44	6.02	5.36	5.30
	10%	12.90	11.38	11.28	10.54	10.48
0.00	1%	2.12	1.18	1.22	1.26	0.98
	5%	7.30	5.86	5.76	6.00	5.20
	10%	13.40	11.26	10.94	11.16	10.22
0.25	1%	2.28	1.42	1.36	0.96	0.82
	5%	7.32	6.14	5.84	5.12	4.68
	10%	13.42	11.40	10.96	9.76	9.50
0.50	1%	2.02	1.80	0.98	1.10	1.22
	5%	6.64	6.44	5.44	5.52	5.54
	10%	11.84	11.54	10.62	10.60	10.74

Table 3: EMPIRICAL LEVELS OF THE WALD TEST FOR SHORT-RUN SYMMETRY (IN PERCENT). The FM-OLS is used in the first step and OLS is used in the second step. The data is generated as follows: $\Delta y_t = -(2/3)u_{t-1} + \varphi_* \Delta y_{t-1} + (1/2)\Delta x_t^+ + (1/2)\Delta x_t^- + e_t$, where $u_t := y_t - 2x_t^+ - x_t^-$, $\Delta x_t = 0.5\Delta x_{t-1} + \sqrt{1-0.5^2}v_t$, and $(e_t, v_t)' \sim \text{IID}N(\mathbf{0}_2, \mathbf{I}_2)$. $H_0^{(s)} : \pi_*^+ - \pi_*^- = 0$ vs. $H_1^{(s)} : H_0 : \pi_*^+ - \pi_*^- \neq 0$. The simulation results are obtained using $R = 5,000$ replications.

φ_*	sample size	100	250	500	750	1,000
-0.50	1%	17.86	44.00	78.48	93.50	98.40
	5%	35.38	66.66	91.82	98.44	99.76
	10%	45.56	76.70	95.76	99.34	99.92
-0.25	1%	17.58	44.84	79.40	93.70	98.32
	5%	34.96	66.66	91.90	98.26	99.72
	10%	46.04	76.64	95.96	99.14	99.86
0.00	1%	17.26	43.16	78.56	93.28	98.60
	5%	35.66	66.68	92.38	98.14	99.72
	10%	46.62	76.14	96.06	99.18	99.90
0.25	1%	17.90	43.02	78.76	93.34	98.72
	5%	35.02	66.14	92.12	98.32	99.68
	10%	45.80	76.24	95.48	99.34	99.98
0.50	1%	17.50	42.82	77.82	92.94	98.54
	5%	34.20	65.78	91.32	98.28	99.78
	10%	44.90	76.06	94.90	99.26	99.92

Table 4: EMPIRICAL POWER OF THE WALD TEST FOR SHORT-RUN SYMMETRY (IN PERCENT). The FM-OLS is used in the first step and OLS is used in the second step. The data is generated as follows: $\Delta y_t = -(2/3)u_{t-1} + \varphi_* \Delta y_{t-1} + \Delta x_t^+ + (1/2)\Delta x_t^- + e_t$, where $u_t := y_t - 2x_t^+ - x_t^-$, $\Delta x_t = 0.5\Delta x_{t-1} + \sqrt{1-0.5^2}v_t$, and $(e_t, v_t)' \sim \text{IID}N(\mathbf{0}_2, \mathbf{I}_2)$. $H_0^{(s)} : \pi_*^+ - \pi_*^- = 0$ vs. $H_1^{(s)} : H_0 : \pi_*^+ - \pi_*^- \neq 0$. The simulation results are obtained using $R = 5,000$ replications.

	Estimate	S.E.
Intercept	7.453	0.381
β_*^+	0.271	0.133
β_*^-	-0.640	0.154

Table 5: LONG-RUN PARAMETER ESTIMATES USING QUARTERLY OBSERVATIONS. This table reports the long-run parameter estimates obtained from our two-step estimation procedure applied to observations from 1960q1 to 2019q4, where FM-OLS is used in the first step.

	Estimate	S.E.
γ_*	0.015	0.005
ρ_*	-0.068	0.016
φ_*	0.255	0.066
π_{0*}^+	-0.555	0.179
π_{1*}^+	-0.029	0.150
π_{0*}^-	-0.359	0.344
π_{1*}^-	0.482	0.176
Adjusted R^2	0.199	
$\mathcal{X}^2_{\text{S.Corr.}}$	0.385	
$\mathcal{X}^2_{\text{Hetero.}}$	0.035	

Table 6: DYNAMIC PARAMETER ESTIMATES USING QUARTERLY OBSERVATIONS. This table reports parameter estimates for the NARDL(2,2) model in error-correction form estimated using the two-step procedure applied to observations from 1960q1 to 2019q4, where FM-OLS is used in the first step and OLS is used in the second step. The lag order is selected by AIC, and the standard errors are computed by HAC robust covariance matrix estimation. $\mathcal{X}^2_{\text{S.Corr.}}$ denotes the Breusch–Godfrey Lagrange multiplier test for serial correlation up to order four. $\mathcal{X}^2_{\text{Hetero.}}$ denotes the Breusch–Pagan–Godfrey Lagrange multiplier test for residual heteroskedasticity. The values reported for these two tests are asymptotic p-values.

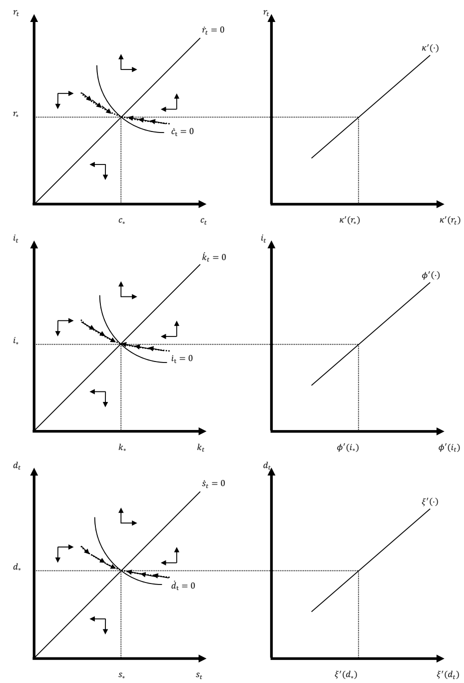


Figure 1: PHASE DIAGRAMS AND STEADY STATE. This figure shows the phase diagrams of (r_t, c_t), (i_t, k_t), and (d_t, k_t) and their relationships with the marginal cost functions.

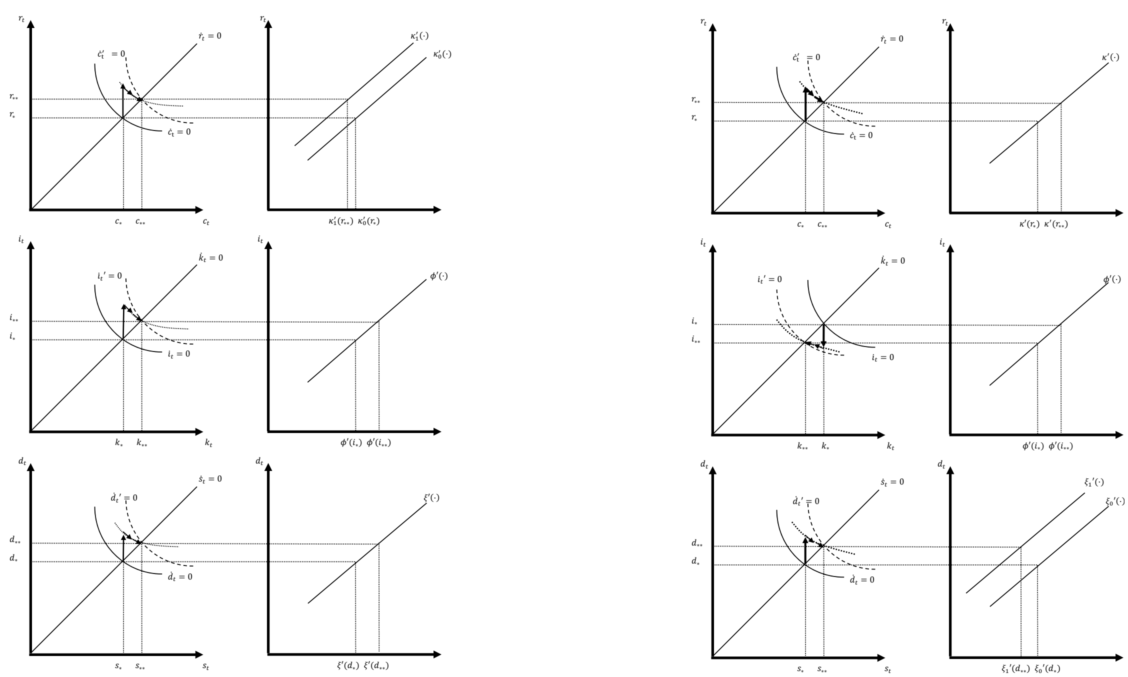
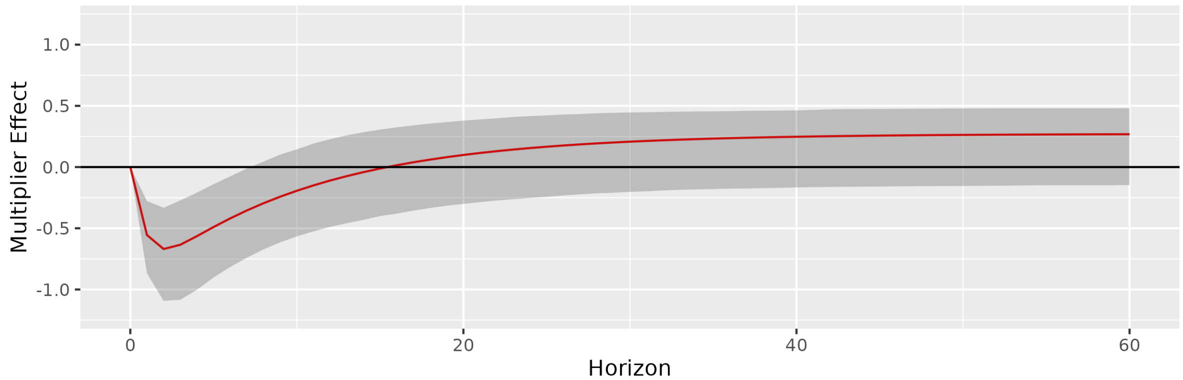
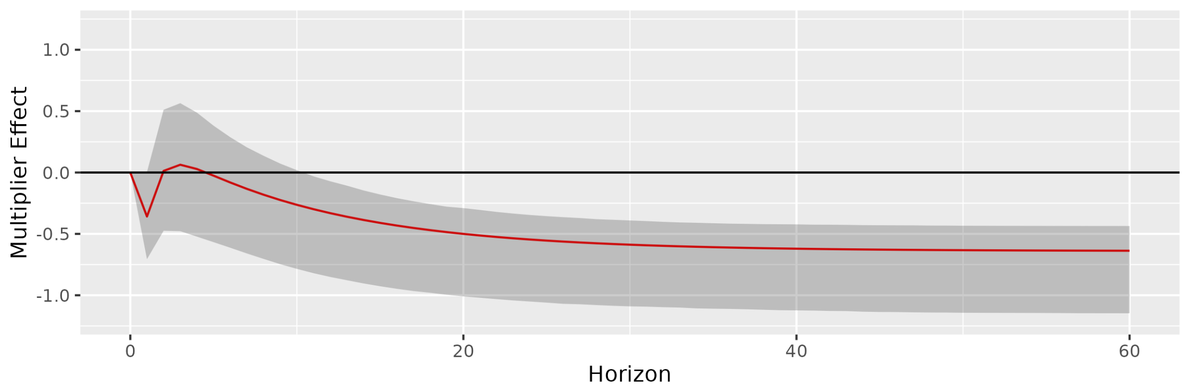


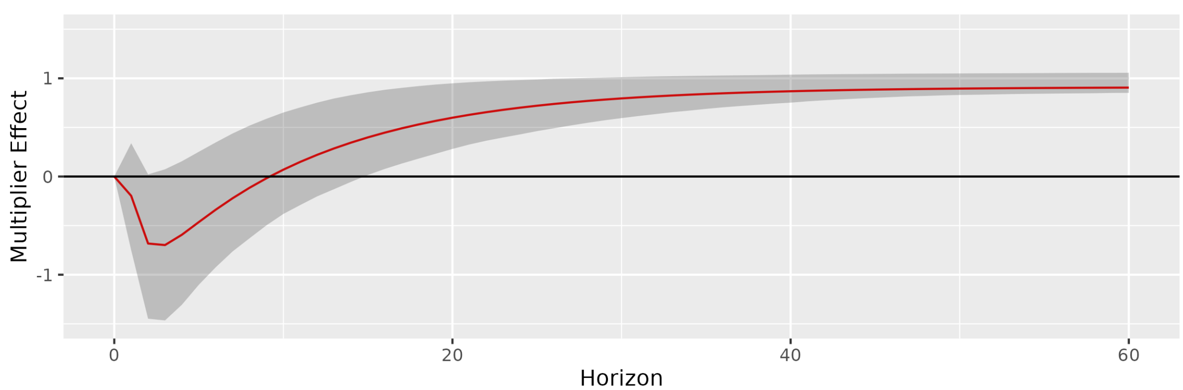
Figure 2: PHASE DIAGRAMS AND STEADY STATE. The left figure demonstrates how the steady-state levels are adjusted as the marginal cost function $\kappa'(\cdot)$ of innovative R&D expenditure decreases from $\kappa_0'(\cdot)$ to $\kappa_1'(\cdot)$. The right figure demonstrates how the steady-state levels are adjusted as the marginal cost function of managerial R&D expenditure decreases from $\xi_0'(\cdot)$ to $\xi_1'(\cdot)$.



(a) Cumulative response of i_{t+h} to a +1 unit shock to r_t^+ in period 1



(b) Cumulative response of i_{t+h} to a +1 unit shock to r_t^- in period 1



(c) Asymmetry across horizons

Figure 3: CUMULATIVE DYNAMIC MULTIPLIERS. Panels (a) and (b) present the cumulative dynamic multiplier effects with respect to unit shocks to r_t^+ and r_t^-, respectively, occurring in period 1. Panel (c) shows the asymmetry at each horizon, i.e., the value of the cumulative dynamic multiplier effect in panel (b) subtracted from the corresponding value in panel (a). Empirical 95% confidence intervals obtained from 5,000 replications of a moving block bootstrap procedure are reported throughout.

Online Supplement for
"Two-Step Nonlinear ARDL Estimation of the Relationship between R&D Intensity and Investment"

by

Jin Seo Cho[a], Matthew Greenwood-Nimmo[b,c,d], and Yongcheol Shin[e]

[a]Yonsei University

[b]University of Melbourne, [c]Australian National University, [d]Codera Analytics

[e]University of York

This Online Supplement consists of six sections. In Sections A.1 and A.2, we provide the proofs of the main claims of the manuscript. In Section A.3, we employ the extreme production function in (1) and conduct a comparative analysis to confirm the theoretical predictions in Section 2.2. In Sections A.4 and A.5, we explore further singularity issues and provide an alternative methodology to consistently estimate the NARDL model with $k > 1$. Sections A.6 and A.7 present additional simulation and estimation results.

A.1 Preliminary Equations

We first provide some equations for an efficient exposition of our proofs. As they are already explained, we simply provide them without reiterating their motivation and derivations.

$$\boldsymbol{x}_t^+ = \boldsymbol{\mu}_*^+ t + \sum_{j=1}^{t} \boldsymbol{s}_j^+ \quad \text{and} \quad \boldsymbol{x}_t^- = \boldsymbol{\mu}_*^- t + \sum_{j=1}^{t} \boldsymbol{s}_j^- \tag{A.1}$$

$$y_t = \delta_* t + \sum_{j=1}^{t} d_j \tag{A.2}$$

$$\Delta y_t = \rho_* y_{t-1} + (\theta_*^+ - \theta_*^-) x_{t-1}^+ + \theta_*^- x_{t-1} + \gamma_* + \sum_{j=1}^{p-1} \varphi_{j*} \Delta y_{t-j} + \sum_{j=0}^{q-1} \left(\pi_{j*}^+ \Delta x_{t-j}^+ + \pi_{j*}^- \Delta x_{t-j}^- \right) + e_t \tag{A.3}$$

$$y_t = \alpha_* + \lambda_* x_t^+ + \eta_* x_t + u_t \tag{A.4}$$

$$u_{t-1} := y_{t-1} - \beta_*^+ x_{t-1}^+ - \beta_*^- x_{t-1}^- \tag{A.5}$$

$$\Delta y_t = \rho_* u_{t-1} + \gamma_* + \sum_{j=1}^{p-1} \varphi_{j*} \Delta y_{t-j} + \sum_{j=0}^{q-1} \left(\boldsymbol{\pi}_{j*}^{+\prime} \Delta \boldsymbol{x}_{t-j}^+ + \boldsymbol{\pi}_{j*}^{-\prime} \Delta \boldsymbol{x}_{t-j}^- \right) + e_t \tag{A.6}$$

$$\widehat{\boldsymbol{\varrho}}_T = \boldsymbol{\varrho}_* + \left(\sum_{t=1}^T \boldsymbol{q}_t \boldsymbol{q}_t' \right)^{-1} \left(\sum_{t=1}^T \boldsymbol{q}_t u_t \right) \tag{A.7}$$

$$\widehat{\boldsymbol{\zeta}}_T := \left(\sum_{t=1}^T \boldsymbol{h}_t \boldsymbol{h}_t' \right)^{-1} \left(\sum_{t=1}^T \boldsymbol{h}_t \Delta y_t \right) = \boldsymbol{\zeta}_* + \left(\sum_{t=1}^T \boldsymbol{h}_t \boldsymbol{h}_t' \right)^{-1} \left(\sum_{t=1}^T \boldsymbol{h}_t e_t \right) \tag{A.8}$$

A.2 Proofs

Proof of Lemma 1. (*i*) By (A.1) and (A.2), we derive the following results:

- $T^{-3} \sum_{t=1}^T y_{t-1}^2 = \frac{1}{3}\delta_*^2 + o_{\mathbb{P}}(1)$;
- $T^{-3} \sum_{t=1}^T y_{t-1} \boldsymbol{x}_t^{+\prime} = \frac{1}{3}\delta_* \boldsymbol{\mu}_*^{+\prime} + o_{\mathbb{P}}(1)$;
- $T^{-3} \sum_{t=1}^T y_{t-1} \boldsymbol{x}_t^{-\prime} = \frac{1}{3}\delta_* \boldsymbol{\mu}_*^{-\prime} + o_{\mathbb{P}}(1)$;
- $T^{-3} \sum_{t=1}^T \boldsymbol{x}_t^+ \boldsymbol{x}_t^{+\prime} = \frac{1}{3}\boldsymbol{\mu}_*^+ \boldsymbol{\mu}_*^{+\prime} + o_{\mathbb{P}}(1)$;
- $T^{-3} \sum_{t=1}^T \boldsymbol{x}_t^+ \boldsymbol{x}_t^{-\prime} = \frac{1}{3}\boldsymbol{\mu}_*^+ \boldsymbol{\mu}_*^{-\prime} + o_{\mathbb{P}}(1)$; and
- $T^{-3} \sum_{t=1}^T \boldsymbol{x}_t^- \boldsymbol{x}_t^{-\prime} = \frac{1}{3}\boldsymbol{\mu}_*^- \boldsymbol{\mu}_*^{-\prime} + o_{\mathbb{P}}(1)$.

These limits imply that $T^{-3} \sum_{t=1}^T \boldsymbol{z}_{1t} \boldsymbol{z}_{1t}' = \mathbf{M}_{11} + o_{\mathbb{P}}(1)$.

(*ii*) By (A.1) and (A.2), we note that:

- $T^{-2} \sum_{t=1}^T y_{t-1} = \frac{1}{2}\delta_* + o_{\mathbb{P}}(1)$;
- $T^{-2} \sum_{t=1}^T y_{t-1} \boldsymbol{w}_{1t}' = T^{-2} \sum_{t=1}^T [\delta_*^2 t, \delta_*^2 t, \ldots, \delta_*^2 t] + o_{\mathbb{P}}(1) = \frac{1}{2}\delta_*^2 \boldsymbol{\iota}_{p-1}' + o_{\mathbb{P}}(1)$;
- $T^{-2} \sum_{t=1}^T y_{t-1} \boldsymbol{w}_{2t}' = T^{-2} \sum_{t=1}^T [\delta_* \boldsymbol{\mu}_*^{+\prime} t, \delta_* \boldsymbol{\mu}_*^{+\prime} t, \ldots, \delta_* \boldsymbol{\mu}_*^{+\prime} t] + o_{\mathbb{P}}(1) = \frac{1}{2}\delta_* \boldsymbol{\iota}_q' \otimes \boldsymbol{\mu}_*^{+\prime} + o_{\mathbb{P}}(1)$;
- $T^{-2} \sum_{t=1}^T y_{t-1} \boldsymbol{w}_{3t}' = T^{-2} \sum_{t=1}^T [\delta_* \boldsymbol{\mu}_*^{-\prime} t, \delta_* \boldsymbol{\mu}_*^{-\prime} t, \ldots, \delta_* \boldsymbol{\mu}_*^{-\prime} t] + o_{\mathbb{P}}(1) = \frac{1}{2}\delta_* \boldsymbol{\iota}_q' \otimes \boldsymbol{\mu}_*^{-\prime} + o_{\mathbb{P}}(1)$;
- $T^{-2} \sum_{t=1}^T \boldsymbol{x}_{t-1}^+ = \frac{1}{2}\boldsymbol{\mu}_*^+ + o_{\mathbb{P}}(1)$;
- $T^{-2} \sum_{t=1}^T \boldsymbol{x}_{t-1}^+ \boldsymbol{w}_{1t}' = T^{-2} \sum_{t=1}^T [\delta_* \boldsymbol{\mu}_*^+ t, \delta_* \boldsymbol{\mu}_*^+ t, \ldots, \delta_* \boldsymbol{\mu}_*^+ t] + o_{\mathbb{P}}(1) = \frac{1}{2}\delta_* \boldsymbol{\mu}_*^+ \boldsymbol{\iota}_{p-1}' + o_{\mathbb{P}}(1)$;

- $T^{-2}\sum_{t=1}^{T}\boldsymbol{x}_{t-1}^{+}\boldsymbol{w}_{2t}' = T^{-2}\sum_{t=1}^{T}[\boldsymbol{\mu}_*^{+}\boldsymbol{\mu}_*^{+\prime}t, \boldsymbol{\mu}_*^{+}\boldsymbol{\mu}_*^{+\prime}t, \ldots, \boldsymbol{\mu}_*^{+}\boldsymbol{\mu}_*^{+\prime}t] + o_{\mathbb{P}}(1) = \frac{1}{2}\boldsymbol{\iota}_q' \otimes \boldsymbol{\mu}_*^{+}\boldsymbol{\mu}_*^{+\prime} + o_{\mathbb{P}}(1)$;
- $T^{-2}\sum_{t=1}^{T}\boldsymbol{x}_{t-1}^{+}\boldsymbol{w}_{3t}' = T^{-2}\sum_{t=1}^{T}[\boldsymbol{\mu}_*^{+}\boldsymbol{\mu}_*^{-\prime}t, \boldsymbol{\mu}_*^{+}\boldsymbol{\mu}_*^{-\prime}t, \ldots, \boldsymbol{\mu}_*^{+}\boldsymbol{\mu}_*^{-\prime}t] + o_{\mathbb{P}}(1) = \frac{1}{2}\boldsymbol{\iota}_q' \otimes \boldsymbol{\mu}_*^{+}\boldsymbol{\mu}_*^{-\prime} + o_{\mathbb{P}}(1)$;
- $T^{-2}\sum_{t=1}^{T}\boldsymbol{x}_{t-1}^{-} = -\frac{1}{2}\boldsymbol{\mu}_*^{+} + o_{\mathbb{P}}(1)$;
- $T^{-2}\sum_{t=1}^{T}\boldsymbol{x}_{t-1}^{-}\boldsymbol{w}_{1t}' = T^{-2}\sum_{t=1}^{T}[\delta_*\boldsymbol{\mu}_*^{-}t, \delta_*\boldsymbol{\mu}_*^{-}t, \ldots, \delta_*\boldsymbol{\mu}_*^{-}t] + o_{\mathbb{P}}(1) = \frac{1}{2}\delta_*\boldsymbol{\mu}_*^{-}\boldsymbol{\iota}_{p-1}' + o_{\mathbb{P}}(1)$;
- $T^{-2}\sum_{t=1}^{T}\boldsymbol{x}_{t-1}^{-}\boldsymbol{w}_{2t}' = T^{-2}\sum_{t=1}^{T}[\boldsymbol{\mu}_*^{-}\boldsymbol{\mu}_*^{+\prime}t, \boldsymbol{\mu}_*^{-}\boldsymbol{\mu}_*^{+\prime}t, \ldots, \boldsymbol{\mu}_*^{-}\boldsymbol{\mu}_*^{+\prime}t] + o_{\mathbb{P}}(1) = \frac{1}{2}\boldsymbol{\iota}_q' \otimes \boldsymbol{\mu}_*^{-}\boldsymbol{\mu}_*^{+\prime} + o_{\mathbb{P}}(1)$;
- $T^{-2}\sum_{t=1}^{T}\boldsymbol{x}_{t-1}^{-}\boldsymbol{w}_{3t}' = T^{-2}\sum_{t=1}^{T}[\boldsymbol{\mu}_*^{-}\boldsymbol{\mu}_*^{-\prime}t, \boldsymbol{\mu}_*^{-}\boldsymbol{\mu}_*^{-\prime}t, \ldots, \boldsymbol{\mu}_*^{-}\boldsymbol{\mu}_*^{-\prime}t] + o_{\mathbb{P}}(1) = \frac{1}{2}\boldsymbol{\iota}_q' \otimes \boldsymbol{\mu}_*^{-}\boldsymbol{\mu}_*^{-\prime} + o_{\mathbb{P}}(1)$.

These results imply that $T^{-1}\sum_{t=1}^{T}\boldsymbol{z}_{1t}\boldsymbol{z}_{2t}' = \mathbf{M}_{12} + o_{\mathbb{P}}(1)$.

(*iii*) We note that:

- $T^{-1}\sum_{t=1}^{T}\boldsymbol{w}_{1t}' = \mathbb{E}[\Delta\boldsymbol{y}_{t-1}]' + o_{\mathbb{P}}(1) = \delta_*\boldsymbol{\iota}_{p-1}' + o_{\mathbb{P}}$;
- $T^{-1}\sum_{t=1}^{T}\boldsymbol{w}_{2t}' = [\mathbb{E}[\Delta\boldsymbol{x}_t^{+\prime}], \mathbb{E}[\Delta\boldsymbol{x}_{t-1}^{+\prime}], \ldots, \mathbb{E}[\Delta\boldsymbol{x}_{t-q+1}^{+\prime}]] + o_{\mathbb{P}}(1) = [\boldsymbol{\mu}_*^{+\prime}, \ldots, \boldsymbol{\mu}_*^{+\prime}] + o_{\mathbb{P}}(1) = \boldsymbol{\iota}_q' \otimes \boldsymbol{\mu}_*^{+} + o_{\mathbb{P}}(1)$;
- $T^{-1}\sum_{t=1}^{T}\boldsymbol{w}_{3t}' = [\mathbb{E}[\Delta\boldsymbol{x}_t^{-\prime}], \mathbb{E}[\Delta\boldsymbol{x}_{t-1}^{-\prime}], \ldots, \mathbb{E}[\Delta\boldsymbol{x}_{t-q+1}^{-\prime}]] + o_{\mathbb{P}}(1) = [\boldsymbol{\mu}_*^{-\prime}, \ldots, \boldsymbol{\mu}_*^{-\prime}] + o_{\mathbb{P}}(1) = \boldsymbol{\iota}_q' \otimes \boldsymbol{\mu}_*^{-} + o_{\mathbb{P}}(1)$; and
- $T^{-1}\sum_{t=1}^{T}\boldsymbol{w}_t\boldsymbol{w}_t' = \mathbb{E}[\boldsymbol{w}_t\boldsymbol{w}_t'] + o_{\mathbb{P}}(1)$.

These limits imply that $T^{-1}\sum_{t=1}^{T}\boldsymbol{z}_{2t}\boldsymbol{z}_{2t}' = \mathbf{M}_{22} + o_{\mathbb{P}}(1)$, as desired. ■

Proof of Lemma 2.

(*i*) We note that:

- $T^{-2}\sum_{t=1}^{T}x_t^{+} = T^{-1}\sum_{t=1}^{T}\mu_*^{+}(t/T) + o_{\mathbb{P}}(1) \overset{\mathbb{P}}{\to} \frac{1}{2}\mu_*^{+}$;
- $T^{-3/2}\sum_{t=1}^{T}x_t = T^{-1}\sum_{t=1}^{T}(T^{-1/2}\sum_{i=1}^{t}\Delta x_i) \Rightarrow \int\mathcal{B}_x$ using that $T^{-1/2}\sum_{i=1}^{[T(\cdot)]}\Delta x_i \Rightarrow \int_0^{(\cdot)} d\mathcal{B}_x$;
- $T^{-3}\sum_{t=1}^{T}x_t^{+}x_t^{+} = T^{-1}\sum_{t=1}^{T}\mu_*^{+}\mu_*^{+}(t/T)^2 + o_{\mathbb{P}}(1) \overset{\mathbb{P}}{\to} \frac{1}{3}\mu_*^{+}\mu_*^{+}$;
- $T^{-5/2}\sum_{t=1}^{T}x_t^{+}x_t = T^{-1}\sum_{t=1}^{T}\mu_*^{+}(t/T)(T^{-1/2}\sum_{i=1}^{t}\Delta x_t) + o_{\mathbb{P}}(1) \Rightarrow \mu_*^{+}\int r\mathcal{B}_x$; and
- $T^{-2}\sum_{t=1}^{T}x_tx_t = T^{-1}\sum_{t=1}^{T}(T^{-1/2}\sum_{i=1}^{t}\Delta x_t)(T^{-1/2}\sum_{i=1}^{t}\Delta x_t) \Rightarrow \int\mathcal{B}_x^2$.

Therefore, $\widehat{\mathbf{Q}}_T \Rightarrow \boldsymbol{\mathcal{Q}}$, as desired.

(*ii*) We note that:

- $T^{-1/2}\sum_{t=1}^{T}u_t \Rightarrow \int d\mathcal{B}_u$ using that $T^{-1/2}\sum_{t=1}^{[T(\cdot)]}u_t \Rightarrow \int_0^{(\cdot)} d\mathcal{B}_u$;

- $T^{-3/2}\sum_{t=1}^T x_t^+ u_t = T^{-1/2}\sum_{t=1}^T \mu_*^+(t/T)u_t + o_{\mathbb{P}}(1) \Rightarrow \mu_*^+ \int r d\mathcal{B}_u$; and
- $T^{-1}\sum_{t=1}^T x_t u_t = T^{-1/2}\sum_{t=1}^T (T^{-1/2}\sum_{i=1}^t \Delta x_i) u_t \Rightarrow \int \mathcal{B}_x d\mathcal{B}_u + \upsilon_*$ using the fact that $\upsilon_* := \lim_{T\to\infty} T^{-1}\sum_{t=1}^T\sum_{i=1}^t \mathbb{E}[\Delta x_i u_t]$ is finite.

Therefore, $\widehat{\mathbf{U}}_T \Rightarrow \boldsymbol{\mathcal{U}}$. ■

Proof of Corollary 1.

Given (A.7), the desired result follows from Lemma 4. ■

Proof of Theorem 1.

We note that $T\{(\widehat{\beta}_T^+ - \widehat{\beta}_T^-) - (\beta_*^+ - \beta_*^-)\} = O_{\mathbb{P}}(T^{-1/2})$ by the definition of $\widehat{\lambda}_T$. Therefore, the weak limit of $T(\widehat{\beta}_T^+ - \beta_*^+)$ is equivalent to that of $T(\widehat{\beta}_T^- - \beta_*^-)$. Furthermore, Corollary 1 implies that $T(\widehat{\beta}_T^- - \beta_*^-) \Rightarrow \mathbf{S}\boldsymbol{\mathcal{Q}}^{-1}\boldsymbol{\mathcal{U}}$, leading to the desired result. ■

Proof of Lemma 3.

Given Assumption 2, we note that $\widetilde{\upsilon}_T \overset{\mathbb{P}}{\to} \upsilon_*$ and $(\widetilde{\sigma}_T^{(1,1)})^{-1}\widetilde{\sigma}_T^{(1,2)} \overset{\mathbb{P}}{\to} \boldsymbol{\nu}_* := (\sigma_*^{(1,1)})^{-1}\sigma_*^{(1,2)}$. Therefore, if we let $\dot{u}_t := u_t - \Delta x_t \boldsymbol{\nu}_*$, $\widetilde{\mathbf{U}}_T = \widetilde{\mathbf{D}}_T^{-1}\sum_{t=1}^T\{\boldsymbol{q}_t \dot{u}_t - \mathbf{S}'\upsilon_*\} + o_{\mathbb{P}}(1)$, then $\widetilde{\mathbf{U}}_T \Rightarrow [\int d\mathcal{B}_{\dot{u}}, \mu_*^+ \int r d\mathcal{B}_{\dot{u}}, \int \mathcal{B}_x d\mathcal{B}_{\dot{u}}]'$, where $\mathcal{B}_{\dot{u}}(\cdot) := \tau_* \mathcal{W}_u(\cdot)$. Therefore, $\widetilde{\mathbf{U}}_T \Rightarrow \widetilde{\boldsymbol{\mathcal{U}}}$. ■

Proof of Corollary 2.

Given that $\widetilde{\mathbf{D}}_T(\widetilde{\boldsymbol{\varrho}}_T - \boldsymbol{\varrho}_*) = [\widetilde{\mathbf{D}}_T^{-1}(\sum_{t=1}^T \boldsymbol{q}_t\boldsymbol{q}_t')\widetilde{\mathbf{D}}_T^{-1}]^{-1}\widetilde{\mathbf{U}}_T$, the desired result follows from Lemmas 2(*i*) and 3. ■

Proof of Theorem 2.

Given that $(\widetilde{\beta}_T^+ - \beta_*^+) - (\widetilde{\beta}_T^- - \beta_*^-) = \widetilde{\lambda}_T - \lambda_* = O_{\mathbb{P}}(T^{-3/2})$ and $(\widetilde{\beta}_T^- - \beta_*^-) = O_{\mathbb{P}}(T^{-1})$, it follows that $(\widetilde{\beta}_T^+ - \beta_*^+) = O_{\mathbb{P}}(T^{-1})$, implying that the weak limit of $T(\widetilde{\beta}_T^+ - \beta_*^+)$ is equivalent to that of $T(\widetilde{\beta}_T^- - \beta_*^-)$. Furthermore, Corollary 1 implies that $T(\widetilde{\eta}_T^- - \eta_*^-) = T(\widetilde{\beta}_T^- - \beta_*^-) \Rightarrow \mathbf{S}\boldsymbol{\mathcal{Q}}^{-1}\widetilde{\boldsymbol{\mathcal{U}}}$, leading to the desired result. ■

Proof of Lemma 4.

This result is easily obtained using the ergodic theorem and the multivariate central limit theorem. ■

Proof of Theorem 3.

(*i*) Given (A.8), we can combine Lemmas 4 (*i* and *ii*) to obtain the desired result.

(*ii*) If it further holds that $\mathbb{E}[e_t^2|\boldsymbol{h}_t] = \sigma_*^2$, then Lemma 4(*iii*) implies that $\boldsymbol{\Omega}_* = \sigma_*^2\boldsymbol{\Gamma}_*$. Therefore, Theorem 3(*i*) now implies that $\sqrt{T}(\widehat{\boldsymbol{\zeta}}_T - \boldsymbol{\zeta}_*) \overset{\text{A}}{\sim} N(\mathbf{0}, \sigma_*^2\boldsymbol{\Gamma}_*^{-1})$. ■

Proof of Theorem 4.

Corollary 2 implies that $T^{3/2}(\widetilde{\boldsymbol{\lambda}}_T - r) \Rightarrow \mathbf{S}\boldsymbol{\mathcal{Q}}^{-1}\widetilde{\boldsymbol{\mathcal{U}}}$ under H_0'', while Lemma 2(*i*) implies that $\widehat{\mathbf{Q}}_T := \widetilde{\mathbf{D}}_T^{-1}(\sum_{t=1}^T \boldsymbol{q}_t\boldsymbol{q}_t')\widetilde{\mathbf{D}}_T^{-1} \Rightarrow \boldsymbol{\mathcal{Q}}$. Furthermore, Assumption 2(*i*) implies that $\widetilde{\tau}_T^2 = \tau_*^2 + o_{\mathbb{P}}(1)$. Given the mixed normal distribution of the FM-OLS estimator for the long-run parameter in Corollary 2, it follows that $\mathcal{W}_T^{(\ell)} \overset{\text{A}}{\sim} \mathcal{X}_1^2$ under $\mathcal{H}_0''$.

In addition, we note that $\widetilde{\mathcal{W}}_T^{(\ell)} = (\widetilde{\mathbf{R}}\widetilde{\boldsymbol{\varrho}}_T - \mathbf{r})'\widetilde{\mathbf{D}}_T(\widetilde{\tau}_T^2\widetilde{\mathbf{R}}\widehat{\mathbf{Q}}_T^{-1}\widetilde{\mathbf{R}}')^{-1}\widetilde{\mathbf{D}}_T(\widetilde{\mathbf{R}}\widetilde{\boldsymbol{\varrho}}_T - \mathbf{r})$, and Theorem 2 implies that $\widetilde{\mathbf{D}}_T(\widetilde{\mathbf{R}}\widetilde{\boldsymbol{\varrho}}_T - \mathbf{r}) \overset{\text{A}}{\sim} N(\mathbf{0}, \tau_*^2\widetilde{\mathbf{R}}\boldsymbol{\mathcal{Q}}^{-1}\widetilde{\mathbf{R}}')$ conditional on $\sigma\{\mathcal{B}_x(r) : r \in (0,1]\}$ under H_0'''. Given the condition that $\widehat{\mathbf{Q}}_T \Rightarrow \boldsymbol{\mathcal{Q}}$ and $\widetilde{\tau}_T^2 \overset{\mathbb{P}}{\to} \tau_*^2$, it now follows that $\widetilde{\mathcal{W}}_T^{(\ell)} \overset{\text{A}}{\sim} \mathcal{X}_2^2$ under H_0'''.

Given that $(\widetilde{\boldsymbol{\lambda}}_T - \boldsymbol{\lambda}_*) = O_{\mathbb{P}}(T^{-3/2})$, $\mathcal{W}_T^{(\ell)} = O_{\mathbb{P}}(T^3)$ under H_1''. Therefore, for any $c_T = o(T^3)$, $\mathbb{P}(\mathcal{W}_T^{(\ell)} > c_T) \to 1$. Furthermore, $(\widetilde{\beta}_T - \beta_*) = O_{\mathbb{P}}(T^{-1})$, implying that $\widetilde{\mathcal{W}}_T^{(\ell)} = O_{\mathbb{P}}(T^2)$ under H_1'''. Therefore, for any $\widetilde{c}_T = o(T^2)$, $\mathbb{P}(\widetilde{\mathcal{W}}_T^{(\ell)} > \widetilde{c}_T) \to 1$. This completes the proof. ■

Proof of Theorem 5.

As this is similar to the standard case, we omit the proof. ■

A.3 A Simple Comparative Static Analysis

We consider a simple comparative static analysis using the production function in Section 2.2, which is helpful in fixing the idea of the two different roles of the R&D activity. We describe the key roles of the capital converted from each type of R&D activity in terms of a complementary or substitutionary relationship with respect to physical capital, as follows:

$$y = \min[c, k + s], \tag{A.9}$$

where y, c, s, and k denote output, the capital converted from the innovative and managerial R&D expenditures and physical capital, respectively.

Suppose that the firm determines the optimal levels of capital by maximizing the profit level. We examine the following optimization problem:

$$\{c_*, k_*, s_*\} := \arg\max_{c,k,s} \ \min[c, k+s] - \kappa_0(c) - p_k k - p_s s,$$

where p_k and p_s denote the unit prices of k and s, respectively, and $\kappa_0(\cdot)$ denotes the cost function needed to convert innovative R&D activity to c. We let $\kappa_0'(\cdot) > 0$ and $\kappa_0''(\cdot) > 0$, and this particular cost function is assumed for the existence of the equilibrium.[1] Given the complementary production function between c and $a := k + s$, the optimization problem is solved in two stages: first, we suppose that c is given. Then, an efficient output is determined by letting $c = k + s$ to exploit the complementary relationship, letting $(k_*(c), s_*(c)) := \arg\max_{k,s} (k+s) - \kappa_0(c) - p_k k - p_s s$ such that $k + s = c$ and obtaining

$$(k_*(c), s_*(c)) = \begin{cases} (c, 0), & \text{if } p_k < p_s; \\ (0, c), & \text{if } p_k > p_s. \end{cases}$$

Second, we plug this optimal demand function back into the profit function and next optimize it with respect to c. That is,

$$c_* = \arg\max_{c} \ c - \kappa_0(c) - p_k k_*(c) - p_s s_*(c) = \arg\max_{c} \begin{cases} (1-p_k)c - \kappa_0(c), & \text{if } p_k < p_s; \\ (1-p_s)c - \kappa_0(c), & \text{if } p_k > p_s, \end{cases}$$

implying that from the first-order condition, the optimal c_* is such that $(1 - p_k) = \kappa_0'(c_*)$, if $p_k < p_s$; and $(1 - p_s) = \kappa_0'(c_*)$, if $p_k > p_s$. If we again plug this optimal level c_* back to $k_*(\cdot)$ and $s_*(\cdot)$, it follows that $(k_*, s_*) = (c_*, 0)$ if $p_k < p_s$; and $(k_*, s_*) = (0, c_*)$ if $p_k > p_s$.

The optimal demands for the capital respond to external shocks. We separately examine the complement and substitution effects by modifying the expenditure function. First, we suppose that $p_k < p_s$ with the initial cost function $\kappa_0(\cdot)$ and let $\kappa_0(\cdot)$ decrease to $\kappa_1(\cdot)$ such that $\kappa_0(\cdot) > \kappa_1(\cdot)$ with $\kappa_1'(\cdot) > 0$ and $\kappa_1''(\cdot) > 0$. By this supposition, the price for the innovative R&D activity uniformly decreases, and the optimal demand for c increases to c_{**} from c_* such that $(1-p_k) = \kappa_1'(c_{**})$ and $(k_{**}, s_{**}) = (c_{**}, 0)$.

[1] For the existence of the equilibrium, it is also possible to suppose a concave production function of c and a unit price for c, so that the optimization problem is modified to $\max_{c,k,s} \min[g(c), k+s] - p_c c - p_k k - p_s s$, where $g(\cdot)$ is such that $g'(\cdot) > 0$, $g''(\cdot) < 0$, and p_c is the unit price of c.

Note that the demands for c and k move in the same direction but the demand for s remains the same, implying a complementary relationship between c and k.

Next, we examine a different static analysis. We suppose that $p_k < p_s^0$ initially but p_s decreases to p_s^1 such that $p_s^1 < p_k$ without modifying the cost function $\kappa_0(\cdot)$. By this decrease in p_s, the optimal demand for c remains the same, so that $c_{**} = c_*$ but the optimal demands for k and s shift $(k_{**}, s_{**}) = (0, c_*)$. That is, demands for s and k move in opposite directions, implying a substitutionary relationship between s and k.

The response of the physical capital to the external shocks describes its core properties in relation to the R&D expenditure. The innovative R&D expenditure complements the capital, so that the production activity cannot extend its scope without investing in the innovative R&D activity. By having an environment favorable to innovative R&D activity, the firm can enhance its potential to produce more outputs by investing more capital. On the contrary, the managerial R&D expenditure substitutes the capital, so that the relative price of the physical capital to the capital converted from the managerial R&D activity determines the optimal level of the physical capital. If the output can be produced more efficiently by conducting the managerial R&D activity which is cheaper than the cost of the capital, the firm can enhance its production efficiency by substituting the managerial R&D activity for the capital.

A.4 Further Singularity Issues Associated with Single-Step NARDL Estimation

It is important to realize that the re-parameterization of the long-run relationship that we propose to resolve the singularity issue under 2-step estimation in (A.4) is insufficient to resolve the singularity issue involved in single-step NARDL estimation. In fact, efforts to estimate the short-run and the long-run parameters in a single step by combining (A.5) with (A.6) will encounter a further singularity problem, even if $k = 1$. Using the definitions of $\lambda_* := \beta_*^+ - \beta_*^-$ and $\eta_* := \beta_*^-$, it follows that $u_{t-1} = y_{t-1} - \lambda_* x_{t-1}^+ - \beta_* x_{t-1}$, such that:

$$\Delta y_t = \rho_* y_{t-1} + (\theta_*^+ - \theta_*^-) x_{t-1}^+ + \theta_*^- x_{t-1} + \gamma_* + \sum_{j=1}^{p-1} \varphi_{j*} \Delta y_{t-j} + \sum_{j=0}^{q-1} \left(\pi_{j*}^+ \Delta x_{t-j}^+ + \pi_{j*}^- \Delta x_{t-j}^- \right) + e_t,$$

where $\boldsymbol{\beta}_*^+ := -\theta_*^+/\rho_*$ and $\boldsymbol{\beta}_*^- := -\theta_*^-/\rho_*$. Let:

$$\boldsymbol{\xi}_* := \left[\begin{array}{c:c} \boldsymbol{\xi}_{1*}' & \boldsymbol{\xi}_{2*}' \end{array}\right]' := \left[\begin{array}{ccc:c} \rho_* & \theta_* & \theta_*^- & \boldsymbol{\alpha}_{2*}' \end{array}\right]' \quad \text{and}$$

$$\boldsymbol{p}_t := \left[\begin{array}{c:c} \boldsymbol{p}_{1t}' & \boldsymbol{p}_{2t}' \end{array}\right]' := \left[\begin{array}{ccc:c} y_{t-1} & x_{t-1}^+ & x_{t-1} & \boldsymbol{z}_{2t}' \end{array}\right]'.$$

Note that $\boldsymbol{\xi}_{2*}$ and $\boldsymbol{p}_{2t}$ are identical to $\boldsymbol{\alpha}_{2*}$ and $\boldsymbol{z}_{2t}$, respectively, where $\theta_* := \theta_*^+ - \theta_*^-$. If we attempt to estimate the vector of unknown parameters, $\boldsymbol{\xi}_*$, in (A.3) by OLS, we obtain:

$$\widehat{\boldsymbol{\xi}}_T := \left(\sum_{t=1}^{T} \boldsymbol{p}_t \boldsymbol{p}_t'\right)^{-1} \left(\sum_{t=1}^{T} \boldsymbol{p}_t \Delta y_t\right).$$

We demonstrate that the inverse matrix in $\widehat{\boldsymbol{\xi}}_T$ is asymptotically singular in Lemma 1:

Lemma A.1. *Given Assumption 1:*

(i) $\ddot{\mathbf{D}}_{1,T}^{-1}\left(\sum_{t=1}^{T} \boldsymbol{p}_{1t}\boldsymbol{p}_{1t}'\right)\ddot{\mathbf{D}}_{1,T}^{-1} \Rightarrow \boldsymbol{\mathcal{P}}_{11}$, *where* $\ddot{\mathbf{D}}_{1,T} := diag[T^{3/2}\mathbf{I}_2, T]$ *and:*

$$\boldsymbol{\mathcal{P}}_{11} := \begin{bmatrix} \frac{1}{3}\delta_*^2 & \frac{1}{3}\delta_*\mu_*^+ & \delta_* \int r\mathcal{B}_x \\ \frac{1}{3}\delta_*\mu_*^+ & \frac{1}{3}\mu_*^+\mu_*^+ & \mu_*^+ \int r\mathcal{B}_x \\ \delta_* \int r\mathcal{B}_x & \int r\mathcal{B}_x\mu_*^+ & \int \mathcal{B}_x^2 \end{bmatrix};$$

(ii) $\ddot{\mathbf{D}}_{1,T}^{-1}\left(\sum_{t=1}^{T} \boldsymbol{p}_{1t}\boldsymbol{p}_{2t}'\right)\ddot{\mathbf{D}}_{2,T}^{-1} \Rightarrow \boldsymbol{\mathcal{P}}_{12}$, *where* $\ddot{\mathbf{D}}_{2,T} := diag[T^{1/2}\mathbf{I}_{1+p+2q}]$ *and:*

$$\boldsymbol{\mathcal{P}}_{12} := \begin{bmatrix} \frac{1}{2}\delta_* & \frac{1}{2}\delta_*^2\boldsymbol{\iota}_{p-1}' & \frac{1}{2}\delta_*\boldsymbol{\iota}_q' \otimes \mu_*^+ & \frac{1}{2}\delta_*\boldsymbol{\iota}_q' \otimes \mu_*^- \\ \frac{1}{2}\mu_*^+ & \frac{1}{2}\delta_*\mu_*^+\boldsymbol{\iota}_{p-1}' & \frac{1}{2}\boldsymbol{\iota}_q' \otimes \mu_*^+\mu_*^+ & \frac{1}{2}\boldsymbol{\iota}_q' \otimes \mu_*^+\mu_*^- \\ \int \mathcal{B}_x & \delta_* \int \mathcal{B}_x\boldsymbol{\iota}_{p-1}' & \boldsymbol{\iota}_q' \otimes \int \mathcal{B}_x\mu_*^+ & \boldsymbol{\iota}_q' \otimes \int \mathcal{B}_x\mu_*^- \end{bmatrix}; \quad \textit{and}$$

(iii) $\ddot{\mathbf{D}}_{2,T}^{-1}\left(\sum_{t=1}^{T} \boldsymbol{p}_{2t}\boldsymbol{p}_{2t}'\right)\ddot{\mathbf{D}}_{2,T}^{-1} \xrightarrow{\mathbb{P}} \mathbf{P}_{22} := \mathbf{M}_{22}$. □

We omit the proof of Lemma A.1, as it can be easily derived from the proof of Lemma 1. Let

$\ddot{\mathbf{D}}_T := \text{diag}[T^{3/2}\mathbf{I}_2, T, T^{1/2}\mathbf{I}_{1+p+2q}]$, then:

$$\ddot{\mathbf{D}}_T^{-1}\left(\sum_{t=1}^{T}\boldsymbol{p}_t\boldsymbol{p}_t'\right)\ddot{\mathbf{D}}_T^{-1} \Rightarrow \boldsymbol{\mathcal{P}} := \begin{bmatrix} \boldsymbol{\mathcal{P}}_{11} & \boldsymbol{\mathcal{P}}_{12} \\ \boldsymbol{\mathcal{P}}_{21} & \mathbf{P}_{22} \end{bmatrix},$$

where $\boldsymbol{\mathcal{P}}_{21} := \boldsymbol{\mathcal{P}}_{12}'$. Since $\boldsymbol{\mathcal{P}}$ is singular, it is difficult to obtain the limit distribution of $\widehat{\boldsymbol{\xi}}_T$ using the one-step approach even after re-parameterizing the long-run level relationship in (A.4).

A.5 An Alternative NARDL Estimation with Multiple Regressors

If there are multiple explanatory variables in the NARDL model, then the two-step estimation procedure described in Section 3 needs to be modified as follows: Let $\boldsymbol{x}_t \equiv \boldsymbol{x}_t^+ + \boldsymbol{x}_t^-$, $\boldsymbol{\lambda}_* = \boldsymbol{\beta}_*^+ - \boldsymbol{\beta}_*^-$ and $\boldsymbol{\eta}_* = \boldsymbol{\beta}_*^-$ with $k > 1$. Then, we have: $y_t = \alpha_* + \boldsymbol{\lambda}_*'\boldsymbol{x}_t^+ + \boldsymbol{\eta}_*'\boldsymbol{x}_t + u_t$. By extending Lemma 2, it follows that:

$$\widehat{\mathbf{Q}}_T := \widetilde{\mathbf{D}}_T^{-1}\left(\sum_{t=1}^{T}\boldsymbol{q}_t\boldsymbol{q}_t'\right)\widetilde{\mathbf{D}}_T^{-1} \Rightarrow \boldsymbol{\mathcal{Q}} := \begin{bmatrix} 1 & \frac{1}{2}\boldsymbol{\mu}_*^+ & \int \boldsymbol{\mathcal{B}}_x' \\ \frac{1}{2}\boldsymbol{\mu}_*^+ & \frac{1}{3}\boldsymbol{\mu}_*^+\boldsymbol{\mu}_*^{+\prime} & \boldsymbol{\mu}_*^+ \int r\boldsymbol{\mathcal{B}}_x' \\ \int \boldsymbol{\mathcal{B}}_x dr & \int r\boldsymbol{\mathcal{B}}_x\boldsymbol{\mu}_*^{+\prime} & \int \boldsymbol{\mathcal{B}}_x\boldsymbol{\mathcal{B}}_x' \end{bmatrix},$$

where $\boldsymbol{q}_t := (1, \boldsymbol{x}_t^{+\prime}, \boldsymbol{x}_t')'$, $\widetilde{\mathbf{D}}_T := \text{diag}[T^{1/2}, T^{3/2}\mathbf{I}_k, T\mathbf{I}_k]$, $\boldsymbol{\mathcal{B}}_x(\cdot)$ is a $k \times 1$ vector of Brownian motions, such that $[\boldsymbol{\mathcal{B}}_x(\cdot)', \mathcal{B}_u(\cdot)]' := \boldsymbol{\Sigma}_*^{1/2}[\boldsymbol{\mathcal{W}}_x(\cdot)', \mathcal{W}_u(\cdot)']'$ and $\boldsymbol{\Sigma}_* := \lim_{T\to\infty} T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathbb{E}[\boldsymbol{g}_t\boldsymbol{g}_s]$, with $\boldsymbol{g}_t := [\Delta\boldsymbol{x}_t, u_t]'$ and $[\boldsymbol{\mathcal{W}}_x(\cdot)', \mathcal{W}_u(\cdot)]'$ being a $(k+1) \times 1$ vector of independent Wiener processes. The blocks on the second row of $\boldsymbol{\mathcal{Q}}$ form a sub-matrix with rank equal to unity. That is, $[\frac{1}{2}\boldsymbol{\mu}_*^+, \frac{1}{3}\boldsymbol{\mu}_*^+\boldsymbol{\mu}_*^{+\prime}, \boldsymbol{\mu}_*^+\int r\boldsymbol{\mathcal{B}}_x'] = \boldsymbol{\mu}_*^+[\frac{1}{2}, \frac{1}{3}\boldsymbol{\mu}_*^{+\prime}, \int r\boldsymbol{\mathcal{B}}_x']$, implying that $\boldsymbol{\mathcal{Q}}$ is a singular matrix with probability 1. Consequently, the two-step procedure in Section 3 cannot be directly applied to estimating the long-run parameters.

A.5.1 Estimation and Inference for $k > 1$

In order to address the above singularity issue, Cho, Greenwood-Nimmo, and Shin (2023a) propose a substantial modification to the estimation and inference procedure. We now provide a summary of this modified methodology.

Define $\boldsymbol{m}_t := \sum_{j=1}^{t} \boldsymbol{s}_j^+$, which is a unit-root process with zero-mean increments. Therefore, if $\boldsymbol{x}_t^+$ is regressed against t, then $\boldsymbol{\mu}_*^+$ can be estimated by $\widehat{\boldsymbol{\mu}}_T^+ := (\sum_{t=1}^{T} t^2)^{-1} \sum_{t=1}^{T} t\boldsymbol{x}_t^+ = \boldsymbol{\mu}_*^+ + (\sum_{t=1}^{T} t^2)^{-1} \sum_{t=1}^{T} t\boldsymbol{m}_t$, where $\widehat{\boldsymbol{m}}_t := \boldsymbol{x}_t^+ - \widehat{\boldsymbol{\mu}}_T^+ t$ is the regression residual. Consequently, $\boldsymbol{m}_t = \widehat{\boldsymbol{m}}_t + t\boldsymbol{d}_T$, such that $\boldsymbol{x}_t^+ = \widehat{\boldsymbol{m}}_t + (\boldsymbol{\mu}_*^+ + \boldsymbol{d}_T)t$, where $\boldsymbol{d}_T := (\sum_{t=1}^{T} t^2)^{-1} \sum_{t=1}^{T} t\boldsymbol{m}_t$. Under mild regularity conditions, $\boldsymbol{d}_T = O_{\mathbb{P}}(T^{-1/2})$. Let $\boldsymbol{\delta}_{*T} := \boldsymbol{\mu}_*^+ + \boldsymbol{d}_T$. Then, $\boldsymbol{\delta}_{*T} = \boldsymbol{\mu}_* + O_{\mathbb{P}}(T^{-1/2})$. We rewrite the equation for y_t using the equation for $\boldsymbol{x}_t^+$ to obtain $y_t = \alpha_* + \boldsymbol{\lambda}_*'(\boldsymbol{\mu}_*^+ + \boldsymbol{d}_T)t + \boldsymbol{\lambda}_*'\widehat{\boldsymbol{m}}_t + \boldsymbol{\eta}_*'\boldsymbol{x}_t + u_t$. Next, we can consistently estimate $\boldsymbol{\lambda}_*$ and $\boldsymbol{\eta}_*$ by regressing y_t on $\boldsymbol{r}_t := (1, t, \widehat{\boldsymbol{m}}_t', \boldsymbol{x}_t')'$. Let $\ddot{\boldsymbol{\varpi}}_T := (\ddot{\alpha}_T, \ddot{\xi}_{*T}, \ddot{\boldsymbol{\lambda}}_T', \ddot{\boldsymbol{\eta}}_T')'$ be the OLS estimator of $\varpi_{*T} := (\alpha_*, \xi_{*T}, \boldsymbol{\lambda}_*', \boldsymbol{\eta}_*')'$, where $\xi_{*T} := \boldsymbol{\lambda}_*'\boldsymbol{\delta}_{*T}$. The long-run estimators of $\boldsymbol{\beta}_*^+$ and $\boldsymbol{\beta}_*^-$ can be obtained as $\ddot{\boldsymbol{\beta}}_T^+ := \ddot{\boldsymbol{\lambda}}_T + \ddot{\boldsymbol{\eta}}_T$ and $\ddot{\boldsymbol{\beta}}_T^- := \ddot{\boldsymbol{\eta}}_T$. Let:

$$\ddot{\mathrm{S}} := \begin{bmatrix} \mathbf{0}_{k\times 1} & \mathbf{0}_{k\times 1} & \mathbf{I}_k & \mathbf{I}_k \\ \mathbf{0}_{k\times 1} & \mathbf{0}_{k\times 1} & \mathbf{0}_{k\times k} & \mathbf{I}_k \end{bmatrix}, \quad \text{then} \quad \begin{bmatrix} \widehat{\boldsymbol{\beta}}_T^+ \\ \widehat{\boldsymbol{\beta}}_T^- \end{bmatrix} = \ddot{\mathrm{S}}\ddot{\boldsymbol{\varpi}}_T.$$

We refer to this as the *first-step transformed OLS (TOLS) estimator*. The intuition is straightforward—as it is the collinear trend in $\boldsymbol{x}_t^+$ that results in the singularity of $\boldsymbol{\mathcal{Q}}$, we detrend $\boldsymbol{x}_t^+$ prior to estimation.

The limit distribution of the first-step TOLS estimator is obtained similarly to that of the first-step OLS estimator. First, note that $\ddot{\boldsymbol{\varpi}}_T = \boldsymbol{\varpi}_{*T} + (\sum_{t=1}^{T} \boldsymbol{r}_t\boldsymbol{r}_t')^{-1} \sum_{t=1}^{T} \boldsymbol{r}_t u_t$. To characterize the limit behaviors of the components, we define $\ddot{\boldsymbol{\Sigma}}_* := \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} \mathbb{E}[\ddot{\boldsymbol{g}}_t\ddot{\boldsymbol{g}}_t']$ and $[\boldsymbol{\mathcal{B}}_m(\cdot)', \boldsymbol{\mathcal{B}}_x(\cdot)', \mathcal{B}_u(\cdot)]'$ $:= \ddot{\boldsymbol{\Sigma}}_*^{1/2}[\boldsymbol{\mathcal{W}}_m(\cdot)', \boldsymbol{\mathcal{W}}_x(\cdot)', \mathcal{W}_u(\cdot)]'$, where we let $\ddot{\boldsymbol{g}}_t := [\Delta\boldsymbol{m}_t', \Delta\boldsymbol{x}_t, u_t]'$ and $[\boldsymbol{\mathcal{W}}_m(\cdot)', \boldsymbol{\mathcal{W}}_x(\cdot)', \mathcal{W}_u(\cdot)]'$ is a $(2k+1)\times 1$ vector of independent Wiener processes.

Lemma A.2. *Under Assumption 1, if $\ddot{\boldsymbol{\Sigma}}_*$ is PD, then (i) $\ddot{\mathrm{R}}_T := \ddot{\mathrm{D}}_T^{-1}(\sum_{t=1}^{T} \boldsymbol{r}_t\boldsymbol{r}_t')\ddot{\mathrm{D}}_T^{-1} \Rightarrow \boldsymbol{\mathcal{R}}$, where:*

$$\boldsymbol{\mathcal{R}} = \begin{bmatrix} 1 & \frac{1}{2} & \int(1-\frac{3}{2}r)\boldsymbol{\mathcal{B}}_m' & \int \boldsymbol{\mathcal{B}}_x' \\ \frac{1}{2} & \frac{1}{3} & \mathbf{0}_{1\times k} & \int r\boldsymbol{\mathcal{B}}_x' \\ \int(1-\frac{3}{2}r)\boldsymbol{\mathcal{B}}_m & \mathbf{0}_{k\times 1} & \int \boldsymbol{\mathcal{B}}_m\boldsymbol{\mathcal{B}}_m' - 3\int r\boldsymbol{\mathcal{B}}_m \int r\boldsymbol{\mathcal{B}}_m' & \int \boldsymbol{\mathcal{B}}_m\boldsymbol{\mathcal{B}}_x' - 3\int r\boldsymbol{\mathcal{B}}_m \int r\boldsymbol{\mathcal{B}}_x' \\ \int \boldsymbol{\mathcal{B}}_x & \int r\boldsymbol{\mathcal{B}}_x & \int \boldsymbol{\mathcal{B}}_x\boldsymbol{\mathcal{B}}_m' - 3\int r\boldsymbol{\mathcal{B}}_x \int r\boldsymbol{\mathcal{B}}_m' & \int \boldsymbol{\mathcal{B}}_x\boldsymbol{\mathcal{B}}_x' \end{bmatrix},$$

and $\ddot{\mathrm{D}}_T := diag[T^{1/2}, T^{3/2}, T\mathbf{I}_{2k}]$; and (ii) if $\boldsymbol{\upsilon}_{x*} := \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} \sum_{i=1}^{t} \mathbb{E}[\Delta\boldsymbol{x}_i u_t]$ and $\boldsymbol{\upsilon}_{m*} := \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} \sum_{i=1}^{t} \mathbb{E}[\Delta\boldsymbol{m}_i u_t]$ are finite, then $\ddot{\mathrm{U}}_T := \ddot{\mathrm{D}}_T^{-1}(\sum_{t=1}^{T} \boldsymbol{r}_t u_t) \Rightarrow \ddot{\boldsymbol{\mathcal{U}}} := [\int d\mathcal{B}_u, \int r d\mathcal{B}_u,$ $\int \boldsymbol{\mathcal{B}}_m' d\mathcal{B}_u - 3\int r d\mathcal{B}_u \int r\boldsymbol{\mathcal{B}}_m' + \boldsymbol{\upsilon}_{m*}', \int \boldsymbol{\mathcal{B}}_x' d\mathcal{B}_u + \boldsymbol{\upsilon}_{x*}']'$. □

$\boldsymbol{\mathcal{R}}$ is no longer singular, because $\sum_{t=1}^{[T(\cdot)]} \ddot{\boldsymbol{g}}_t$ obeys the FCLT using partially correlated increments.

Corollary A.1. *Given Assumption 1,* $\ddot{\mathrm{D}}_T(\ddot{\boldsymbol{\varpi}}_T - \boldsymbol{\varpi}_{*T}) \Rightarrow \boldsymbol{\mathcal{R}}^{-1}\ddot{\boldsymbol{\mathcal{U}}}$ *and* $T^{1/2}(\widehat{\xi}_T - \boldsymbol{\lambda}_*'\boldsymbol{\mu}_*^+) \Rightarrow 3\boldsymbol{\lambda}_*' \int r\boldsymbol{\mathcal{B}}_m$. □

The first part of Corollary A.1 follows from the first-step TOLS estimator. For the second part, $\widehat{\xi}_T$ is not of primary interest. Although the convergence rate of $(\widehat{\xi}_T - \xi_{*T})$ is $T^{3/2}$, $\xi_{*T} := \boldsymbol{\lambda}_*'\boldsymbol{\delta}_{*T}$ converges to $\boldsymbol{\lambda}_*'\boldsymbol{\mu}_*^+$ at the rate $T^{1/2}$. This implies that $T^{1/2}(\widehat{\xi}_T - \boldsymbol{\lambda}_*'\boldsymbol{\mu}_*^+)$ is asymptotically bounded in probability.

Theorem A.1. *Under Assumption 1, and if* $\ddot{\boldsymbol{\Sigma}}_*$ *is PD, then* $T[(\ddot{\boldsymbol{\beta}}_T^+ - \boldsymbol{\beta}_*^+)', (\ddot{\boldsymbol{\beta}}_T^- - \boldsymbol{\beta}_*^-)']' \Rightarrow \ddot{\mathrm{S}}\boldsymbol{\mathcal{R}}^{-1}\ddot{\boldsymbol{\mathcal{U}}}$. □

The limit distribution of the TOLS estimator in Theorem A.1 does not provide a basis for inference, as it is non-standard and exhibits asymptotic bias driven by $\boldsymbol{\upsilon}_{*m}$ and $\boldsymbol{\upsilon}_{*x}$. Thus, we provide an alternative first-step FM-OLS estimator. We make the following assumptions:

Assumption A.1. *(i)* $\ddot{\boldsymbol{\Sigma}}_*$ *is finite and PD, such that there is a consistent estimator for* $\ddot{\boldsymbol{\Sigma}}_*$*:*

$$\bar{\boldsymbol{\Sigma}}_T := \begin{bmatrix} \bar{\boldsymbol{\Sigma}}_T^{(1,1)} & \bar{\boldsymbol{\sigma}}_T^{(1,2)} \\ \bar{\boldsymbol{\sigma}}_T^{(2,1)} & \bar{\sigma}_T^{(2,2)} \end{bmatrix} \xrightarrow{\mathbb{P}} \ddot{\boldsymbol{\Sigma}}_* := \begin{bmatrix} \ddot{\boldsymbol{\Sigma}}_*^{(1,1)} & \ddot{\boldsymbol{\sigma}}_*^{(1,2)} \\ \ddot{\boldsymbol{\sigma}}_*^{(2,1)} & \sigma_*^{(2,2)} \end{bmatrix};$$

and (ii) if we let $\bar{\boldsymbol{\Pi}}_T := T^{-1}\sum_{k=0}^{\ell}\sum_{t=k+1}^{T} \bar{\boldsymbol{g}}_{t-k}\bar{\boldsymbol{g}}_t'$*, then:*

$$\begin{bmatrix} \bar{\boldsymbol{\Pi}}_T^{(1,1)} & \bar{\boldsymbol{\pi}}_T^{(1,2)} \\ \bar{\boldsymbol{\pi}}_T^{(2,1)} & \bar{\pi}_T^{(2,2)} \end{bmatrix} := \bar{\boldsymbol{\Pi}}_T \xrightarrow{\mathbb{P}} \ddot{\boldsymbol{\Pi}}_* := \begin{bmatrix} \ddot{\boldsymbol{\Pi}}_*^{(1,1)} & \ddot{\boldsymbol{\pi}}_*^{(1,2)} \\ \ddot{\boldsymbol{\pi}}_*^{(2,1)} & \pi_*^{(2,2)} \end{bmatrix} := \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{t} \mathbb{E}[\ddot{\boldsymbol{g}}_t\ddot{\boldsymbol{g}}_t'],$$

which is finite, where $\bar{\boldsymbol{g}}_t := [\Delta\widehat{\boldsymbol{m}}_t', \Delta\boldsymbol{x}_t', \ddot{u}_t]'$ *and* $\ddot{u}_t := y_t - \ddot{\alpha}_T - \ddot{\boldsymbol{\beta}}_T^{+\prime}\boldsymbol{x}_t^+ - \ddot{\boldsymbol{\beta}}_T^{-\prime}\boldsymbol{x}_t^-$. □

Assumption A.1 corresponds to Assumption 2 in the case with $k > 1$. The *fully-modified TOLS (FM-TOLS) estimator* is defined as $\bar{\boldsymbol{\varpi}}_T := (\sum_{t=1}^{T}\boldsymbol{r}_t\boldsymbol{r}_t')^{-1}(\sum_{t=1}^{T}\boldsymbol{r}_t\bar{y}_t - T\bar{\mathbf{S}}'\bar{\boldsymbol{\upsilon}}_T)$, where $\bar{y}_t := y_t - \boldsymbol{\ell}_t'(\bar{\boldsymbol{\Sigma}}_T^{(1,1)})^{-1}\bar{\boldsymbol{\sigma}}_T^{(1,2)}$, $\boldsymbol{\ell}_t := (\Delta\widehat{\boldsymbol{m}}_t', \Delta\boldsymbol{x}_t')'$, $\bar{\boldsymbol{\upsilon}}_T := \bar{\boldsymbol{\pi}}_T^{(1,2)} - \bar{\boldsymbol{\Pi}}_T^{(1,1)}(\bar{\boldsymbol{\Sigma}}_T^{(1,1)})^{-1}\bar{\boldsymbol{\sigma}}_T^{(1,2)}$ and $\bar{\mathbf{S}} := [\mathbf{0}_{2k\times 2}, \mathbf{I}_{2k}]$.

The following lemma corresponds to Lemma 3 for the case with $k > 1$:

Lemma A.3. *Under Assumptions 1 and A.1,* $\bar{\mathrm{U}}_T := \ddot{\mathrm{D}}_T^{-1}\{\sum_{t=1}^{T}\boldsymbol{r}_t(u_t - \boldsymbol{\ell}_t'(\bar{\boldsymbol{\Sigma}}_T^{(1,1)})^{-1}\bar{\boldsymbol{\sigma}}_T^{(1,2)}) - T\bar{\mathbf{S}}'\bar{\boldsymbol{\upsilon}}_T\} \Rightarrow \bar{\boldsymbol{\mathcal{U}}} := \ddot{\tau}[\int d\mathcal{W}_u, \int r d\mathcal{W}_u, \int \boldsymbol{\mathcal{B}}_m' d\mathcal{W}_u - 3\int r d\mathcal{W}_u \int r\boldsymbol{\mathcal{B}}_m', \int \boldsymbol{\mathcal{B}}_x' d\mathcal{W}_u]'$*, where* $\ddot{\tau}^2 := \mathrm{plim}_{T\to\infty}\bar{\tau}_T^2$ *and* $\bar{\tau}_T^2 := \bar{\sigma}_T^{(2,2)} - \bar{\boldsymbol{\sigma}}_T^{(2,1)}(\bar{\boldsymbol{\Sigma}}_T^{(1,1)})^{-1}\bar{\boldsymbol{\sigma}}_T^{(1,2)}$. □

By Lemmas A.2 and A.3, the limit distribution of the FM-TOLS estimator is obtained as the product of $\boldsymbol{\mathcal{R}}^{-1}$ and $\bar{\boldsymbol{\mathcal{U}}}$. Letting $[\bar{\boldsymbol{\beta}}_T^{+\prime}, \bar{\boldsymbol{\beta}}_T^{-\prime}]' := \ddot{\mathrm{S}}\bar{\boldsymbol{\varpi}}_T$, we obtain its weak limit as follows:

Theorem A.2. *Under Assumptions 1 and A.1,* $\ddot{\mathrm{D}}_T(\bar{\boldsymbol{\varpi}}_T - \bar{\boldsymbol{\varpi}}_{*T}) \Rightarrow \boldsymbol{\mathcal{R}}^{-1}\bar{\boldsymbol{\mathcal{U}}}$ *and* $T[(\bar{\boldsymbol{\beta}}_T^{+} - \boldsymbol{\beta}_*^{+})', (\bar{\boldsymbol{\beta}}_T^{-} - \boldsymbol{\beta}_*^{-})']' \Rightarrow \ddot{\mathrm{S}}\boldsymbol{\mathcal{R}}^{-1}\bar{\boldsymbol{\mathcal{U}}}$. □

The limit distribution of the FM-TOLS estimator is mixed normal: $\bar{\mathrm{D}}_T(\bar{\boldsymbol{\varpi}}_T - \bar{\boldsymbol{\varpi}}_{*T}) \overset{\mathrm{A}}{\sim} N(\mathbf{0}, \ddot{\tau}^2\boldsymbol{\mathcal{R}}^{-1})$ conditional on $\sigma\{(\boldsymbol{\mathcal{B}}_m(r)', \boldsymbol{\mathcal{B}}_x(r)')', r \in (0,1]\}$.

We can treat the FM-TOLS estimator of the long-run parameters as known when estimating the short-run parameters in the second step because it has a convergence rate faster than $T^{1/2}$. Let $u_{t-1} := y_{t-1} - \boldsymbol{\beta}_*^{+\prime}\boldsymbol{x}_{t-1}^{+} - \boldsymbol{\beta}_*^{-\prime}\boldsymbol{x}_{t-1}^{-} = y_{t-1} - \boldsymbol{\lambda}_*'\boldsymbol{\mu}_*^{+}(t-1) - \boldsymbol{\lambda}_*'\boldsymbol{m}_{t-1} - \boldsymbol{\eta}_*'\boldsymbol{x}_{t-1}$, where $\boldsymbol{\lambda}_*$, $\boldsymbol{\mu}_*^{+}$, and $\boldsymbol{\eta}_*$ are assumed to be known. This allows us to rewrite (3) as described in Section 3.3, so that the short-run parameters can be estimated by the OLS estimator $\widehat{\boldsymbol{\zeta}}_T$ and Theorem 3 can be applied.

We next develop a testing procedure for the presence of asymmetries in both the long-run and short-run. Define $\boldsymbol{\beta}_* := (\boldsymbol{\beta}_*^{+\prime}, \boldsymbol{\beta}_*^{-\prime})'$. Consider $H_0^{(4)} : \mathbf{R}\boldsymbol{\beta}_* = \boldsymbol{r}$ vs. $H_1^{(4)} : \mathbf{R}\boldsymbol{\beta}_* \neq \boldsymbol{r}$ for some $\mathbf{R} \in \mathbb{R}^{r \times 2k}$ and $\boldsymbol{r} \in \mathbb{R}^r$ ($r \leq 2k$). We then construct the Wald test as $\ddot{\mathcal{W}}_T := T^2(\mathbf{R}\bar{\boldsymbol{\beta}}_T - \boldsymbol{r})'\{\bar{\tau}_T^2\mathbf{R}\ddot{\mathrm{S}}\ddot{\mathbf{R}}_T^{-1}\ddot{\mathrm{S}}'\mathbf{R}'\}^{-1}(\mathbf{R}\bar{\boldsymbol{\beta}}_T - \boldsymbol{r})$, where $\bar{\boldsymbol{\beta}}_T := (\bar{\boldsymbol{\beta}}_T^{+\prime}, \bar{\boldsymbol{\beta}}_T^{-\prime})'$.

Theorem A.3. *Given Assumptions 1 and A.1,* $\ddot{\mathcal{W}}_T^{(\ell)} \overset{A}{\sim} \mathcal{X}_r^2$ *under* $H_0^{(4)}$*. Further, for any sequence* c_T *such that* $c_T = o(T^2)$*,* $\mathbb{P}(\ddot{\mathcal{W}}_T^{(\ell)} > c_T) \to 1$ *under* $H_1^{(4)}$*.* □

To examine the test for additive symmetry of the short-run dynamic parameters, we consider $H_0 : \mathbf{R}_s\boldsymbol{\zeta}_* = \mathbf{r}$ vs. $H_1 : \mathbf{R}_s\boldsymbol{\zeta}_* \neq \mathbf{r}$, where $\mathbf{R}_s \in \mathbb{R}^{r \times (1+p+2qk)}$ and $\mathbf{r} \in \mathbb{R}^r$ are selection matrices, and $\boldsymbol{\zeta}_*$ generalizes our prior definition for $k = 1$, viz., $\boldsymbol{\zeta}_* := (\rho_*, \gamma_*, \varphi_{1*}, \ldots, \varphi_{p-1*}, \boldsymbol{\pi}_{0*}^{+\prime}, \ldots, \boldsymbol{\pi}_{q-1*}^{+\prime}, \boldsymbol{\pi}_{0*}^{-\prime}, \ldots, \boldsymbol{\pi}_{q-1*}^{-\prime})$. Let $\mathbf{R}_s := \left[\mathbf{0}_{k \times (1+p)}, \boldsymbol{\iota}_q' \otimes \mathbf{I}_k, -\boldsymbol{\iota}_q \otimes \mathbf{I}_k\right]$ and $\mathbf{r} = \mathbf{0}$. Then, we can test the null hypothesis of additive short-run symmetry: $H_0 : \sum_{j=0}^{q-1}\boldsymbol{\pi}_{j*}^{+} = \sum_{j=0}^{q-1}\boldsymbol{\pi}_{j*}^{-}$ vs. $H_1 : \sum_{j=0}^{q-1}\boldsymbol{\pi}_{j*}^{+} \neq \sum_{j=0}^{q-1}\boldsymbol{\pi}_{j*}^{-}$. If the cointegration residuals are obtained as in Section 3.3, then we can employ the same Wald test statistic introduced in Section 4 for $k > 1$.

A.6 Additional Monte Carlo Simulations

We provide additional simulation evidence. First, we set $k = 2$ and conduct simulations to examine the finite sample performance of the Wald tests presented in Section A.5. Next, we treat the cases of

$k = 1$ and $k > 1$ separately and study the finite sample bias and mean squared error (MSE) of the parameters estimated by our two-step procedure, where the estimator for the long-run parameters is either OLS or FM-OLS (TOLS or FM-TOLS) and the estimator for the short-run parameters is OLS.

A.6.1 Monte Carlo Simulations when $k = 2$

We examine the finite sample performance of the Wald tests in Section A.5 for $k = 2$. The data are generated by the NARDL(1,0) DGP: $\Delta y_t = \gamma_* + \rho_* u_{t-1} + \varphi_* \Delta y_{t-1} + \boldsymbol{\pi}_{0*}^{+\prime} \Delta \boldsymbol{x}_t^+ + \pi_{0*}^{-\prime} \Delta \boldsymbol{x}_t^- + e_t$, where $u_{t-1} := y_{t-1} - \alpha_* - \boldsymbol{\beta}_*^{+\prime} \boldsymbol{x}_{t-1}^+ - \boldsymbol{\beta}_*^{-\prime} \boldsymbol{x}_{t-1}^-$, $\Delta \boldsymbol{x}_t := \kappa_* \Delta \boldsymbol{x}_{t-1} + \sqrt{1-\kappa_*^2} \boldsymbol{v}_t$, and $(e_t, \boldsymbol{v}_t)' \sim$ IID$N(\mathbf{0}_3, \mathbf{I}_3)$.

Testing the Long-Run Parameters We focus on the case where the FM-TOLS estimator is used. We generate the data by setting $(\alpha_*, \gamma_*, \rho_*, \varphi_*, \kappa_*) = (0, 0, -1, \varphi_*, 0.5)$, $(\boldsymbol{\beta}_*^{+\prime}, \boldsymbol{\beta}_*^{-\prime})' = (-1, 0.5, 0.75, -1.5)'$, and $(\boldsymbol{\pi}_{0*}^{+\prime}, \boldsymbol{\pi}_{0*}^{-\prime})' = (0.5, -0.5, -1, 1)'$. We test $\dot{H}_0^{(\ell)} : \boldsymbol{\iota}_2' \boldsymbol{\beta}_*^+ = -0.50$ and $\boldsymbol{\iota}_2' \boldsymbol{\beta}_*^- = -0.75$ vs. $\dot{H}_1^{(\ell)} : \boldsymbol{\iota}_2' \boldsymbol{\beta}_*^+ \neq -0.50$ or $\boldsymbol{\iota}_2' \boldsymbol{\beta}_*^- \neq -0.75$ while allowing φ_* to vary over -0.2, -0.1, 0, 0.1 and 0.2. The simulation results are reported in Table A.1. As T increases, the distribution of the Wald test statistic becomes well approximated by the chi-squared distribution with two degrees of freedom. For $\varphi_* = 0.3$, a larger sample size is required to achieve a satisfactory approximation.[2]

— Insert Table A.1 Here —

To examine the empirical power of the Wald test statistic, we test $\dot{H}_0^{(\ell)} : \boldsymbol{\iota}_2' \boldsymbol{\beta}_*^+ = -0.40$ and $\boldsymbol{\iota}_2' \boldsymbol{\beta}_*^- = -0.65$ vs. $\dot{H}_1^{(\ell)} : \boldsymbol{\iota}_2' \boldsymbol{\beta}_*^+ \neq -0.40$ or $\boldsymbol{\iota}_2' \boldsymbol{\beta}_*^- \neq -0.65$. Table A.2 shows that the Wald test statistic is consistent under the alternative hypothesis, irrespective of the value of φ_*.

— Insert Table A.2 Here —

Testing the Short-Run Parameters Here, we set $(\alpha_*, \gamma_*, \rho_*, \varphi_*, \kappa_*) = (0, 0, -1, \varphi_*, 0.5)$, $(\boldsymbol{\beta}_*^{+\prime}, \boldsymbol{\beta}_*^{-\prime})' = (-1, 0.5, 0.75, -1.5)'$, and $(\boldsymbol{\pi}_{0*}^{+\prime}, \boldsymbol{\pi}_{0*}^{-\prime})' = (0.5, 0.2, 0.5, 0.2)'$, while allowing φ_* to vary as before. We limit our attention to the case where we estimate the long-run parameters by the FM-TOLS estimator and compute $\widehat{u}_t$ prior to estimating the short-run parameters by OLS. We then test

[2] In this situation the use of resampling methods may be advisable.

$\dot{H}_0^{(s)}$: $\boldsymbol{\pi}_{0*}^{+} - \boldsymbol{\pi}_{0*}^{-} = \mathbf{0}$ vs. $\dot{H}_1^{(s)}$: $\boldsymbol{\pi}_{0*}^{+} - \boldsymbol{\pi}_{0*}^{-} \neq \mathbf{0}$, using the Wald test with the heteroskedasticity consistent covariance estimator, $\widehat{\boldsymbol{\Omega}}_T$.

Table A.3 shows that the finite sample distribution of the Wald test statistic is well-approximated by the chi-squared distribution. For each significance level, its empirical level tends to the nominal level, once T reaches 1,000. In addition, the empirical levels display little sensitivity to φ_*, even for moderate T.

— Insert Table A.3 Here —

Finally, we examine the empirical power of the Wald test statistic. We work with the same DGP and test $\ddot{H}_0^{(s)}$: $\boldsymbol{\pi}_{0*}^{+} - \boldsymbol{\pi}_{0*}^{-} = 0.3\boldsymbol{\iota}$ vs. $\ddot{H}_1^{(s)}$: $\boldsymbol{\pi}_{0*}^{+} - \boldsymbol{\pi}_{0*}^{-} \neq 0.3\boldsymbol{\iota}$. We use the same Wald test statistic and report the simulation results in Table A.4. The empirical power of the test increases sharply with T and exhibits little sensitivity to the degree of autocorrelation measured by φ_*.

— Insert Table A.4 Here —

A.6.2 Monte Carlo Simulations with $k = 1$

In this section, we examine the bias and MSE of the single-step and two-step NARDL estimations. We generate simulated data using the same NARDL(1,0) DGP in Section 5 of the manuscript:

$$\Delta y_t = \gamma_* + \rho_* u_{t-1} + \varphi_* \Delta y_{t-1} + \pi_*^{+} \Delta x_t^{+} + \pi_*^{-} \Delta x_t^{-} + e_t,$$

where $u_{t-1} := y_{t-1} - \alpha_* - \beta_*^{+} x_{t-1}^{+} - \beta_*^{-} x_{t-1}^{-}$, $\Delta x_t := \kappa_* \Delta x_{t-1} + \sqrt{1-\kappa_*^2} v_t$, and $(e_t, v_t)' \sim$ IID$N(\mathbf{0}_2, \mathbf{I}_2)$. We set $(\alpha_*, \beta_*^{+}, \beta_*^{-}, \gamma_*, \rho_*, \varphi_*, \pi_*^{+}, \pi_*^{-}, \kappa_*) = (0, 2, 1, 0, -2/3, \varphi_*, 1, 1/2, 1/2)$ and we allow the sample size, T, and the parameter φ_* to vary. Note that Δx_t is generated by an AR(1) process with normally distributed disturbances and that u_t is both serially correlated and contemporaneously correlated with Δx_t.

Next, we specify the following long-run and short-run models:

$$y_t = \alpha + \lambda x_t^{+} + \eta x_t + u_t \quad \text{and} \quad \Delta y_t = \gamma + \rho \widehat{u}_{t-1} + \varphi_1 \Delta y_{t-1} + \pi_0^{+} \Delta x_t^{+} + \pi_0^{-} \Delta x_t^{-} + e_t,$$

where $\widehat{u}_t := y_t - \widehat{\alpha}_T - \widehat{\lambda}_T x_t^+ - \widehat{\eta}_T x_t$. We estimate these models in two steps. In the first step, we estimate the parameters of the long-run relationship using either OLS or FM-OLS. In the second step, we estimate the short-run parameters by OLS. In each case, we evaluate the performance of the estimators by comparing their finite sample bias and MSE. We calculate the bias as follows:

$$\text{Bias}_T(\beta_*^+) := R^{-1}\sum_{j=1}^{R}(\widehat{\beta}_{T,j}^+ - \beta_*^+) \quad \text{and} \quad \text{Bias}_T(\varphi_*) := R^{-1}\sum_{j=1}^{R}(\widehat{\varphi}_{T,j} - \varphi_*),$$

where R is the number of replications used in the simulation experiment, $\widehat{\beta}_T^+$ is obtained in the first step by OLS or FM-OLS and $\widehat{\varphi}_T$ is obtained in the second-step by OLS. Likewise, we calculate the finite sample MSE of $\widehat{\beta}_T^+$ and $\widehat{\varphi}_T$ as:

$$\text{MSE}_T(\beta_*^+) := R^{-1}\sum_{j=1}^{R}(\widehat{\beta}_{T,j}^+ - \beta_*^+)^2 \quad \text{and} \quad \text{MSE}_T(\varphi_*) := R^{-1}\sum_{j=1}^{R}(\widehat{\varphi}_{T,j} - \varphi_*)^2.$$

The finite sample bias and MSE of the estimated parameters based on $R = 5,000$ replications of the simulation experiments are reported in Tables A.5 and A.6, respectively. To conserve space, we do not report the finite sample bias or MSE for the intercepts, α and γ, but these results are available from the authors on request.

— Insert Tables A.5 and A.6 Here —

First, consider the long-run parameter estimators obtained in the first step. The finite sample bias of the FM-OLS estimator is substantially smaller than that of the first-step OLS estimator. Recall that FM-OLS yields normally distributed estimators for the long-run parameters, β_*^+ and β_*^-. Consequently, in most cases, we find that the finite sample bias of the FM-OLS estimator is close to zero, because $T(\widetilde{\beta}_T^+ - \beta_*^+)$ and $T(\widetilde{\beta}_T^- - \beta_*^-)$ are asymptotically mixed-normally distributed around zero. By contrast, the OLS estimator is not asymptotically distributed around zero and exhibits non-negligible bias. In addition, our simulation results indicate that the FM-OLS estimator is often more efficient than its OLS counterpart, resulting in a smaller MSE as the sample size increases. This tendency is particularly apparent for small and/or negative values of φ_*, although it is likely to be different for other nuisance parameters. Taken as a whole, these results strongly favor the use of FM-OLS in the first step.

Now, consider the short-run parameter estimators obtained by OLS in the second step. We find that the finite sample biases of the second step OLS estimators of the dynamic parameters become negligible as the sample size increases. This is true irrespective of whether we use OLS or FM-OLS in the first step, although the smallest biases are obtained in almost all cases when the first-step estimator is FM-OLS. The MSEs of the second-step OLS estimators are similar irrespective of the use of OLS or FM-OLS in the first step. Even for a small sample of just 50 observations, the bias is minor in all cases. This is an encouraging observation because many existing applications of the NARDL model rely on small datasets, constrained by the low sampling frequency and limited history of many macroeconomic databases.

A.6.3 Monte Carlo Simulations when $k = 2$

We generate simulated data using the following NARDL(1,0) DGP, which is the same DGP as in Section A.6.1.

$$\Delta y_t = \gamma_* + \rho_* u_{t-1} + \varphi_* \Delta y_{t-1} + \boldsymbol{\pi}_{0*}^{+\prime} \Delta \boldsymbol{x}_t^+ + \pi_{0*}^{-\prime} \Delta \boldsymbol{x}_t^- + e_t,$$

where $u_{t-1} := y_{t-1} - \alpha_* - \boldsymbol{\beta}_*^{+\prime} \boldsymbol{x}_{t-1}^+ - \boldsymbol{\beta}_*^{-\prime} \boldsymbol{x}_{t-1}^-$, $\Delta \boldsymbol{x}_t := \kappa_* \Delta \boldsymbol{x}_{t-1} + \sqrt{1 - \kappa_*^2} \boldsymbol{v}_t$, and $(e_t, \boldsymbol{v}_t)' \sim$ IID$N(\mathbf{0}_3, \mathbf{I}_3)$. We set $(\alpha_*, \gamma_*, \rho_*, \varphi_*, \kappa_*) = (0, 0, -1, \varphi_*, 0.5)$, $(\boldsymbol{\beta}_*^{+\prime}, \boldsymbol{\beta}_*^{-\prime})' = (-1, 0.5, 0.75, -1.5)'$, and $(\boldsymbol{\pi}_{0*}^{+\prime}, \boldsymbol{\pi}_{0*}^{-\prime})' = (0.5, -0.5, -1, 1)'$. As before, we allow T and φ_* to vary and examine the effect of different degrees of serial correlation.

We next specify the long-run and short-run models as: $y_t = \alpha + \boldsymbol{\lambda}' \boldsymbol{x}_t^+ + \boldsymbol{\eta} \boldsymbol{x}_t + u_t$ and $\Delta y_t = \gamma + \rho \widehat{u}_{t-1} + \varphi_1 \Delta y_{t-1} + \boldsymbol{\pi}_0^+ \Delta \boldsymbol{x}_t^+ + \boldsymbol{\pi}_0^- \Delta \boldsymbol{x}_t^- + e_t$, where $\widehat{u}_t$ is the regression residual obtained from the first step estimation. In the first step, we estimate the long-run parameters by TOLS or FM-TOLS, and then subsequently estimate the short-run parameters by OLS. As in the case with $k = 1$, we evaluate the performance of the estimators by studying their respective finite sample bias and MSE, which are obtained using $R = 5,000$ replications. The results are reported in Tables A.7 and A.8.

— Insert Tables A.7 and A.8 Here —

We first examine the long-run parameter estimators obtained in the first step. As T increases, the finite sample bias of the FM-TOLS estimator becomes smaller than that of the TOLS estimator.

Recall that FM-OLS yields normally distributed estimators for the long-run parameters, $\boldsymbol{\beta}_*^+$ and $\boldsymbol{\beta}_*^-$. Consequently, we find that the finite sample bias of the FM-TOLS estimator approaches zero much faster than the TOLS estimator as T increases. In addition, the FM-TOLS estimator becomes more efficient than the TOLS estimator as T increases, resulting in a smaller MSE at larger sample sizes. These results favor the use of the FM-TOLS estimator in samples of small to moderate size.

We finally consider the short-run dynamic parameter estimators obtained by OLS. We find that the finite sample biases of the OLS estimator become negligible as the sample size rises. This is true irrespective of the use of either TOLS or FM-TOLS in the preceding step. Likewise, the MSEs of the OLS estimators of the dynamic parameters are similar irrespective of whether we use TOLS or FM-TOLS in the first step, particularly as the sample size becomes larger.

A.7 Additional Empirical Evidence

In Table A.9, we report the descriptive statistics of both the R&D ratio to GDP and the log of GPDI. The descriptive statistics reveal that the differenced R&D ratio is more standard than the differenced log of GPDI in the sense that the first has the characteristics of a normal distribution centered at zero, whereas the latter is more widely distributed, centered at a non-zero value with a negative skew and excess kurtosis.

— Insert Table A.9 Here —

We apply Phillips and Perron's (1988) unit root test to both series. We test the unit root hypothesis both including and excluding the time trend. The results are reported in Table A.10. Irrespective of the presence of the time trend, the unit-root hypothesis cannot be rejected, implying that both series are unit-root nonstationary. Furthermore, the time trend in the Phillips and Perron equation for unit root testing is statistically significant for the log of GPDI but insignificant for the R&D ratio to GDP, implying that the log of GPDI is a unit root process with a time drift, while the R&D ratio to GDP is a unit root process without a time drift, viz., $\mathbb{E}[\Delta r_t] = 0$ but $\mathbb{E}[\Delta i_t] > 0$.

— Insert Table A.10 Here —

Finally, we replicate the NARDL estimation exercise using the single-step NARDL estimator popularized by SYG. The cumulative dynamic multipliers obtained from the single-step parameter estimates reported in Figure A.1 are very similar to those obtained from the two-step estimates in Figure A.1, indicating that both estimation procedures yield similar results in practice.

— Insert Figure A.1 Here —

φ_*	sample size	100	250	500	1,000	3,000	5,000
-0.20	1%	44.58	17.42	7.64	3.48	1.20	1.42
	5%	58.30	32.32	18.08	10.60	5.56	5.64
	10%	65.52	40.84	26.28	17.64	10.98	10.68
-0.10	1%	42.96	18.12	8.12	3.32	1.48	1.24
	5%	57.82	31.10	19.04	10.76	6.02	5.48
	10%	65.34	39.20	26.66	17.48	11.64	10.34
0.00	1%	42.50	16.78	8.44	4.30	1.42	1.28
	5%	56.38	29.90	19.20	11.62	5.80	5.60
	10%	64.06	38.70	28.20	18.74	11.36	10.76
0.10	1%	41.24	16.18	8.56	4.02	1.68	1.16
	5%	56.14	30.00	19.30	11.32	6.68	5.80
	10%	63.30	38.78	27.12	19.24	12.50	10.88
0.20	1%	40.68	14.64	8.34	4.22	2.12	1.34
	5%	53.64	28.22	18.44	11.70	6.74	5.82
	10%	61.40	37.22	26.22	18.58	12.22	11.32

Table A.1: EMPIRICAL LEVELS OF THE WALD TEST FOR LONG-RUN SYMMETRY. The FM-TOLS estimator is used in the first step. The data is generated as follows: $\Delta y_t = -u_{t-1} + \varphi_* \Delta y_{t-1} + \boldsymbol{\pi}_{0*}^{+\prime} \Delta \boldsymbol{x}_t^+ + \boldsymbol{\pi}_{0*}^{-\prime} \Delta \boldsymbol{x}_t^- + e_t$, where $u_t := y_t - \boldsymbol{\beta}_*^{+\prime} \boldsymbol{x}_t^+ - \boldsymbol{\beta}_*^{-\prime} \boldsymbol{x}_t^-$, $\Delta \boldsymbol{x}_t = 0.5 \Delta \boldsymbol{x}_{t-1} + \sqrt{1 - 0.5^2} \boldsymbol{v}_t$, and $(e_t, \boldsymbol{v}_t')' \sim \text{IID}N(\mathbf{0}_3, \mathbf{I}_3)$. $\dot{H}_0^{(\ell)} : \boldsymbol{\iota}_2' \boldsymbol{\beta}_*^+ = -0.5$ and $\boldsymbol{\iota}_2' \boldsymbol{\beta}_*^- = -0.75$ vs. $\dot{H}_1^{(\ell)} : \boldsymbol{\iota}_2' \boldsymbol{\beta}_*^+ \neq -0.5$ or $\boldsymbol{\iota}_2' \boldsymbol{\beta}_*^- \neq -0.75$. The simulation results are obtained using $R = 5,000$ replications. When $\varphi_* = 0.2$ and $n = 10,000$, the empirical levels are 1.18%, 5.52%, and 10.50% for the empirical levels 1%, 5%, and 10%, respectively.

φ_*	sample size	100	200	300	400	500
-0.20	1%	47.46	70.16	90.44	98.10	99.74
	5%	62.40	80.76	94.26	99.00	99.88
	10%	70.16	85.42	96.18	99.32	99.92
-0.10	1%	48.76	71.48	91.24	98.04	99.68
	5%	62.14	81.50	94.80	99.08	99.84
	10%	70.28	86.02	96.42	99.36	99.86
0.00	1%	47.32	70.54	90.56	98.20	99.78
	5%	61.72	81.70	94.02	99.12	99.94
	10%	69.78	86.56	95.44	99.46	99.94
0.10	1%	46.36	69.20	90.22	98.06	99.86
	5%	60.26	79.82	94.00	98.98	99.96
	10%	67.16	84.56	95.86	99.36	99.98
0.20	1%	43.10	63.52	88.70	97.58	99.44
	5%	56.98	75.22	93.74	98.88	99.78
	10%	65.50	80.66	95.68	99.18	99.88

Table A.2: EMPIRICAL POWER OF THE WALD TEST FOR LONG-RUN SYMMETRY (IN PERCENT). The FM-TOLS estimator is used in the first step. The data is generated as follows: $\Delta y_t = -u_{t-1} + \varphi_* \Delta y_{t-1} + \boldsymbol{\pi}_{0*}^{+\prime} \Delta \boldsymbol{x}_t^+ + \boldsymbol{\pi}_{0*}^{-\prime} \Delta \boldsymbol{x}_t^- + e_t$, where $u_t := y_t - \boldsymbol{\beta}_*^{+\prime} \boldsymbol{x}_t^+ - \boldsymbol{\beta}_*^{-\prime} \boldsymbol{x}_t^-$, $\Delta \boldsymbol{x}_t = 0.5 \Delta \boldsymbol{x}_{t-1} + \sqrt{1 - 0.5^2} \boldsymbol{v}_t$, and $(e_t, \boldsymbol{v}_t')' \sim \text{IID}N(\mathbf{0}_3, \mathbf{I}_3)$. $\ddot{H}_0^{(\ell)} : \boldsymbol{\iota}_2' \boldsymbol{\beta}_*^+ = -0.4$ and $\boldsymbol{\iota}_2' \boldsymbol{\beta}_*^- = -0.65$ vs. $\ddot{H}_1^{(\ell)} : \boldsymbol{\iota}_2' \boldsymbol{\beta}_*^+ \neq -0.4$ or $\boldsymbol{\iota}_2' \boldsymbol{\beta}_*^- \neq -0.65$. The simulation results are obtained using $R = 5,000$ replications.

φ_*	sample size	100	200	400	600	800	1,000
-0.20	1%	10.92	3.98	2.10	1.88	1.42	1.06
	5%	22.80	11.20	7.20	6.60	6.40	5.62
	10%	32.52	17.88	13.38	12.12	11.40	10.56
-0.10	1%	10.58	3.72	2.02	1.34	1.20	1.22
	5%	22.50	11.12	7.56	6.36	6.06	5.84
	10%	32.18	18.66	13.38	12.28	11.38	10.94
0.00	1%	10.90	3.92	2.00	1.54	1.48	1.28
	5%	23.34	11.80	7.90	6.36	6.64	5.50
	10%	31.46	19.30	13.96	11.36	12.46	10.50
0.10	1%	10.80	3.98	2.00	1.36	1.44	1.26
	5%	22.60	11.96	7.36	6.54	6.46	5.40
	10%	30.78	18.68	12.74	12.38	12.16	10.82
0.20	1%	10.26	3.76	2.38	1.70	1.50	1.38
	5%	23.06	11.68	7.76	6.58	6.28	5.72
	10%	32.16	19.26	14.02	12.28	11.68	10.68

Table A.3: EMPIRICAL LEVELS OF THE WALD TEST FOR SHORT-RUN SYMMETRY (IN PERCENT). The FM-TOLS estimator is used in the first step and OLS is used in the second step. The data is generated as follows: $\Delta y_t = -u_{t-1} + \varphi_* \Delta y_{t-1} + \boldsymbol{\pi}_{0*}^{+\prime} \Delta \boldsymbol{x}_t^+ + \boldsymbol{\pi}_{0*}^{-\prime} \Delta \boldsymbol{x}_t^- + e_t$, where $u_t := y_t - \boldsymbol{\beta}_*^{+\prime} \boldsymbol{x}_t^+ - \boldsymbol{\beta}_*^{-\prime} \boldsymbol{x}_t^-$, $\Delta \boldsymbol{x}_t = 0.5 \Delta \boldsymbol{x}_{t-1} + \sqrt{1 - 0.5^2} \boldsymbol{v}_t$, and $(e_t, \boldsymbol{v}_t')' \sim$ IID$N(\mathbf{0}_3, \mathbf{I}_3)$. $\dot{H}_0^{(s)} : \boldsymbol{\pi}_{0*}^+ - \boldsymbol{\pi}_{0*}^- = \mathbf{0}_2$ vs. $\dot{H}_1^{(s)} : \boldsymbol{\pi}_{0*}^+ - \boldsymbol{\pi}_{0*}^- \neq \mathbf{0}_2$. The simulation results are obtained using $R = 5,000$ replications.

φ_*	sample size	100	200	400	600	800	1,000
-0.20	1%	20.74	21.68	40.30	61.64	76.56	86.20
	5%	36.78	40.48	62.66	80.90	90.96	95.76
	10%	46.58	51.82	73.06	87.76	95.04	98.26
-0.10	1%	21.56	22.94	41.36	61.54	76.68	86.96
	5%	37.70	41.56	63.86	80.14	91.18	95.40
	10%	46.98	52.92	74.76	87.70	95.16	97.72
0.00	1%	21.04	23.38	42.32	61.96	76.86	86.66
	5%	36.70	42.28	64.08	81.12	90.50	95.58
	10%	46.52	53.74	74.38	88.60	94.46	97.34
0.10	1%	20.86	24.26	41.74	61.58	76.62	87.28
	5%	37.22	42.92	63.96	81.92	90.42	95.84
	10%	46.84	54.56	74.36	88.70	94.80	97.94
0.20	1%	20.48	23.32	42.08	62.02	76.58	86.82
	5%	36.56	42.84	63.60	80.88	90.20	95.28
	10%	46.76	53.20	73.78	87.98	94.48	97.42

Table A.4: EMPIRICAL POWER OF THE WALD TEST FOR SHORT-RUN SYMMETRY (IN PERCENT). The FM-TOLS estimator is used in the first step and OLS is used in the second step. The data is generated as follows: $\Delta y_t = -u_{t-1} + \varphi_* \Delta y_{t-1} + \boldsymbol{\pi}_{0*}^{+\prime} \Delta \boldsymbol{x}_t^+ + \boldsymbol{\pi}_{0*}^{-\prime} \Delta \boldsymbol{x}_t^- + e_t$, where $u_t := y_t - \boldsymbol{\beta}_*^{+\prime} \boldsymbol{x}_t^+ - \boldsymbol{\beta}_*^{-\prime} \boldsymbol{x}_t^-$, $\Delta \boldsymbol{x}_t = 0.5 \Delta \boldsymbol{x}_{t-1} + \sqrt{1 - 0.5^2} \boldsymbol{v}_t$, and $(e_t, \boldsymbol{v}_t')' \sim$ IID$N(\mathbf{0}_3, \mathbf{I}_3)$. $\ddot{H}_0^{(s)} : \boldsymbol{\pi}_{0*}^+ - \boldsymbol{\pi}_{0*}^- = 0.3\boldsymbol{\iota}_2$ vs. $\ddot{H}_1^{(s)} : \boldsymbol{\pi}_{0*}^+ - \boldsymbol{\pi}_{0*}^- \neq 0.3\boldsymbol{\iota}_2$. The simulation results are obtained using $R = 5,000$ replications.

	Sample Size	50		100		150		200		250	
	First Step	OLS	FM-OLS	OLS	FM-OLS	OLS	FM-OLS	OLS	FM-OLS	OLS	FM-OLS
φ_*	Second Step	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
-0.50	β_*^+	-0.263	-0.130	-0.140	-0.038	-0.095	-0.017	-0.072	-0.010	-0.058	-0.006
	β_*^-	-0.269	-0.038	-0.140	-0.009	-0.095	-0.004	-0.071	-0.002	-0.058	-0.001
	ρ_*	-0.101	-0.083	-0.036	-0.029	-0.022	-0.017	-0.015	-0.012	-0.012	-0.009
	φ_*	0.112	0.074	0.055	0.028	0.036	0.016	0.028	0.012	0.023	0.009
	π_*^+	-0.062	-0.022	-0.026	0.004	-0.018	0.007	-0.012	0.005	-0.010	0.006
	π_*^-	-0.107	-0.038	-0.044	-0.008	-0.032	-0.008	-0.024	-0.006	-0.020	-0.005
-0.25	β_*^+	-0.185	-0.073	-0.098	-0.020	-0.066	-0.007	-0.050	-0.002	-0.040	-0.001
	β_*^-	-0.192	0.004	-0.099	0.002	-0.066	0.003	-0.050	0.004	-0.041	0.003
	ρ_*	-0.084	-0.070	-0.030	-0.026	-0.018	-0.017	-0.012	-0.012	-0.011	-0.010
	φ_*	0.088	0.044	0.048	0.016	0.033	0.008	0.025	0.006	0.020	0.004
	π_*^+	-0.042	-0.002	-0.024	0.002	-0.017	0.003	-0.010	0.005	-0.009	0.005
	π_*^-	-0.069	-0.011	-0.032	-0.005	-0.026	-0.007	-0.019	-0.004	-0.015	-0.003
0.00	β_*^+	-0.112	-0.033	-0.057	-0.010	-0.037	-0.001	-0.027	0.000	-0.022	0.002
	β_*^-	-0.117	0.023	-0.057	0.004	-0.037	0.005	-0.027	0.005	-0.022	0.004
	ρ_*	-0.081	-0.069	-0.035	-0.034	-0.024	-0.023	-0.017	-0.016	-0.013	-0.012
	φ_*	0.051	0.016	0.028	0.007	0.019	0.003	0.015	0.001	0.011	0.000
	π_*^+	-0.035	-0.010	-0.018	-0.006	-0.009	0.001	-0.008	0.000	-0.006	0.002
	π_*^-	-0.047	-0.007	-0.026	-0.011	-0.017	-0.007	-0.015	-0.007	-0.009	-0.003
0.25	β_*^+	-0.024	0.000	-0.009	-0.006	-0.006	-0.003	-0.004	-0.001	-0.004	-0.001
	β_*^-	-0.027	0.024	-0.010	-0.003	-0.007	-0.001	-0.004	0.000	-0.004	-0.001
	ρ_*	-0.072	-0.068	-0.035	-0.036	-0.022	-0.023	-0.017	-0.018	-0.014	-0.014
	φ_*	0.015	0.001	0.006	0.002	0.005	0.001	0.003	0.000	0.003	0.000
	π_*^+	-0.001	-0.004	-0.004	-0.011	0.001	-0.004	-0.003	-0.007	-0.001	-0.005
	π_*^-	-0.024	-0.022	-0.007	-0.015	-0.007	-0.012	-0.004	-0.008	-0.005	-0.008
0.50	β_*^+	0.065	0.015	0.034	-0.023	0.024	-0.016	0.018	-0.013	0.014	-0.011
	β_*^-	0.062	-0.004	0.034	-0.035	0.024	-0.021	0.017	-0.016	0.014	-0.012
	ρ_*	-0.046	-0.046	-0.022	-0.019	-0.015	-0.013	-0.011	-0.010	-0.008	-0.007
	φ_*	-0.016	-0.008	-0.009	0.002	-0.005	0.003	-0.004	0.002	-0.003	0.002
	π_*^+	0.023	-0.017	0.011	-0.021	0.010	-0.012	0.006	-0.012	0.005	-0.009
	π_*^-	0.026	-0.019	0.013	-0.021	0.007	-0.017	0.005	-0.012	0.004	-0.010

Table A.5: FINITE SAMPLE BIAS. This table reports the finite sample bias associated with our two-step estimation procedure, both in the case where OLS is used in the first step and in the case where FM-OLS is used in the first step. In all cases, OLS is used in the second step. The data is generated as follows: $\Delta y_t = -(2/3)u_{t-1} + \varphi_* \Delta y_{t-1} + \Delta x_t^+ + (1/2)\Delta x_t^- + e_t$, where $u_t := y_t - 2x_t^+ - x_t^-$, $\Delta x_t = 0.5\Delta x_{t-1} + \sqrt{1-0.5^2}v_t$, and $(e_t, v_t)' \sim \text{IID}N(\mathbf{0}_2, \mathbf{I}_2)$. The simulation results are obtained using $R = 5,000$ replications.

	Sample Size	50		100		150		200		250	
	First Step	OLS	FM-OLS	OLS	FM-OLS	OLS	FM-OLS	OLS	FM-OLS	OLS	FM-OLS
φ_*	Second Step	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
-0.50	β_*^+	0.104	0.057	0.029	0.009	0.013	0.003	0.007	0.001	0.005	0.001
	β_*^-	0.120	0.129	0.029	0.010	0.014	0.003	0.007	0.001	0.005	0.001
	ρ_*	0.026	0.024	0.006	0.006	0.003	0.003	0.002	0.002	0.002	0.002
	φ_*	0.020	0.013	0.006	0.004	0.003	0.002	0.002	0.001	0.002	0.001
	π_*^+	0.153	0.134	0.062	0.053	0.037	0.032	0.025	0.022	0.019	0.017
	π_*^-	0.180	0.152	0.062	0.053	0.038	0.032	0.026	0.022	0.020	0.017
-0.25	β_*^+	0.060	0.034	0.016	0.006	0.007	0.002	0.004	0.001	0.003	0.001
	β_*^-	0.068	0.096	0.016	0.007	0.007	0.002	0.004	0.001	0.003	0.001
	ρ_*	0.023	0.022	0.007	0.007	0.004	0.004	0.003	0.003	0.002	0.002
	φ_*	0.016	0.011	0.007	0.004	0.004	0.003	0.003	0.002	0.002	0.002
	π_*^+	0.136	0.125	0.055	0.051	0.034	0.031	0.023	0.021	0.018	0.017
	π_*^-	0.140	0.130	0.056	0.050	0.033	0.030	0.023	0.021	0.018	0.017
0.00	β_*^+	0.032	0.023	0.007	0.004	0.003	0.001	0.002	0.001	0.001	0.000
	β_*^-	0.036	0.045	0.007	0.005	0.003	0.001	0.002	0.001	0.001	0.000
	ρ_*	0.022	0.021	0.008	0.008	0.005	0.005	0.003	0.003	0.003	0.003
	φ_*	0.011	0.010	0.005	0.004	0.003	0.003	0.002	0.002	0.002	0.002
	π_*^+	0.126	0.121	0.050	0.048	0.031	0.030	0.022	0.022	0.018	0.017
	π_*^-	0.128	0.126	0.050	0.048	0.030	0.029	0.023	0.022	0.018	0.017
0.25	β_*^+	0.015	0.018	0.002	0.003	0.001	0.001	0.001	0.001	0.000	0.000
	β_*^-	0.015	0.034	0.003	0.004	0.001	0.001	0.001	0.001	0.000	0.000
	ρ_*	0.019	0.019	0.007	0.007	0.004	0.004	0.003	0.003	0.002	0.002
	φ_*	0.007	0.009	0.004	0.004	0.002	0.002	0.002	0.002	0.001	0.001
	π_*^+	0.116	0.117	0.047	0.046	0.030	0.029	0.021	0.021	0.017	0.017
	π_*^-	0.112	0.114	0.047	0.047	0.030	0.030	0.022	0.022	0.017	0.017
0.50	β_*^+	0.022	0.020	0.004	0.004	0.002	0.002	0.001	0.001	0.001	0.001
	β_*^-	0.022	0.033	0.004	0.007	0.002	0.002	0.001	0.001	0.001	0.001
	ρ_*	0.011	0.011	0.005	0.004	0.003	0.003	0.002	0.002	0.002	0.002
	φ_*	0.007	0.007	0.003	0.003	0.002	0.002	0.001	0.001	0.001	0.001
	π_*^+	0.116	0.115	0.045	0.046	0.030	0.030	0.021	0.021	0.017	0.017
	π_*^-	0.113	0.111	0.046	0.047	0.029	0.029	0.021	0.021	0.017	0.017

Table A.6: FINITE SAMPLE MEAN SQUARED ERROR (MSE) OF THE ESTIMATORS. This table reports the finite sample MSE associated with our two-step estimation procedure, both in the case where OLS is used in the first step and in the case where FM-OLS is used in the first step. In all cases, OLS is used in the second step. The data is generated as follows: $\Delta y_t = -(2/3)u_{t-1} + \varphi_* \Delta y_{t-1} + \Delta x_t^+ + (1/2)\Delta x_t^- + e_t$, where $u_t := y_t - 2x_t^+ - x_t^-$, $\Delta x_t = 0.5\Delta x_{t-1} + \sqrt{1 - 0.5^2} v_t$, and $(e_t, v_t)' \sim$ IID$N(\mathbf{0}_2, \mathbf{I}_2)$. The simulation results are obtained using $R = 5,000$ replications.

	Sample Size	50		100		200		300		400		500	
	First Step	TOLS	FM-TOLS	TOLS	FM-TOLS	TOLS	FM-TOLS	TOLS	FM-TOLS	TOLS	FM-TOLS	TOLS	FM-TOLS
φ_*	Second Step	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
-0.20	β_{1*}^+	0.393	0.685	0.199	0.248	0.093	0.069	0.063	0.037	0.047	0.020	0.038	0.012
	β_{2*}^+	-0.329	-0.691	-0.162	-0.264	-0.080	-0.072	-0.053	-0.042	-0.039	-0.024	-0.031	-0.015
	β_{1*}^-	-0.426	-0.790	-0.207	-0.289	-0.100	-0.079	-0.067	-0.043	-0.049	-0.024	-0.040	-0.015
	β_{2*}^-	0.549	1.025	0.272	0.384	0.136	0.106	0.090	0.059	0.066	0.033	0.052	0.019
	ρ_*	-0.024	0.009	-0.007	-0.023	-0.002	-0.029	-0.001	-0.019	-0.001	-0.015	0.000	-0.011
	φ_*	0.018	0.053	0.007	0.044	0.002	0.025	0.001	0.015	0.001	0.010	0.001	0.007
	π_{1*}^+	0.009	0.214	0.003	0.078	0.000	0.025	0.000	0.011	0.000	0.007	0.000	0.004
	π_{2*}^+	-0.016	-0.288	-0.006	-0.113	-0.002	-0.027	-0.001	-0.014	-0.001	-0.007	-0.001	-0.002
	π_{1*}^-	-0.005	-0.233	0.001	-0.075	0.001	-0.019	0.000	-0.008	0.000	-0.006	0.000	-0.002
	π_{2*}^-	0.000	0.240	-0.003	0.085	-0.002	0.024	-0.001	0.014	-0.001	0.008	-0.001	0.004
-0.10	β_{1*}^+	0.390	0.694	0.190	0.259	0.090	0.075	0.059	0.040	0.044	0.021	0.034	0.015
	β_{2*}^+	-0.318	-0.683	-0.153	-0.265	-0.074	-0.078	-0.048	-0.043	-0.037	-0.024	-0.029	-0.016
	β_{1*}^-	-0.414	-0.779	-0.198	-0.293	-0.098	-0.086	-0.062	-0.047	-0.047	-0.025	-0.037	-0.018
	β_{2*}^-	0.522	0.999	0.252	0.384	0.121	0.111	0.080	0.062	0.059	0.034	0.047	0.022
	ρ_*	-0.158	0.019	-0.063	-0.016	-0.026	-0.026	-0.016	-0.017	-0.012	-0.013	-0.009	-0.011
	φ_*	0.124	0.050	0.068	0.042	0.034	0.024	0.024	0.015	0.017	0.010	0.013	0.007
	π_{1*}^+	0.073	0.227	0.017	0.079	0.003	0.024	0.001	0.013	0.003	0.008	-0.001	0.006
	π_{2*}^+	-0.109	-0.280	-0.048	-0.111	-0.021	-0.031	-0.010	-0.012	-0.010	-0.009	-0.008	-0.003
	π_{1*}^-	-0.030	-0.222	0.005	-0.071	0.004	-0.022	0.010	-0.009	0.004	-0.008	0.007	-0.005
	π_{2*}^-	-0.017	0.233	-0.036	0.088	-0.032	0.033	-0.026	0.016	-0.018	0.011	-0.015	0.008
0.00	β_{1*}^+	0.379	0.698	0.173	0.264	0.083	0.078	0.053	0.042	0.040	0.026	0.031	0.016
	β_{2*}^+	-0.317	-0.686	-0.146	-0.267	-0.071	-0.081	-0.047	-0.045	-0.034	-0.026	-0.028	-0.017
	β_{1*}^-	-0.409	-0.811	-0.186	-0.310	-0.092	-0.090	-0.058	-0.049	-0.044	-0.030	-0.034	-0.019
	β_{2*}^-	0.501	1.014	0.232	0.391	0.111	0.116	0.075	0.065	0.055	0.038	0.044	0.025
	ρ_*	-0.141	0.033	-0.057	-0.007	-0.025	-0.021	-0.014	-0.015	-0.011	-0.013	-0.008	-0.010
	φ_*	0.112	0.044	0.061	0.040	0.032	0.023	0.021	0.015	0.017	0.010	0.013	0.007
	π_{1*}^+	0.084	0.237	0.015	0.089	0.003	0.027	0.002	0.016	0.001	0.011	-0.002	0.006
	π_{2*}^+	-0.106	-0.287	-0.043	-0.118	-0.023	-0.029	-0.014	-0.014	-0.009	-0.012	-0.007	-0.007
	π_{1*}^-	-0.041	-0.239	0.011	-0.086	0.008	-0.025	0.010	-0.013	0.007	-0.009	0.007	-0.008
	π_{2*}^-	-0.027	0.226	-0.045	0.094	-0.033	0.037	-0.023	0.019	-0.019	0.016	-0.015	0.009
0.10	β_{1*}^+	0.345	0.729	0.169	0.272	0.077	0.083	0.049	0.048	0.036	0.030	0.029	0.019
	β_{2*}^+	-0.306	-0.706	-0.148	-0.291	-0.067	-0.085	-0.042	-0.049	-0.032	-0.030	-0.026	-0.020
	β_{1*}^-	-0.389	-0.829	-0.182	-0.311	-0.086	-0.095	-0.053	-0.054	-0.040	-0.034	-0.033	-0.022
	β_{2*}^-	0.476	1.059	0.223	0.411	0.101	0.122	0.066	0.071	0.049	0.043	0.039	0.029
	ρ_*	-0.125	0.051	-0.050	0.005	-0.020	-0.018	-0.014	-0.012	-0.010	-0.011	-0.007	-0.009
	φ_*	0.101	0.035	0.055	0.038	0.029	0.022	0.021	0.013	0.015	0.010	0.012	0.007
	π_{1*}^+	0.056	0.258	0.018	0.092	0.004	0.036	0.001	0.019	-0.001	0.014	-0.001	0.008
	π_{2*}^+	-0.105	-0.294	-0.042	-0.129	-0.019	-0.035	-0.008	-0.020	-0.009	-0.014	-0.007	-0.007
	π_{1*}^-	-0.022	-0.239	0.013	-0.081	0.010	-0.029	0.012	-0.014	0.009	-0.015	0.006	-0.008
	π_{2*}^-	-0.027	0.265	-0.041	0.100	-0.033	0.042	-0.022	0.026	-0.019	0.018	-0.014	0.011
0.20	β_{1*}^+	0.338	0.747	0.150	0.283	0.069	0.089	0.044	0.052	0.033	0.031	0.027	0.022
	β_{2*}^+	-0.313	-0.754	-0.138	-0.282	-0.063	-0.091	-0.040	-0.051	-0.031	-0.031	-0.025	-0.021
	β_{1*}^-	-0.373	-0.850	-0.174	-0.323	-0.077	-0.101	-0.050	-0.059	-0.037	-0.035	-0.030	-0.024
	β_{2*}^-	0.461	1.086	0.203	0.423	0.091	0.131	0.059	0.075	0.044	0.046	0.034	0.032
	ρ_*	-0.110	0.068	-0.045	0.011	-0.019	-0.013	-0.012	-0.011	-0.009	-0.010	-0.006	-0.008
	φ_*	0.086	0.026	0.049	0.033	0.026	0.020	0.018	0.013	0.014	0.008	0.011	0.006
	π_{1*}^+	0.061	0.263	0.016	0.108	-0.003	0.033	-0.002	0.025	0.001	0.015	0.001	0.012
	π_{2*}^+	-0.112	-0.327	-0.043	-0.119	-0.022	-0.042	-0.010	-0.020	-0.007	-0.013	-0.008	-0.011
	π_{1*}^-	-0.014	-0.243	0.007	-0.094	0.016	-0.031	0.011	-0.023	0.008	-0.015	0.007	-0.010
	π_{2*}^-	-0.025	0.277	-0.043	0.108	-0.035	0.046	-0.025	0.028	-0.022	0.020	-0.016	0.014

Table A.7: FINITE SAMPLE BIAS. This table reports the finite sample bias associated with our two-step estimation procedure, both in the case where the TOLS is used in the first step and in the case where the FM-TOLS is used in the first step. In all cases, OLS is used in the second step. The data is generated as follows: $\Delta y_t = -u_{t-1} + \varphi_* \Delta y_{t-1} + \boldsymbol{\pi}_{0*}^{+\prime} \Delta \boldsymbol{x}_t^+ + \boldsymbol{\pi}_{0*}^{-\prime} \Delta \boldsymbol{x}_t^- + e_t$, where $u_t := y_t - \boldsymbol{\beta}_*^{+\prime} \boldsymbol{x}_t^+ - \boldsymbol{\beta}_*^{-\prime} \boldsymbol{x}_t^-$, $\Delta \boldsymbol{x}_t = 0.5 \Delta \boldsymbol{x}_{t-1} + \sqrt{1 - 0.5^2} \boldsymbol{v}_t$, and $(e_t, \boldsymbol{v}_t')' \sim \text{IID}N(\mathbf{0}_3, \mathbf{I}_3)$. The simulation results are obtained using $R = 5,000$ replications.

	Sample Size	50		100		200		300		400		500	
	First Step	TOLS	FM-TOLS	TOLS	FM-TOLS	TOLS	FM-TOLS	TOLS	FM-TOLS	TOLS	FM-TOLS	TOLS	FM-TOLS
φ_*	Second Step	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
-0.20	β_{1*}^+	0.427	1.201	0.105	0.193	0.024	0.019	0.011	0.006	0.006	0.002	0.004	0.001
	β_{2*}^+	0.397	1.254	0.091	0.197	0.022	0.020	0.010	0.007	0.005	0.003	0.003	0.001
	β_{1*}^-	0.471	1.436	0.105	0.219	0.026	0.022	0.011	0.007	0.006	0.003	0.004	0.001
	β_{2*}^-	0.537	1.865	0.126	0.293	0.031	0.028	0.014	0.009	0.008	0.003	0.005	0.002
	ρ_*	0.001	0.047	0.000	0.011	0.000	0.003	0.000	0.002	0.000	0.001	0.000	0.001
	φ_*	0.001	0.020	0.000	0.006	0.000	0.002	0.000	0.001	0.000	0.001	0.000	0.001
	π_{1*}^+	0.008	0.547	0.001	0.133	0.000	0.031	0.000	0.017	0.000	0.011	0.000	0.009
	π_{2*}^+	0.008	0.614	0.001	0.135	0.000	0.033	0.000	0.017	0.000	0.011	0.000	0.008
	π_{1*}^-	0.008	0.587	0.001	0.128	0.000	0.031	0.000	0.017	0.000	0.011	0.000	0.008
	π_{2*}^-	0.007	0.511	0.001	0.122	0.000	0.032	0.000	0.018	0.000	0.011	0.000	0.009
-0.10	β_{1*}^+	0.445	1.257	0.098	0.192	0.023	0.019	0.010	0.006	0.005	0.002	0.003	0.001
	β_{2*}^+	0.378	1.234	0.087	0.192	0.020	0.019	0.008	0.007	0.005	0.002	0.003	0.001
	β_{1*}^-	0.438	1.438	0.098	0.224	0.024	0.022	0.010	0.008	0.006	0.003	0.003	0.001
	β_{2*}^-	0.509	1.812	0.114	0.289	0.027	0.028	0.011	0.010	0.006	0.003	0.004	0.002
	ρ_*	0.043	0.041	0.010	0.010	0.003	0.003	0.002	0.002	0.001	0.001	0.001	0.001
	φ_*	0.024	0.018	0.008	0.006	0.003	0.002	0.001	0.001	0.001	0.001	0.001	0.001
	π_{1*}^+	0.382	0.549	0.102	0.130	0.036	0.031	0.020	0.017	0.014	0.011	0.010	0.009
	π_{2*}^+	0.378	0.569	0.111	0.133	0.035	0.032	0.020	0.017	0.014	0.011	0.010	0.009
	π_{1*}^-	0.355	0.565	0.102	0.126	0.034	0.031	0.020	0.017	0.013	0.011	0.010	0.009
	π_{2*}^-	0.341	0.507	0.105	0.126	0.037	0.032	0.021	0.018	0.014	0.012	0.011	0.009
0.00	β_{1*}^+	0.429	1.254	0.088	0.190	0.020	0.020	0.008	0.006	0.005	0.003	0.003	0.001
	β_{2*}^+	0.371	1.350	0.077	0.196	0.019	0.020	0.008	0.007	0.004	0.003	0.003	0.001
	β_{1*}^-	0.440	1.516	0.093	0.235	0.021	0.024	0.009	0.008	0.005	0.003	0.003	0.001
	β_{2*}^-	0.476	1.933	0.099	0.305	0.023	0.030	0.010	0.010	0.006	0.004	0.004	0.002
	ρ_*	0.037	0.042	0.009	0.009	0.003	0.003	0.001	0.002	0.001	0.001	0.001	0.001
	φ_*	0.021	0.018	0.007	0.006	0.002	0.002	0.001	0.001	0.001	0.001	0.001	0.001
	π_{1*}^+	0.365	0.549	0.099	0.127	0.034	0.030	0.019	0.017	0.013	0.012	0.010	0.009
	π_{2*}^+	0.352	0.609	0.096	0.132	0.033	0.031	0.019	0.017	0.013	0.012	0.009	0.009
	π_{1*}^-	0.343	0.559	0.102	0.124	0.033	0.030	0.020	0.017	0.013	0.012	0.010	0.009
	π_{2*}^-	0.332	0.505	0.098	0.122	0.035	0.031	0.020	0.017	0.014	0.012	0.010	0.009
0.10	β_{1*}^+	0.397	1.286	0.084	0.209	0.017	0.021	0.007	0.007	0.004	0.003	0.003	0.001
	β_{2*}^+	0.372	1.294	0.078	0.219	0.017	0.022	0.007	0.007	0.004	0.003	0.002	0.001
	β_{1*}^-	0.416	1.551	0.087	0.238	0.019	0.026	0.008	0.008	0.004	0.003	0.003	0.002
	β_{2*}^-	0.461	2.088	0.097	0.323	0.020	0.032	0.009	0.011	0.005	0.004	0.003	0.002
	ρ_*	0.030	0.041	0.007	0.009	0.002	0.002	0.001	0.001	0.001	0.001	0.001	0.001
	φ_*	0.018	0.016	0.006	0.006	0.002	0.002	0.001	0.001	0.001	0.001	0.001	0.001
	π_{1*}^+	0.340	0.539	0.095	0.131	0.032	0.031	0.018	0.017	0.013	0.012	0.010	0.009
	π_{2*}^+	0.336	0.567	0.091	0.143	0.032	0.031	0.018	0.017	0.013	0.012	0.010	0.009
	π_{1*}^-	0.317	0.553	0.091	0.126	0.032	0.032	0.020	0.018	0.013	0.011	0.010	0.009
	π_{2*}^-	0.318	0.532	0.095	0.126	0.032	0.032	0.020	0.018	0.014	0.012	0.011	0.009
0.20	β_{1*}^+	0.388	1.444	0.074	0.210	0.016	0.023	0.006	0.008	0.004	0.003	0.002	0.001
	β_{2*}^+	0.392	1.475	0.072	0.209	0.015	0.022	0.006	0.008	0.004	0.003	0.002	0.001
	β_{1*}^-	0.412	1.707	0.080	0.257	0.017	0.027	0.007	0.009	0.004	0.003	0.002	0.002
	β_{2*}^-	0.460	2.209	0.091	0.342	0.018	0.034	0.007	0.012	0.004	0.004	0.003	0.002
	ρ_*	0.025	0.040	0.006	0.009	0.002	0.002	0.001	0.001	0.001	0.001	0.001	0.001
	φ_*	0.015	0.015	0.005	0.005	0.002	0.002	0.001	0.001	0.001	0.001	0.001	0.001
	π_{1*}^+	0.320	0.555	0.086	0.136	0.030	0.032	0.018	0.018	0.012	0.011	0.009	0.009
	π_{2*}^+	0.336	0.613	0.090	0.137	0.032	0.033	0.018	0.017	0.013	0.011	0.010	0.009
	π_{1*}^-	0.323	0.561	0.089	0.136	0.031	0.031	0.017	0.017	0.012	0.012	0.010	0.009
	π_{2*}^-	0.307	0.532	0.092	0.125	0.033	0.033	0.019	0.018	0.013	0.012	0.010	0.009

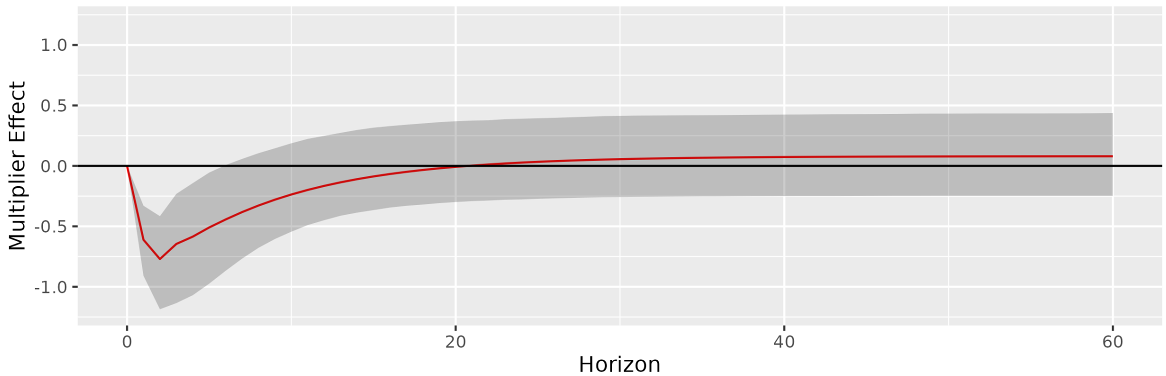
Table A.8: FINITE SAMPLE MEAN SQUARED ERROR (MSE) OF THE ESTIMATORS. This table reports the finite sample MSE associated with our two-step estimation procedure, both in the case where the TOLS is used in the first step and in the case where the FM-TOLS is used in the first step. In all cases, OLS is used in the second step. The data is generated as follows: $\Delta y_t = -u_{t-1} + \varphi_* \Delta y_{t-1} + \boldsymbol{\pi}_{0*}^{+\prime} \Delta \boldsymbol{x}_t^+ + \boldsymbol{\pi}_{0*}^{-\prime} \Delta \boldsymbol{x}_t^- + e_t$, where $u_t := y_t - \boldsymbol{\beta}_*^{+\prime} \boldsymbol{x}_t^+ - \boldsymbol{\beta}_*^{-\prime} \boldsymbol{x}_t^-$, $\Delta \boldsymbol{x}_t = 0.5 \Delta \boldsymbol{x}_{t-1} + \sqrt{1 - 0.5^2} \boldsymbol{v}_t$, and $(e_t, \boldsymbol{v}_t')' \sim \text{IID}N(\mathbf{0}_3, \mathbf{I}_3)$. The simulation results are obtained using $R = 5,000$ replications.

	$\Delta \log$ GPDI	Δ(R&D/GDP)
Mean	0.009	0.004
Median	0.009	0.001
Maximum	0.110	0.082
Minimum	-0.176	-0.082
Standard Deviation	0.039	0.028
Skewness	-0.752	0.190
Excess Kurtosis	2.701	0.256
Sample Size	240	240

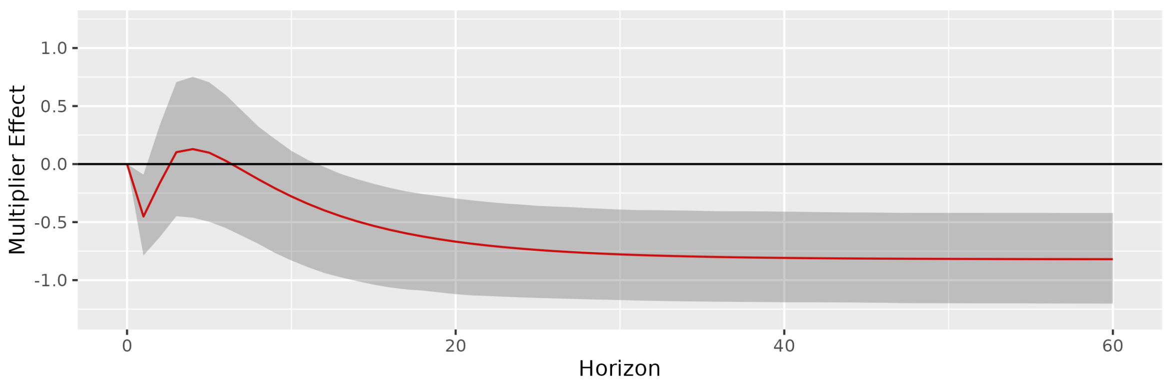
Table A.9: DESCRIPTIVE STATISTICS. Descriptive statistics are computed over 240 quarters from 1960q1 to 2019q4. GPDI is measured in US Dollars at 2012 prices and seasonally adjusted. R&D and GDP are seasonally adjusted nominal values.

PP test	$\log$ GPDI	R&D/GDP
PP test w/o trend	-1.138	-1.946
p-value	0.701	0.311
PP test w/ trend	-3.198	-2.255
p-value	0.087	0.457

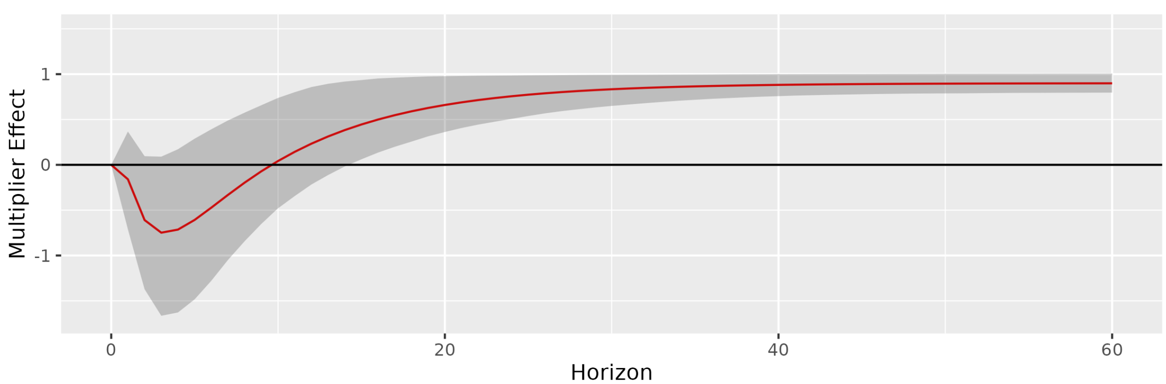
Table A.10: PHILLIPS AND PERRON'S (1988) UNIT-ROOT TEST STATISTICS. Two Phillips and Perron test statistics are computed, one including and the other excluding a time trend. When the time trend is included, it is statistically significant for the log of GPDI but not for R&D intensity. The lag lengths are selected by the SIC when estimating the ADF equations.



(a) Cumulative response of i_{t+h} to a +1 unit shock to r_t^+ in period 1



(b) Cumulative response of i_{t+h} to a +1 unit shock to r_t^- in period 1



(c) Asymmetry across horizons

Figure A.1: CUMULATIVE DYNAMIC MULTIPLIERS (SINGLE-STEP). Panels (a) and (b) present the cumulative dynamic multiplier effects with respect to unit shocks to r_t^+ and r_t^-, respectively, occurring in period 1. Panel (c) shows the asymmetry at each horizon, i.e., the value of the cumulative dynamic multiplier effect in panel (b) subtracted from the corresponding value in panel (a). Empirical 95% confidence intervals obtained from 5,000 replications of a moving block bootstrap procedure are reported throughout.