The Coherence Side of Rationality: Theory and Evidence from Firm Plans

Finance Working Paper N° 1106/2025
May 2026

Pamela Giustinelli
University of Padova

Stefano Rossi
Bocconi University, CEPR and ECGI

ECGI Working Paper Series in Finance

The Coherence Side of Rationality: Theory and Evidence from Firm Plans

Working Paper N° 1106/2025
May 2026

Pamela Giustinelli
Stefano Rossi

We acknowledge the financial support of the Baffi-Carefin Center at Bocconi University. We are grateful to Alexei Verkhovtsev and Nazrul Shafiq bin Shahril Effendi for outstanding research assistance and to Victor Aguirregabiria, Oriana Bandiera, Nick Barberis, Manudeep Bhuller, Alon Brav, Francesco Corielli, Max Croce, Francesco Decarolis, Alexia Delfino, Xavier Gabaix, Fabian Gouret, John Graham, Marius Guenzel, Max Kellogg, Deborah Kim, Edwin Leuven, Chen Lian, Lance Lochner, Francesca Molinari, Sean Myers, Andras Niedermayer, Marco Ottaviani, Nicola Pavoni, Vincenzo Pezone, Collin Raymond, David Rivers, Matthias Rodemeier, Julien Sauvagnat, Matthew Shapiro, Kelly Shue, Todd Stinebrickner, Johannes Stroebel, Mirko Wiederholt, Wei Xiong, conference participants at the NBER Behavioral Finance Meetings, the Miami Behavioral Finance Conference, the 7th Workshop on Subjective Expectations, the 2nd Northwestern University Former Econometrics Students Conference, the HKU/TLV Finance Forum, the 2023 Western Finance Association Conference, the 2023 International Association of Applied Econometrics Annual Conference, the 2023 Summer Meeting of the European Economic Association, the 2024 CESifo Summer Institute on Expectation Formation, the CSEF-IGIER Symposium on Economics and Institutions, the 5th Joint BoC-ECB-NY FED Conference on Expectations Surveys, the 2025 World Congress of the Econometric Society, and seminar participants at Bocconi University, IWH Halle, Munich TUM, University of Western Ontario, Paris Dauphine University, University of Bristol, University of Oslo, University of Queensland, CY Cergy Paris University, Birkbeck University of London, University of Padova, and Norges Bank for very helpful comments.

Abstract

Using the Duke Survey data on internal firm plans, we show that over 80% of Chief Financial Officer (CFO) forecasts align closely with simple heuristics taught in MBA textbooks. We introduce forecast coherence--internal consistency across forecasts of jointly determined variables--as a benchmark for evaluating these heuristics. Nearly half of CFOs issue forecasts closely aligned with incoherent heuristics. Such forecasts are associated with predictable reversals in forecast errors, lower firm performance, and underinvestment. We develop a parsimonious model illustrating how reliance on restricted forecasting heuristics can generate incoherence and resource misallocation within firms, consistent with our empirical evidence.

Keywords: Coherence, Rules of Thumb, Narrow Bracketing, Firm Expectations

JEL Classifications: D84, D22, L2, M2, G32

Pamela Giustinelli
Assistant Professor
Unoversity of Padova
Palazzo Levi Cases
via del Santo 33, 35123, Padova, Italy
e-mail: pamela.giustinelli@unipd.it

Stefano Rossi*
Professor of Finance
Bocconi University
Via Röntgen 1
20136 Milano, Italy
e-mail: rossi.stefano@unibocconi.it

*Corresponding Author

The Coherence Side of Rationality

Theory and evidence from firm plans *

Pamela Giustinelli† and Stefano Rossi‡

This version: May 2026

Abstract

Using the Duke Survey data on internal firm plans, we show that over 80% of Chief Financial Officer (CFO) forecasts align closely with simple heuristics taught in MBA textbooks. We introduce forecast coherence—internal consistency across forecasts of jointly determined variables—as a benchmark for evaluating these heuristics. Nearly half of CFOs issue forecasts closely aligned with incoherent heuristics. Such forecasts are associated with predictable reversals in forecast errors, lower firm performance, and underinvestment. We develop a parsimonious model illustrating how reliance on restricted forecasting heuristics can generate incoherence and resource misallocation within firms, consistent with our empirical evidence.

JEL classification: D84, D22, L2, M2, G32.

Keywords: Firm Expectations, Coherence, Rules of Thumb, Bracketing.

*We acknowledge the financial support of the Baffi-Carefin Center at Bocconi University. We are grateful to Alexei Verkhovtsev and Nazrul Shafiq bin Shahril Effendi for outstanding research assistance and to Victor Aguirregabiria, Oriana Bandiera, Nick Barberis, Manudeep Bhuller, Alon Brav, Francesco Corielli, Max Croce, Francesco Decarolis, Alexia Delfino, Xavier Gabaix, Fabian Gouret, John Graham, Marius Guenzel, Max Kellogg, Deborah Kim, Edwin Leuven, Chen Lian, Lance Lochner, Francesca Molinari, Sean Myers, Andras Niedermayer, Marco Ottaviani, Nicola Pavoni, Vincenzo Pezone, Collin Raymond, David Rivers, Matthias Rodemeier, Julien Sauvagnat, Matthew Shapiro, Kelly Shue, Todd Stinebrickner, Johannes Stroebel, Mirko Wiederholt, Wei Xiong, conference participants at the NBER Behavioral Finance Meetings, the Miami Behavioral Finance Conference, the 7th Workshop on Subjective Expectations, the 2nd Northwestern University Former Econometrics Students Conference, the HKU/TLV Finance Forum, the 2023 Western Finance Association Conference, the 2023 International Association of Applied Econometrics Annual Conference, the 2023 Summer Meeting of the European Economic Association, the 2024 CESifo Summer Institute on Expectation Formation, the CSEF-IGIER Symposium on Economics and Institutions, the 5th Joint BoC-ECB-NY FED Conference on Expectations Surveys, the 2025 World Congress of the Econometric Society, and seminar participants at Bocconi University, IWH Halle, Munich TUM, University of Western Ontario, Paris Dauphine University, University of Bristol, University of Oslo, University of Queensland, CY Cergy Paris University, Birkbeck University of London, University of Padova, and Norges Bank for very helpful comments.

†University of Padova; CESifo; IZA@LISER; email: pamela.giustinelli@unipd.it.

‡Bocconi, IGIER and Baffi-CAREFIN; CEPR and ECGI; email: stefano.rossi@unibocconi.it.

Coherence—or "the consistency of the elements of the person's judgment" (Hammond (2007), p. xvi)—is one of the two standards of rationality together with accuracy.[1] In the context of multidimensional forecasting, coherence requires forecasts of individual variables to incorporate the connections among those variables; whereas accuracy requires individual forecasts not to differ systematically from realizations ex post. While forecast accuracy has been widely studied—see, e.g., Tversky and Kahneman (1974) and Benjamin (2019) for a recent review—forecast coherence has received limited attention.

We study forecast coherence in firm plans, internal forecasts over multiple balance-sheet variables, outlining all resource allocation within the firm. When preparing such plans, Chief Financial Officers (CFOs) start from a given output target ('top line', or sales, forecast), and then proceed to forecast all other items (e.g., input costs).

Internal planning is a challenging multidimensional forecasting problem in a high-stakes setting, and is still not well understood (Graham, 2022). MBA textbooks and case studies offer a series of heuristics or 'rules of thumb' (RoTs) to help CFOs make internal forecasts (e.g., Ruback (2004), Welch (2017), Koller et al. (2020)). These rules range from forecasting each item on its own, to anchoring each forecast to the output target, to multivariate rules, thus varying in their concern for coherence and degree of sophistication. To the best of our knowledge, these rules have not previously been assessed. In this paper we provide a first assessment in terms of forecast coherence and accuracy.

We consider the RoTs as defined in Welch (2017)'s taxonomy: (R1) plain growth, (R2) proportion of sales, (R3) economies of scale, (R4) industry based, and (R5) disaggregated. Conceptually, rules (R2)-(R5) reflect some concern for coherence, as they link the forecast of each input's growth to that of output growth ((R2)-(R4)), and also of other inputs' growth (R5); whereas (R1) prescribes forecasting each input's growth by projecting that input's past growth into the future, implicitly disregarding the input's relationship with the output and other inputs. Operationally, rules (R1)-(R4) are straightforward, involving a mere sample average (R1) or univariate mean linear regressions ((R2)-(R4)); whereas (R5) is less restricted, using multivariate regression.

[1] Coherence can be traced back at least to Aristotle (Fogelin, 2003). Tversky and Kahneman (1981), p. 453, write, "there is general agreement that rational choices should satisfy some elementary requirements of consistency and coherence." See also Sen (1993), Becker (1996), Posner (2014).

Using balance-sheet and income-statement data from Compustat, we start by documenting that these RoTs give rise to very different forecasts from one another. Next, we assess firm plans by exploiting unique data on simultaneous forecasts of multiple balance-sheet items from the Duke Survey of CFOs of large- and mid-size US corporations (Ben-David et al. (2013), Graham (2022)), and linking it to realizations from Compustat.

We find that many CFOs make forecasts that align closely with textbook RoTs and that these rules are not equally good. Some RoTs are associated with more predictable forecast errors, and can be naturally interpreted as more incoherent because they fail to account fully for structural relationships across variables. Firms whose CFO forecasts align more closely with these poorer rules display lower performance and lower investment. These findings are empirical and do not depend on any specific model of belief formation.

We then develop a simple model to organize and interpret our findings. The model provides a benchmark notion of forecast coherence and generates empirical implications for how different RoTs can generate systematically predictable forecast errors and real outcomes, in line with our empirical evidence.

We make three main contributions. First, we document the extent to which (R1)-(R5) describe the forecasts CFOs actually make. We find more than 80% of CFOs provide forecasts closely aligned with (R1)-(R5). The implication is that for a large majority of CFOs it is possible to view their forecasts *as if* they were generated by one of the RoTs, with small discrepancies due to rounding, or small differences in implementation or data.

These findings stop short of demonstrating that CFOs have necessarily used any of the RoTs in Welch (2017). The reason is that our data do not directly elicit the CFOs' forecasting rules, so CFOs could be using some other forecasting process, which turns out to yield forecasts closely aligned with those implied by the RoTs. At the same time, our empirical results suggest that, whether or not CFOs actually use these RoTs, it is useful to think of firm plans in terms of the RoTs in Welch (2017), because those RoTs give rise to forecasts that are significantly different from one another, and because so many CFOs provide forecasts that align closely with those implied by one of the RoTs.

Our second contribution is to document that forecasts in line with different RoTs covary systematically with firm outcomes in the cross section. Firms whose CFO forecasts are closest to those implied by (R1) and (R2) have lower firm performance as measured by return on assets (ROA) than those whose forecasts are closest to a multivariate rule. These findings are confirmed using an event-study approach around the dates of CFO appointments. Firms with these same CFO forecasts are also associated with predictable reversals in forecast errors and lower capital expenditures (hence CapEx). Albeit descriptive, this evidence establishes the relevance of the RoTs and suggests the importance of forecast (in)coherence.

Our third contribution is to develop a theory of optimal multidimensional forecasting in firm plans, nesting the five RoTs we study empirically in the first part. The model illustrates formally how different forecasting rules can generate varying degrees of (in)coherence, yielding distinct empirical implications for firm outcomes. In the model, the use of suboptimal forecasting rules provides a novel mechanism generating incoherent plans and, as a result, resource misallocation within the firm.

In our theory, a CFO is given a sales growth target by the CEO and the board of directors. The CFO's task is to minimize total costs while meeting the target. The CFO issues a vector of forecasts over growth rates of firm variables, given knowledge of the firm's technology, which is assumed stable over time and not subject to aggregate shocks. Forecasts are made by minimizing expected inaccuracy under square loss, subject to a coherence constraint provided by the firm's technology. In our benchmark with complete information, we establish that the elements of the optimal forecast vector are coherently linked by the firm's technology parameters, rationalizing a specific version of (R5) as consistent with a first-best benchmark, unlike more restricted RoTs. Our model implies that more restricted RoTs such as (R1) and (R2) are likely to yield lower CapEx forecasts than unrestricted (R5), explaining our evidence of a systematic reversal in forecast errors. In turn, in a world with ex post adjustment costs these lower CapEx forecasts directly translate into underinvestment, consistent with our evidence.

In fact, (R1) is reminiscent of "narrow bracketing" (e.g., Thaler (1985), Read et al. (1999)), (R5) of "broad bracketing," and (R2)-(R4) appear to lie in between these two extremes. "Narrow bracketing" refers to the behavior of decision makers who, when confronted with multiple related choices (e.g., consumption bundles), make each choice in isolation instead of jointly ("broad bracketing"), thus disregarding how budget constraint and utility function tie choice variables together and obtaining lower utility.

Starting from this observation and drawing on related frameworks of sparse cognition (Gabaix, 2014) and narrow thinking (Lian, 2021), we model CFOs' internal plans as an incomplete-information, common-interest game among multiple selves. Each CFO self is in charge of forecasting one balance sheet item, while observing noisy signals about other items. Each CFO self's imperfect knowledge of the other selves' forecasts creates intra-personal frictions in coordinating forecasts of different variables among CFO selves, generating incoherent forecasts. The resulting form of forecast incoherence depends on the signal structure, giving rise to different RoTs.

In our model, "narrow thinking" generates incoherent plans via the use of suboptimal forecasting rules. Incoherent plans, in turn, tend to induce suboptimal resource allocation and, hence, lower profits. In the benchmark case, the multivariate rule (R5) dominates more restricted rules, (R3)-(R4) are intermediate, and (R2)-(R1) perform worse, providing one possible microfoundation for incoherence.

In our data, among the CFOs whose forecasts are closely aligned with (R1)-(R5), 6% are aligned with (R1) and 41% with (R2), with 32% of these exactly equal to (R2). Despite its simplicity and popularity in managerial and consulting textbooks, (R2) is problematic. We show theoretically that under decreasing returns to scale CapEx growth needs to be strictly larger than the sales growth target, implying (R2) is suboptimal. This can also be seen by simply comparing (R2) with (R3) empirically. (R2) can be expressed as a mean regression of each input's growth on the output's growth with a zero intercept and a unit slope, which is why we relabel it 'sales anchoring'. (R3) can be expressed as the same mean regression, but with the intercept and slope actually estimated using Compustat data. The estimated slope is multiplied by the targeted sales growth and added to the estimated intercept. In Compustat, while the slope is close to one, the intercept is positive, large, and significant. Intuitively, the large positive intercept is consistent with replacement and maintenance needs in addition to expansion, implying CapEx growth typically exceeds sales growth. Importantly, this applies to all industries.

(R1) is also problematic, as it dictates to extrapolate past CapEx growth into the future; whereas in the data CapEx is mean reverting, implying also (R1) yields CapEx growth forecasts that are systematically too low relative to planned sales growth.

Our model also allows to distinguish between forecast coherence and forecast accuracy. Accuracy requires the FEs of each variable to be zero on average, whereas coherence requires the output-input relationship among FEs to mirror that for realizations. This latter theoretical prediction is robust to a large set of technological environments and popular investment frameworks, including Q-theory, time-to-build, investment spikes, financial constraints, irreversibility, and uncertainty. In our data, we reject forecast coherence on average and forecast accuracy for CapEx, but not for sales. Distinguishing coherence from accuracy is important, as under some conditions coherence can be assessed ex ante, and ex ante coherence may aid accuracy ex post.

In our data CFOs are reasonably accurate in forecasting output, reflecting their incentives (Graham, 2022), but much less accurate in forecasting CapEx, likely because their forecast of CapEx growth is incoherently linked to that of sales growth.

Our evidence is also hard to reconcile with several alternative interpretations. Incoherence is distinct from other managerial traits such as optimism and overconfidence (e.g., Ben-David et al. (2013) and Guenzel and Malmendier (2020)'s review): controlling for these traits, we continue to find that incoherence is associated with lower investment, whereas optimism and overconfidence predict the opposite. More generally, our findings are difficult to explain solely by differential managerial control over variables or by firms' optimal responses to idiosyncratic shocks, because CFOs are more accurate about sales than about CapEx, the results are relatively stable across industries, and incoherence is systematically associated with lower firm performance.

Studying coherence is challenging, as it requires referring to "the entire web of beliefs held by the individual" (Tversky and Kahneman (1974), p. 1130) and "to something external to choice behavior (such as objectives, values, or norms)" (Sen (1993), p. 495). Our corporate planning setting is naturally amenable to study forecast coherence, for two reasons. First, executives of large mature firms should have a good sense of their firm's technology and agree on the firm's objectives. Second, these executives should be optimizing agents, or at least they should be seeking to achieve certain pre-defined goals.

Incoherence in firm plans provides a possible mechanism for understanding resource allocation within the firm (e.g., see Maksimovic and Phillips (2002, 2013), Hoberg and Phillips (2025)). Prior literature on managerial beliefs has focused on forecast accuracy of individual variables (e.g., Ben-David et al. (2013), Graham (2022)). Prior work on behavioral bracketing has not studied a firm production setting (e.g., Thaler (1985), Barberis et al. (2006), Lian (2021)). Prior behavioral literature proposes different mechanisms to generate predictable FEs (e.g., Coibion and Gorodnichenko (2015), Bordalo et al. (2020)). Prior behavioral corporate finance literature has focused on other psychological traits such as optimism and overconfidence (e.g., Baker et al. (2007), Guenzel and Malmendier (2020)). We show incoherence differs from both optimism and overconfidence and generates different predictions from prior literature, which we validate in the data. We discuss the paper's contributions in greater detail in Section 4.

1. Data

We use two main sources of data, one on CFO expectations and one on firm realizations. CFO expectations come from the Duke Survey, run by John Graham and Campbell Harvey. The study surveys 2,000–3,000 CFOs quarterly, asking their views about the US economy and corporate policies, as well as their expectations of future firm performance and operational plans. The typical response rate is 5% to 8%, with most responses arriving within a couple of days. Since the end of the 1990s, the survey has consistently asked respondents their expectations of the future twelve-month growth of key firm variables, including revenues, capital expenditures, employment, and earnings.[2]

Our data comprises 72 quarterly surveys conducted between March 2001 and December 2018. We observe corporate forecasts as a single number per variable, which we interpret as the CFO's expected value, corresponding to the firm's base case scenario.[3]

Forecasts are elicited for all variables jointly as follows:

Relative to the previous 12 months, what will be your company's PERCENTAGE CHANGE during

[2]Historical surveys and aggregated responses can be accessed at https://cfosurvey.fuqua.duke.edu/.

[3]For many firms, this is the only scenario leading to fleshed-out forecasts in internal planning. Some firms consider also a downside scenario to plan for contingencies and an upside scenario to lay out stretch goals. However, the latter do not usually lead to fleshed-out forecasts (see Graham (2022), p. 1997).

the next 12 months? (e.g., +3%, −2%, etc.) [Leave blank if not applicable] Revenues: _____; Capital spending: _____; R&D spending: _____; Technology spending: _____; Prices of your product: _____; Earnings: _____; Cash on balance sheet: _____; Number of domestic full-time employees: _____; Wage: _____; Dividends: _____. Advertising: _____. Share repurchases: _____.

Figure A1 of the Web Appendix shows a screenshot from the survey. It is worth noting that by eliciting all forecasts jointly, the question's format might aid coherence relative to alternative formats eliciting individual forecasts in separate questions.

Firm realizations come from Compustat, which extracts the information from the Security and Exchange Commission (SEC)-required public filing of financial statements. Compustat covers all publicly traded firms across all sectors of the US economy since 1955. We exclude firms with negative assets and we winsorize at the 1% level.

Table A1 of the Web Appendix shows that, on average, Duke firms are larger in sales and assets, more profitable, and hoard more cash than Compustat firms, but are otherwise similar in terms of market-to-book ratio, investment, and leverage. These patterns concur with prior work using the Duke data (e.g., Ben-David et al. (2013)).

When matching Duke and Compustat by firm ID, we face four sources of attrition: (1) due to privacy restrictions, not all Duke respondents report their firm ID; (2) not all Duke respondents give forecasts about all variables; (3) some of the variables whose forecasts are elicited in the Duke Survey do not have a precise counterpart in Compustat, namely, technology spending, outsourced employees, health spending, productivity, product prices, and share repurchases; and (4) some of the variables with a counterpart in Compustat do not have full coverage. Chiefly, wages are missing for about 90% of Compustat firms. R&D and advertising expenditures are also missing for many Compustat firms.[4]

Table A2 of the Web Appendix reports summary statistics on twelve-month ahead growth forecasts (Panel A), followed by the corresponding realizations (Panel B), in the matched Duke-Compustat sample.[5] Comparing the two panels shows that while CFOs are on average slightly optimistic about sales (output), the empirical medians of forecasts

[4]Points (1) and (2) raise the possibility of non-random selection. However, under the assumption that missing forecasts reflect lack of knowledge, our respondents are likely positively selected from those more likely to provide complete and high-quality answers. Points (3) and (4) imply that our analysis of FEs needs to be limited to items with full coverage in both Duke and Compustat. Finally, the matched Duke-Compustat sample mostly refers to the first part of the sample period, until 2011Q4. This is not a problem for us, as we conduct most of our regression analysis in the pre-financial crisis period.

[5]For some firm-year pairs there might be multiple CFO forecasts, as we match Duke and Compustat by firm ID. Table A3 of the Web Appendix reports the same statistics in the full Compustat population.

and realizations are quite close to each other, consistent with Graham (2022)'s observation that CFOs care about getting sales forecasts right. Conversely, CFOs' forecasts of capital expenditures are systematically too low, with the distribution of forecasts shifted to the left relative to the distribution of realizations.

2. Empirical Analysis

When preparing firm plans, CFOs typically start by issuing forecasts of sales growth for a number of years, reflecting targets set by the CEO and board members, and then proceed to forecast all other balance-sheet variables (Graham, 2022). To help CFOs with this multidimensional forecasting problem, managerial textbooks and case studies propose simple methods, surveyed in the following taxonomy by Welch (2017) (p. 593-594):

(R1) A **plain growth** forecast, separately projecting past growth rates of each item; henceforth 'narrow bracketing' rule. Welch (2017) implements this rule by computing for each item the average of the two most recent annual growth rates.

(R2) A pure **proportion of sales** forecast, forecasting each item as a fixed proportion of sales; henceforth 'sales anchoring' rule. Welch (2017) implements this rule by assigning to each item the same growth rate as that forecasted for sales. Thus, this rule can be expressed as a mean regression of each item's growth on sales' growth with a zero intercept and a unit slope.

(R3) An **economies-of-scale** forecast, positing for each item a fixed component and a variable component, the latter itself a proportion of sales' growth. Welch (2017) implements this rule by estimating a best linear predictor under square loss of each item's growth on contemporaneous sales growth using Compustat data. The estimated regression intercept is the fixed component and the estimated slope multiplied by the forecasted sales growth is the variable component.

(R4) An **industry-based** forecast, drawing on information from other firms in the same industry. Welch (2017) implements this rule exactly as (R3), but using only data on firms in the same industry as the firm under consideration.

(R5) A **disaggregated** forecast, allowing each item to comove not only with sales but also with other items. Welch (2017) implements this rule by using a 'kitchen sink' specification with all Compustat variables.

There is no consensus on which of the above rules, if any, constitutes best practice. In their best-selling guide to valuation, Koller et al. (2020) advocate (R2) writing, "*net Property, Plant and Equipment should be forecast as a percentage of revenues*" (p. 286).[6] Similarly, Ruback (2004) advocates (R2) by stating "*capex should be proportional to sales*" and Harvard Business School case studies provide solutions in which CapEx growth forecasts equal sales growth forecasts, in line with (R2), and also suggest (R1) and (R4) (e.g., Luehrman and Heilprin (2009) and Stafford and Heilprin (2011)). Proponents of other rules include Titman and Martin (2016) for (R3) and Holthausen and Zmijewski (2020) for (R5). This literature lacks a framework to evaluate these methods and establish whether they are equivalent to one another or, if not, which ones perform best.

This section provides such a framework. Subsection 2.1 examines the extent to which (R1)-(R5) describe the forecasts CFOs make. Subsection 2.2 proposes a measure of (in)coherence. Subsection 2.3 presents tests of coherence and accuracy based on forecast errors. Subsection 2.4 examines whether forecasts aligned with RoTs are associated with predictable forecast errors. Subsections 2.5 and 2.6 examine the associations between incoherence (RoTs) and performance and CapEx. Subsection 2.7 presents an event study of changes in incoherence (RoTs) and firm outcomes around CFO appointments.

2.1 *RoTs and their Distance from CFO Forecasts*

We begin in Table 1 with a dual purpose. First, Table 1 describes how to implement specific RoTs following Welch (2017) and using Compustat data; second, Table 1 provides a first assessment of the RoTs, exclusively based on Compustat data.

The first row of Table 1 reports the results of a regression designed to assess (R1), which dictates to forecast future CapEx growth as the average of the past two CapEx

[6] In turn, Koller et al. (2020) indicate that depreciation should be forecasted as a proportion of net Property, Plant, and Equipment (PPE), and that capital expenditures should be calculated by summing the projected increase in net PPE to depreciation (p. 286-287). As a result of this chain of calculations, capital expenditures is forecasted as a percent of revenues, and thus the forecast of capital expenditures growth should equal the forecast of sales growth, as in (R2).

growth rates. We estimate, CapEx Growth$_{i,t\to t+1}$ $= \alpha + \beta$Avg CapEx Growth$_{i,t-2\to t}$ $+$ $\varepsilon_{i,t+1}$, where Avg CapEx Growth$_{i,t-2\to t}$ denotes the average CapEx annual growth rate from $t-2$ to t. If $\alpha = 0$ and $\beta = 1$, (R1) would yield on average an accurate forecast of CapEx growth. In the data, however, $\widehat{\alpha} = 0.316$ and $\widehat{\beta} =$ -0.089, precisely estimated and statistically different from 0 and 1, respectively, implying CapEx growth is mean-reverting, which is not captured by (R1).

The second row of Table 1 reports the results of a regression designed to implement (R3) and assess (R2). We estimate, CapEx Growth$_{i,t\to t+1}$ $= \alpha + \beta \cdot$ Sales Growth$_{i,t\to t+1}$ $+$ $\varepsilon_{i,t+1}$. We find $\widehat{\alpha} = 0.217$ and $\widehat{\beta} = 1.055$, precisely estimated. Thus, a firm wishing to achieve a 5% sales growth rate would, under (R3), forecast a CapEx growth rate of 1.055 x 5% + 21.7% $\approx$ 27%. These estimates imply that (R2), which assumes $\alpha = 0$ and $\beta = 1$, provides much lower CapEx growth forecasts than (R3). Indeed, the same firm wishing to achieve a 5% sales growth rate would forecast under (R2) a CapEx growth rate of only 5%. The discrepancy between the two rules is driven by the intercept: equal to zero according to (R2), positive and large according to (R3). The estimated slope of 1.055 under (R3) is statistically indistinguishable from 1 assumed by (R2). Relative to (R3), (R2) underpredicts the growth rates of other inputs as well, although to a smaller extent. For example, a firm aiming at a 5% sales growth should under (R3) forecast a 12.7% advertising growth and a 7.7% wage growth.

The subsequent ten rows implement (R4) by estimating the same regression separately for each of ten industries, defined by one-digit SIC codes.[7] The intercept ranges from 0.163 to 0.402, and the slope ranges from 0.859 to 2.050, and both are statistically significant in all ten industries. These results confirm that in all industries (R2) implies a lower CapEx growth forecast relative to sales growth than in the data.

The last row reports results from a multivariate regression implementing (R5), CapEx Growth$_{i,t\to t+1}$ $= \alpha + \beta \cdot$ Sales Growth$_{i,t\to t+1}$ $+ \theta \cdot$ Earnings Growth$_{i,t\to t+1}$ $+ \varepsilon_{i,t+1}$. Earnings forecasts are widely reported and earnings realizations are widely available, unlike other variables. We find $\widehat{\alpha} = 0.217$, $\widehat{\beta} = 1.042$, and $\widehat{\theta} = 0.018$, precisely estimated.

As a next step, we measure the distance between the CFO forecast reported in the

[7] Agriculture, forestry, fishing; Mining; Construction; Manufacturing; Transportation, communication, utilities; Wholesale trade; Retail trade; Finance, insurance, real estate; Services; Public administration.

Duke data and the forecasts implied by each RoT. First, we compute five values, each equal to the orthogonal distance between the CFO's forecast and that implied by each RoT. Second, we compute the minimum among these five values. For each CFO, i, we refer to the distance-minimizing RoT, $\tau^i \in \{1, ..., 5\}$, as to the CFO's 'type.'[8]

Panel A of Table 2 shows that, among the 396 CFOs for whom we observe their identity and the joint forecasts and realizations of all variables, a plurality of about 40% makes a forecast that is closest to that implied by (R2), 27% of whom appear to use exactly (R2). Another 7.6% of CFOs are closest to (R1). Thus, about 48% of CFOs make forecasts closest to an incoherent RoT. We also find that about 11% of CFOs make forecasts closest to (R3), 27% to (R4), and 15% to (R5). These results underscore the large heterogeneity in CFO forecasts, possibly reflecting the challenges of making coherent forecasts and the absence of a consensus on which RoTs to use.

The average minimum distance of 0.033 is reasonably small, while masking substantial heterogeneity as revealed by a standard deviation of 0.059. Small discrepancies between actual and minimum-distance forecasts may reflect rounding or truncation in CFOs' reports and/or small differences in rule implementation between CFOs and us, e.g., the use of a different number of past observations for (R1) or a different sample size for (R3)-(R5). Larger discrepancies raise the possibility forecasts reflect different methods. To investigate this possibility, we compute the fraction of CFOs for whom their actual forecast falls outside of a pre-specified interval around the minimum-distance one. We find that the actual forecast falls within a $+/-0.05$ interval around the forecast implied by one of the five RoTs for 81.3% of CFOs and within a $+/-0.025$ interval for close to 60% of CFOs. By way of comparison, the standard deviation of sales growth forecasts is 0.271 (see Panel A of Table A2). Hence, the widths of the $+/-0.05$ and $+/-0.025$ intervals are about 1/3 and 1/6, respectively, of that standard deviation. These figures suggest our classification is plausible.

To account for the possibility that some CFOs use a different method, in Panel B we

[8]To focus on coherence and minimize confounding with accuracy, we compute the distance relative to the sales forecast, as CFOs are more accurate at forecasting sales than any other item, as we show later. In the Web Appendix, we present multiple robustness checks. Table A4 implements (R5) using advertising instead of earnings. Tables A5 and A6 report the distributions of RoTs, when the distance is measured relative to earnings and CapEx, respectively.

report an alternative categorization in which we assign to a residual category, (R6), the 18.7% CFOs whose actual forecasts are outside of a $+/-0.05$ interval around the forecasts implied by (R1)-(R5). Of the CFOs still categorized as (R1)-(R5), 6.2% make forecasts closely aligned with (R1), 40.7% with (R2), 12.7% with (R3), 25.8% with (R4), and 14.6% with R5. We use this alternative categorization in robustness analyses.

2.2 *Incoherence*

To construct a measure of incoherence, we require a reference forecasting rule that captures the contemporaneous linear relationship among multiple balance-sheet items. Among the rules in Welch (2017), (R5) provides the least restrictive linear specification, allowing each item to comove with sales and at least one additional item. The other rules can be viewed as restricted versions of this specification: (R3)-(R4) omit additional inputs or focus on subsamples, (R2) imposes a zero intercept and unit slope, and (R1) projects each item in isolation. Because (R5) spans the space of linear joint forecasts considered in this taxonomy, we use its implied hyperplane as a reference and define incoherence as the orthogonal distance of a CFO's forecast from that multivariate benchmark.[9]

We implement (R5) as, $\Delta \log I_t = \alpha + \beta \cdot \Delta \log Y_t + \theta \cdot \Delta \log \Psi_t + \varepsilon_t$, where Ψ denotes profits. This equation corresponds to the standard log-linear investment relationship derived from cost-minimization with quadratic adjustment costs, as we show in the Web Appendix A.1 (e.g., Hayashi (1982), Abel and Eberly (1994)). It represents the contemporaneous relationship among CapEx, output, and earnings growth. We measure output growth, $\Delta \log Y_t$, with sales growth, as standard in the literature (e.g., Syverson (2011)). This allows us to bypass the paucity of data on labor costs in Compustat.

To focus on coherence while minimizing confounding with forecast accuracy, we measure distance in the sales-forecast dimension. We therefore rewrite the same hyperplane in terms of sales growth as, $\Delta \log Y_t = \alpha^Y + \beta^Y \cdot \Delta \log I_t + \theta^Y \cdot \Delta \log \Psi_t + \varepsilon_t^Y$, and define incoherence as the orthogonal distance between the CFO's forecast of sales

[9]By doing this, we do not assume that (R5) is the optimal rule. Rather, we use it as a statistical reference because it represents the most general linear relationship among the items considered in the Welch (2017) taxonomy. Other rules can be expressed as restrictions of this multivariate form. Our incoherence measure captures the extent to which a CFO's forecasts depart from this benchmark.

growth and the sales-growth forecast implied by this hyperplane:

$$\text{Incoherence}_{i,t} = \frac{\left| \Delta\mathbb{F}_{i,t}[Y_{i,t+1}] - \widehat{\beta}^{Y} \Delta\mathbb{F}_{i,t}[I_{i,t+1}] - \widehat{\theta}^{Y} \Delta\mathbb{F}_{i,t}[\Psi_{i,t+1}] - \widehat{\alpha}^{Y} \right|}{\sqrt{1 + \left(\widehat{\beta}^{Y}\right)^2 + \left(\widehat{\theta}^{Y}\right)^2}} \tag{1}$$

where $\Delta\mathbb{F}_{i,t}\left[x_{i,t+1}\right] = \log\mathbb{F}_{i,t}\left[x_{i,t+1}\right] - \log\bar{x}_{i,t}$ denotes the forecast of the growth rate of a generic variable x, and $\widehat{\alpha}^{Y}$, $\widehat{\beta}^{Y}$, and $\widehat{\theta}^{Y}$ are the estimated coefficients from the sales-side representation of the multivariate hyperplane implied by (R5). The denominator normalizes the numerator to produce the shortest orthogonal distance from the forecast point to that hyperplane, as we show in the Web Appendix A.2.

We have 396 CFOs in our main sample. When using advertising instead of earnings growth (as in column 8 of Table A7 of the Web Appendix), we obtain similar results, although based on 130 CFOs only. Similarly, using wages or R&D instead of earnings growth, or more than two regressors, results in broadly similar patterns, although with even smaller samples of 50 CFOs or less. Rather than the specific regressors, what turns out to be crucial is moving from a univariate to a multivariate specification, that is, that forecasters account for the contemporaneous relationship among multiple items.

Because different forecasting rules impose different restrictions on the joint behavior of balance-sheet items, one would expect incoherence to vary systematically across rules. To examine this variation, we regress our incoherence measure in (1) on binary indicators for (R1)-(R4), with (R5) as the reference category.

Table 3 presents the results and shows that (R1) has the largest incoherence relative to (R5), followed by (R2). The average differences in incoherence of (R3) and (R4) relative to (R5) are much smaller in magnitude and not statistically significant.

Robustness analyses yield similar results. When using advertising instead of earnings to implement (R5), we still find (R1) is most distant from (R5), followed by (R2), and both are strongly statistically significant (see Table A8). When using the alternative categories (R1)-(R6) in Table A9, we find as expected that (R6) is the most distant from (R5) and, more remarkably, that the relative ranking of (R1)-(R4) is unchanged, with only (R1) and (R2) statistically significantly different from (R5) as in the baseline case.

Table 4 reports sample characteristics and investigates heterogeneity in ex ante incoherence by CFO and firm characteristics. In Panel A, 45% of CFOs have an MBA and 9% are women. On average, CFOs are 50.4 years old and have been on the job for 4.3 years. Average firm size is 2.5 billion USD of sales and 64% of firms pay a dividend. These figures are in line with those in Ben-David et al. (2013).

Panel B shows incoherence is lowest at the intermediate age range, 41-50, suggesting experience may help CFOs form more coherent forecasts, and coherence might decline at older ages. Incoherence is unrelated to optimism and miscalibration, consistent with the idea from psychology that incoherence and overconfidence are different constructs. Having an MBA does not correlate with incoherence, likely reflecting that, on the one hand, some RoTs are quite simple and CFOs may come up with them on their own and, on the other hand, there is no consensus in MBA textbooks on which rule is best; although in practice, as shown in Table 3, different rules perform differently in terms of ex ante coherence. Finally, CFOs of larger firms and firms paying dividends display lower incoherence, although the statistical significance is marginal. These results suggest that in stable or predictable environments it might be easier to make (more) coherent forecasts.[10]

2.3 *Tests of Coherence (and Accuracy) Using Forecast Errors*

Coherence imposes intuitive restrictions on how forecast errors of output and inputs must relate to one another. Consider the following empirical formulation of a generic production function,

$$y^f_{t+1} = \alpha + \alpha^f + \sum_{i=1}^{n} \beta^f_i x^f_{i,t+1} + \sum_{i=1}^{n} \sum_{s=0}^{t} \delta^f_{i,t-s} x^f_{i,t-s} + \sum_{j=1}^{m} \sum_{s=0}^{t} \gamma^f_{j,t-s} z^f_{j,t-s} + \varepsilon^f_{t+1}, \quad (2)$$

where f indexes firms, i inputs, t time, s lags, and j any relevant state variable (e.g., inventory, cash, etc.). The loadings of inputs and state variables, β^f_i, $\delta^f_{i,t-s}$, and $\gamma^f_{j,t-s}$, may vary across firms and over time, allowing for a wide array of lead-lag relationships. For example, firms in short-term industries will load more on inputs and state variables in recent years, whereas firms in long-term industries will load more on inputs and state variables in more distant years. The variables could be in levels, growth rates, logarithms,

[10]CFO tenure does not correlate with incoherence. This, combined with the fact that longer tenured CFOs delegate less (Graham et al., 2015), suggests that incoherence is unrelated to delegation of authority.

or any other transformation that preserves linearity in parameters. Note this formulation allows for many empirically relevant cases, including standard Q-theory, investment spikes, time-to-build considerations, financing costs and other state variables, and allowing for a quite general lead-lag relationship between investment and output.

Coherent forecasts of firm-specific variables should be cross-sectionally related as are their realizations in equation (2), that is,

$$\mathbb{F}_t\left[y^f_{t+1}\right] = \alpha + \alpha^f + \sum_{i=1}^{n} \beta^f_i \mathbb{F}_t\left[x^f_{i,t+1}\right] + \sum_{i=1}^{n}\sum_{s=0}^{t} \delta^f_{i,t-s} x^f_{i,t-s} + \sum_{j=1}^{m}\sum_{s=0}^{t} \gamma^f_{j,t-s} z^f_{j,t-s}. \quad (3)$$

Subtracting (3) from (2) *within firms* yields

$$\mathbb{FE}_t\left[y^f_{t+1}\right] \equiv y^f_{t+1} - \mathbb{F}_t\left[y^f_{t+1}\right] = \sum_{i=1}^{n} \beta^f_i \mathbb{FE}_t\left[x^f_{i,t+1}\right] + \varepsilon^f_{t+1}, \quad (4)$$

Differently from equation (3), equation (4) does not depend on any firm-level heterogeneity known to, or predictable by, the CFO at the time of forecast—but likely unobserved by the econometrician—as such heterogeneity gets differenced away when subtracting the forecast from the realization.

This discussion suggests two intuitive restrictions on FEs of contemporaneous inputs and output. First, as long as inputs contribute positively to output, that is, input loadings are positive, one would expect FEs of output and each input to be positively associated. Second, as long as no individual input has increasing returns, that is, input loadings are not larger than one, one would expect forecast errors to lie below the 45 degree line. Together, these restrictions imply the shaded areas in Figure 1.

Figure 2 plots the joint distribution of FEs of Sales and CapEx. While the regression slope is positive as expected, 42% of CFOs have FEs of Sales and CapEx of opposite sign, violating the first restriction of positive loadings. An additional 10.4% of FE pairs imply an input loading larger than one, thus violating the second restriction. Therefore, more than 50% of CFOs report contemporaneous FEs that violate intuitive coherence bounds. We find similar results in all ten industries, consistent with the idea that examining contemporaneous FEs is informative about coherence irrespective of the lead-lag relationships or the short- vs- long-term outlooks of different industries. We also find similar results for additional FE pairs, including sales and earnings.

Next, we assess equation (4) by regressing FEs of output on FEs of inputs,

$$\mathbb{FE}_t \log y^f_{t+1} = \alpha + \beta_1 \cdot \mathbb{FE}_t \log x^f_{1,t+1} + \beta_2 \cdot \mathbb{FE}_t \log x^f_{2,t+1} + \varepsilon^f_{t+1}, \tag{5}$$

whereby under the null of forecast coherence one would expect the constant to equal zero and the slope of the FE of each input to equal the loading of the corresponding input in (2). In the Cobb-Douglas benchmark with capital stock and labor as inputs, this implies $\alpha = 0$, $\beta_1 = a$, and $\beta_2 = b$, where a and b are the capital and labor shares,[11] which we compute using data from the Bureau of Economic Analysis.[12]

This framework further allows to distinguish between forecast accuracy and coherence, as accuracy can be assessed by testing for each variable that the mean of the FEs is zero, Avg $\mathbb{FE}_t\left[y^f_{t+1}\right] = 0$ and Avg $\mathbb{FE}_t\left[x^f_{i,t+1}\right] = 0$ for each input i. Equivalently, accuracy can be tested by regressing realizations on forecasts for each variable.[13]

Table 5 reports the results. In columns 1 and 2, we test for accuracy of CapEx and sales growth. In column 1, the constant is statistically different from zero at 10%; in column 2 the constant is not statistically different from zero. Hence, CFOs are on average inaccurate about CapEx and accurate about sales. The latter is consistent with Graham (2022)'s observation that top executives care the most about sales forecasts.

In column 3 we test for coherence. We find $\hat{\alpha} = 0.046$, statistically different from 0 at 5%; $\hat{\beta}_1 = 0.113$, marginally significant at 10%; and $\hat{\beta}_2 = 0.023$, not statistically different from zero. Hence, forecast errors of output and inputs do not satisfy the joint restrictions implied by coherence, and we reject the null of coherence.

The coherence test in column 3 requires FEs on wages, implying we could perform it

[11] The interpretation $\beta_1 = a$ holds exactly when the capital-side input is capital stock. In our empirical implementation, we use capital expenditures, so the mapping from theory to equation (5) is reduced-form: under $K_{t+1} = K_t(1-\delta) + I_t$, forecast errors in CapEx proxy locally for forecast errors in the capital-side input, but the exact coefficient need not equal the capital share.

[12] To implement the tests, we need $\text{FE}_t \log x_{1,t+1}$ and $\text{FE}_t \log y_{t+1}$, which are not directly observed in our data as the forecasts were not elicited in logs. To address this issue, we use $\mathbb{E}_t \log x_{t+1} \approx \log \mathbb{E}_t x_{t+1} - \frac{1}{2} V_t \log x_{t+1}$, a transformation written for a generic input x, and $V_t \log x_{t+1}$ is the conditional variance of $\log x$. Hence, $V_t \log x_{i,t+1} = \sigma_i^2\left(1-\gamma_i^2\right)$ for $i = 1, 2$, where σ_i^2 is the unconditional variance and γ_i the coefficient of an AR(1) regression of $\log x_{i,t}$. The conditional variance of y is $V_t \log y_{t+1} = a^2\sigma_1^2 + b^2\sigma_2^2$.

[13] Under rationality and no aggregate shocks, one expects a unit slope and a zero intercept. There is a long tradition of rational expectations tests of this type (e.g., Muth (1961), Lovell (1986), Pesaran and Weale (2006), D'Haultfoeuille et al. (2021), Crossley et al. (2024), and Giustinelli and Shapiro (2024)).

on 51 observations only, raising issues of statistical power. In column 4, we implement the regression test without wages. We now find $\hat{\alpha} = -0.004$, not statistically different from zero, and $\hat{\beta}_1 = 0.135$, precisely estimated. This reduced-form specification yields a similar qualitative pattern, although omitting wages weakens the direct structural interpretation of the coefficient. In column 5 we estimate (4) without CapEx and find the coefficient on wages to remain statistically insignificant.

In sum, CFO forecasts are on average accurate for sales, inaccurate for CapEx, and incoherent across the two. These results indicate that accuracy and coherence are not only distinct concepts in theory, as illustrated in Figure 3, but also in our data. These conclusions are confirmed in individual-level tests based on stronger assumptions, which we derive, implement, and report in Web Appendix C.

2.4 *Incoherence, Rules of Thumb, Predictable Forecast Errors*

We next examine whether incoherent forecasts are associated with systematically predictable forecast errors. First, we estimate the following specification,

$$\mathbb{FE}_t\text{CapEx Growth}_{t+1}^f = \beta_0 + \beta_1 \cdot \text{CapEx Growth}_t^f + \beta_2 \cdot \text{Sales Growth}_t^f + \beta_3 \cdot \text{Earnings Growth}_t^f + \varepsilon_{t+1}^f.$$

Table 6 reports the results. In columns 1 and 2, we find $\widehat{\beta}_1 < 0$, statistically significant, indicating high current CapEx growth predicts low subsequent FEs. Hence, CFOs overreact to realizations of CapEx growth when forecasting future CapEx growth.

Next, we examine whether this pattern differs by RoT. In columns 3 and 4 we find the largest overreaction for CFOs whose forecasts are closest to (R1) and (R2), with the latter strongly statistically significant.

These results differentiate our paper from recent literature that links specific deviations from perfect rationality to systematic forecast errors. Coibion and Gorodnichenko (2015) show how investor inattention generates underreaction. Bordalo et al. (2020) show how diagnostic expectations generate overreaction. We also document overreaction, but we link it to a different mechanism, namely, incoherent forecasting heuristics, as in Table 6. We discuss later in Section 3 one possible theoretical microfoundation of this result.

So far, we have distinguished forecast coherence from accuracy, and shown that incoherence generates predictable FEs. Next, we test whether incoherence and incoherent heuristics are associated with real effects on performance and investment.

2.5 *Incoherence, Rules of Thumb, and Firm Performance*

We now examine whether corporate performance varies with managerial incoherence, by estimating the regression,

$$\mathrm{ROA}_{i,j,t} = \alpha + \lambda_j + \delta_t + \beta \cdot \mathrm{Incoherence}_{i,j,t} + \theta \cdot X_{i,j,t} + \varepsilon_{i,j,t},$$

where i indexes CFO-firm pairs, j indexes industries, t indexes time, the dependent variable $\mathrm{ROA}_{i,j,t}$ is the percent return on the firm's assets, λ_j are industry fixed effects, δ_t are survey fixed effects, and $X_{i,j}$ includes firm-level regressors (firm size, market-to-book, and dividends) and CFO-level regressors (miscalibration and optimism, measured at both short- and long-term horizons). If deviations from coherence in internal plans reflect misallocation of inputs, higher incoherence should be associated with lower firm performance. Accordingly, we expect $\beta < 0$.

We conduct the estimation over 2001-2007 to avoid the effect of the financial crisis, an aggregate shock which impacted managerial beliefs (Boutros et al., 2025), and to focus on the peak of the 'great moderation,' when aggregate volatility was not a concern. We compute bootstrap standard errors and cluster them at the firm level. Given the lack of exogenous variation in incoherence, the results should be interpreted as correlational.

Table 7 shows the results. In column 1, a 1-standard deviation increase in incoherence (0.079 from Table 3) is associated with a 3 percentage points (p.p.) lower ROA, significant at the 5% level. Results are stable when we further condition on CFO miscalibration and optimism (columns 2 and 3), industry and survey fixed effects (column 4), and firm-level regressors (columns 5-8). Thus, while sample size changes across columns reflecting availability of regressors, the estimated negative association between CFO incoherence and firm performance remains quite stable in magnitude and statistically significant.

Next, we assess the extent to which this result reflects different RoTs by regressing firm performance on RoT indicators and other regressors. Consistent with the ranking documented in Table 3, in Table 8 we find that firms whose CFO forecasts are closest to (R1) have a 5-to-6 p.p. lower ROA relative to (R5). These are large differences in economic terms. We also find firms whose CFO forecasts are closest to (R2) or (R3) have a 2-3 p.p. lower ROA relative to (R5), whereas (R4) are indistinguishable from (R5).

Table A10 in the Web Appendix reports similar results when using an alternative definition of (R5) based on advertising growth.

In a robustness analysis using the (R1)-(R6) categorization, we find that firms whose CFOs are classified as (R6) have a 4 p.p. lower average ROA and results for (R1)-(R4) firms are very similar to those presented in Table 8. Therefore, CFOs whose forecasts are quite distant from those implied by all the five rules we study are unlikely to use a more sophisticated or 'better' approach, further supporting our focus on Welch (2017)'s taxonomy and our version of (R5) as a benchmark rule. In sum, firm performance is negatively associated with incoherence, and is on average lowest for firms whose CFO forecasts are closest to (R1) or (R2).

2.6 *Incoherence, Rules of Thumb, and Corporate Investment*

If downward-biased capital expenditure forecasts are only partially corrected ex post when firms implement their plans, firms relying on incoherent rules should exhibit lower realized investment. In Table 9, we assess this prediction by estimating,

$$\mathrm{Y}_{i,j,t} = \alpha + \lambda_j + \delta_t + \beta \cdot \mathrm{RoT}_{i,j,t} + \theta \cdot X_{i,j,t} + \varepsilon_{i,j,t},$$

where the indexes i, j, and t are defined as in the earlier regression for ROA; $\mathrm{RoT}_{i,j,t}$ is a vector of indicators for (R1)-(R4); $\mathrm{Y}_{i,j,t}$ is CapEx over assets (columns 1-3), and $X_{i,j,t}$ is a vector of covariates including firm size, market-to-book ratio, and dividends.

In columns 1 and 2, (R1) and (R2) are associated with 1.3-1.6 p.p. lower CapEx relative to (R5). The difference is large in economic terms and for (R2) statistically significant. These results continue to hold after conditioning on miscalibration and optimism. We confirm Ben-David et al. (2013)'s finding that miscalibration and optimism are positively (rather than negatively) associated with investment spending, underscoring that miscalibration, optimism, and incoherence are distinct constructs. In sum, (R1) and (R2) are associated with lower investment relative to (R5).

2.7 *Change in Firm Behavior when CFOs Take Office*

The negative association between performance and incoherence may reflect different possibilities. Higher incoherence may lead to lower investment and performance. Or, lower investment may induce CFOs to make incoherent forecasts, for example, to forecast too high a sales growth. Finally, more incoherent CFOs might be selected by—or self-select into—firms with lower investment spending and worse performance.

To make a step toward causality and shed light on its direction, we study how firm performance and investment evolve in the years around a CFO's hiring. We extract the dates when CFOs join firms from Execucomp and Boardex data and supplement this information by hand-collecting data from 10-K filings. A CFO is considered to take office in a firm when he or she first signs the financial reports. We match performance, investment, and characteristics from Compustat for the year of taking office. We use as dependent variables the within-firm differences in ROA and investment between the two years after and prior the CFO's taking office.

Table 10 presents the results. Performance declines following the appointment of incoherent CFOs, especially those with forecasts closest to (R1). (R1) is also associated with a 2.2 p.p. lower investment intensity. While not unambiguously causal, these results align with the evidence documented above.

3. A Simple Model of Forecast Coherence

We present a model to illustrate forecast coherence and interpret our empirical results. Subsection 3.1 introduces the benchmark model and shows incoherent RoTs yield lower performance and overreaction. Subsection 3.2 shows incoherent RoTs yield underinvestment. Subsection 3.3 discusses one possible microfoundation of incoherence.

3.1 *A Benchmark Model of Optimal Corporate Forecasts*

Consider a firm producing output y_t at time t using n inputs, where x_i is the quantity of input i, p_i is its price, and p_y the output price, which we normalize to 1. To map the theory into our data, in the main text we develop the Cobb-Douglas case with 2 inputs (K, L), $y = \left(x_1^a \cdot x_2^b\right)$, and we consider the more general CES case in Web Appendix A.6.

The CFO's task at any time t is to choose inputs to minimize costs subject to meeting a given sales growth target and taking input prices as given. While capital stock follows, $K_{t+1} = K_t(1-\delta) + I_t$, where δ is the depreciation rate and I_t is capital expenditures, at any time t at which the CFO forecasts one-year ahead variables, the current stock of capital, K_t, is given, and all that matters for expected one-year ahead profits are the expected output and the capital (and labor) expenditures. As a result, the CFO aims to produce input growth forecasts that are both individually accurate and jointly coherent, in the sense that they are consistent with the firm's production technology.

We assume the technology is stable over the forecasting horizon and abstract from aggregate shocks, so factor-augmenting productivities are constant and normalized to one. We further maintain standard producer-theory assumptions: firms take input and output prices as given and choose production plans to minimize costs.

The CFO is in charge of internal plans, which consist of forecasts of growth rates of multiple balance-sheet variables. We denote the forecast of sales growth at time t with $\Delta\mathbb{F}y_{t+1} = \log\mathbb{F}y_{t+1} - \log\bar{y}_t$, and the forecasts of the inputs' growth rates at time t with $\{\Delta\mathbb{F}x_{i,t+1} = \log\mathbb{F}x_{i,t+1} - \log\bar{x}_{i,t}\}_{i=1,2}$, where $\bar{\eta}_t$ denotes the realization of the generic variable η. It is important to note that the forecasters observe realizations of own and other firms' output and inputs, e.g., by accessing Compustat. On the other hand, the forecasters observe only their own past and current forecasts, but not those of other firms.

We consider the problem of a CFO who takes $\Delta\mathbb{F}y_{t+1}$, set by the CEO and board of directors, as given, and forecasts the growth rate of input expenditures under square loss,

$$\min_{\Delta\mathbb{F}x_{1,t+1},\,\Delta\mathbb{F}x_{2,t+1}} \mathbb{E}\left[(\Delta x^*_{1,t+1} - \Delta\mathbb{F}x_{1,t+1})^2 + (\Delta x^*_{2,t+1} - \Delta\mathbb{F}x_{2,t+1})^2 \,\middle|\, \Omega_t\right],$$

where Ω_t denotes the CFO's information set at t, which includes knowledge of the production function, and $\{\Delta x^*_{1,t+1}, \Delta x^*_{2,t+1}\}$ denote the cost-minimizing inputs growth needed to obtain the sales growth target, $\Delta\mathbb{F}y_{t+1}$.

In words, the CFO chooses the vector of input-growth forecasts that minimizes expected inaccuracy, subject to a coherence constraint that forecasts be consistent with the firm's technology and sales growth target. This formulation maps the CFO's forecasting problem into a standard cost-minimization problem. An analogous analysis would obtain under profit maximization. Square loss lets the model nest RoTs (R1)-(R5).

3.1.1 Optimal Forecasts

Lemma 1 (Coherent forecasts under Cobb-Douglas). *Forecast coherence requires that forecasts of output and inputs, $\mathbb{F}_t\left[y_{t+1}\right]$, $\mathbb{F}_t\left[x_{1,t+1}\right]$, and $\mathbb{F}_t\left[x_{2,t+1}\right]$, satisfy,*

$$\mathbb{F}_t \log\left[y_{t+1}\right] = a \cdot \mathbb{F}_t \log\left[x_{1,t+1}\right] + b \cdot \mathbb{F}_t \log\left[x_{2,t+1}\right],$$
$$\Delta\mathbb{F}_t\left[y_{t+1}\right] = a \cdot \Delta\mathbb{F}_t\left[x_{1,t+1}\right] + b \cdot \Delta\mathbb{F}_t\left[x_{2,t+1}\right].$$

The Web Appendices A.3 and A.4 contain all proofs.

Next, define the forecast error for a generic variable, η, as realization minus forecast, $\mathbb{FE}_t\left[\eta_{t+1}\right] = \eta_{t+1} - \mathbb{F}_t\left[\eta_{t+1}\right]$, and recall that these forecast errors are defined at the CFO (and, hence, firm) level. We obtain the following key result.

Lemma 2 (Coherence, accuracy, and forecast errors under Cobb-Douglas). *Forecast coherence requires that,*

$$\mathbb{FE}_t\left[\Delta y_{t+1}\right] \equiv \mathbb{FE}_t\left[\log y_{t+1}\right] = a \cdot \mathbb{FE}_t\left[\log x_{1,t+1}\right] + b \cdot \mathbb{FE}_t\left[\log x_{2,t+1}\right].$$

Lemma 2 links coherence and accuracy: accuracy requires forecast errors to be zero on average item by item, whereas coherence requires forecast errors of different items to be cross-sectionally linked by the technology parameters. While the exact equality relies on Cobb-Douglas log-linearity, the distinction between coherence and accuracy does not: under CES, coherence yields local monotonicity and sign restrictions rather than exact equalities (see Web Appendix A.6).

3.1.2 Optimal Forecasts and Rules of Thumb

We now consider the forecasting problem when a and b are unknown to the forecaster.

Lemma 3 (Unknown parameters). *If parameters a and b are unknown, a forecaster can estimate them using a linear projection operator, with forecasted variables in logs.*

In what follows, we assume past data on output, inputs, and prices are observed by CFOs and are informative.[14] A CFO can then use such data to estimate the parameters of interest and provide forecasts. We define the following RoTs for $\Delta\mathbb{F}x_{1,t+1}$:

(R5) Estimate a multivariate linear projection of $\Delta\bar{x}_{1,t}$ on a constant, $\Delta\bar{y}_t$, and $\Delta\bar{x}_{2,t}$, with coefficients c_0, c_1, and c_2, and analogously for input 2. Then define $\Delta\mathbb{F}x^{R5}_{1,t+1} = c_0 + c_1\Delta\mathbb{F}y_{t+1} + c_2\Delta\mathbb{F}x^{R5}_{2,t+1}$, where $\Delta\mathbb{F}x^{R5}_{2,t+1}$ is the forecast of the other input implied by the same multivariate rule.

(R3) Estimate a univariate linear projection of $\Delta\bar{x}_{1,t}$ on a constant and $\Delta\bar{y}_t$, with coefficients d_0 and d_1. Then define $\Delta\mathbb{F}x^{R3}_{1,t+1} = d_0 + d_1\Delta\mathbb{F}y_{t+1}$.

(R2) Define $\Delta\mathbb{F}x^{R2}_{1,t+1} = \Delta\mathbb{F}y_{t+1}$.

(R1) Define $\Delta\mathbb{F}x^{R1}_{1,t+1} = \dfrac{1}{k}\sum_{j=1}^{k}\Delta\bar{x}_{1,t+1-j}$, where $\Delta\bar{x}_{1,t+1-j}$ is allowed to differ from the cost minimizing growth of input 1, $\Delta x^*_{1,t}$, needed to achieve the sales growth target.

Relative to the empirical definitions in Welch (2017), three simplifications apply here: (R5) contains only technologically relevant inputs; (R4) is indistinguishable from (R3) because there is no industry heterogeneity; and (R2) is a special case of (R3) with $d_0 = 0$ and $d_1 = 1$.

From now on, we maintain the following assumption.

Assumption. *In all previous time periods the firm's actual input spending equals the cost-minimizing level, $\bar{x}_{i,s} = x^*_{i,s}$, $\forall s \leq t$, $i = 1, 2$, needed to attain the sales growth target.*

The CFO has access to an observational dataset of past input prices, in which the input prices of the CFO's firm represent a single time-series.

We proceed to obtain a (partial) ranking of these RoTs based on their expected loss.

Proposition 1 (Ranking RoTs under Cobb-Douglas). *In the Cobb-Douglas case, when the co-movement of input prices is sufficiently large, a forecasting rule that conditions on all relevant variables (output and other inputs) coincides with the optimal forecast and minimizes expected squared error. In particular,*

$$\Delta\mathbb{F}x^{R5}_{1,t+1} = \mathbb{E}_t\left[\Delta x^*_{1,t+1}\right].$$

[14] More formally, we assume all firms have the same technology, all CFOs face the same forecasting and cost minimization problems, and past data is generated by optimizing CFOs.

Hence, (R5) is unbiased and yields weakly lower expected loss than (R3), (R2), or (R1). Moreover, (R3) yields weakly lower expected loss than (R1) and (R2). Finally, (R2) yields weakly lower expected loss than (R1).

Proposition 1 yields a benchmark ranking in which the multivariate rule dominates more restrictive rules. We next ask whether using these RoTs predicts systematic FEs.

Implication 1 (Predictable forecast errors). *When forecasts are generated using incoherent rules, forecast errors become systematically predictable from past realizations. In particular, these rules generate overreaction: high past investment growth leads to overly optimistic forecasts and negative forecast errors ex post.*

More formally, consider the local linear projection,

$$\Delta x^*_{1,t+1} - \Delta \mathbb{F} x_{1,t+1} = m_0 + m_1 \Delta \bar{x}_{1,t}. \tag{6}$$

We show in Web Appendix A.4 that under the coherent rule (R5), $m_0 = m_1 = 0$, so FEs are not predictable from past realizations. Under (R1) and (R2), instead, $m_1 < 0$: following high CapEx growth, CFOs forecast excessively high future CapEx growth, which then mean reverts and generates negative FEs ex post. Thus incoherent heuristics give rise to overreaction, consistent with Table 6.

3.2 *Rules of Thumb and Underinvestment*

We have established that incoherent forecasting rules can generate biased input forecasts relative to the cost-minimizing benchmark. We now show how such forecast bias can translate into real underinvestment when adjustment is costly.

Let $\Delta x^*_{1,t+1}$ denote the cost-minimizing growth rate of input 1 required to achieve the sales target $\Delta F y_{t+1}$. Under rule (Rk) $\in \{$(R1), (R2), (R3), (R5)$\}$, the CFO issues a forecast $\Delta \mathbb{F} x^{Rk}_{1,t+1}$. Suppose that after the forecast is issued, the firm can adjust input growth, but faces convex adjustment costs. Let realized input growth be,

$$\Delta x_{1,t+1} = \Delta \mathbb{F} x^{Rk}_{1,t+1} + \lambda \big(\Delta x^*_{1,t+1} - \Delta \mathbb{F} x^{Rk}_{1,t+1}\big), \quad \lambda \in (0, 1], \tag{7}$$

where λ captures the degree of adjustment. When $\lambda = 1$, the firm fully corrects forecast errors and attains the cost-minimizing level. When $\lambda < 1$, adjustment is partial due to costs, so deviations between forecasted and optimal growth translate into realized under- or overinvestment. Define the investment gap under rule (Rk) as,

$$G_{t+1}^{Rk} = \Delta x_{1,t+1}^{*} - \Delta x_{1,t+1}. \tag{8}$$

Substituting the adjustment equation yields,

$$G_{t+1}^{Rk} = (1-\lambda)\big(\Delta x_{1,t+1}^{*} - \Delta \mathbb{F} x_{1,t+1}^{Rk}\big). \tag{9}$$

Thus, realized underinvestment is proportional to the forecast bias, $\Delta x_{1,t+1}^{*} - \Delta \mathbb{F} x_{1,t+1}^{Rk}$. When forecasts are unbiased (as under coherent (R5)), $\mathbb{E}[G_{t+1}^{R5}] = 0$. When forecasts are downward biased (as under (R1) or (R2)), the expected investment gap is strictly positive:

$$\mathbb{E}[G_{t+1}^{Rk}] > 0 \quad \text{if} \quad \mathbb{E}\big[\Delta x_{1,t+1}^{*} - \Delta \mathbb{F} x_{1,t+1}^{Rk}\big] > 0. \tag{10}$$

Implication 2 (Underinvestment and RoTs). *Under convex adjustment costs, forecasting rules that generate downward-biased input forecasts produce persistent underinvestment relative to the cost-minimizing benchmark. In particular, if* $\mathbb{E}[\Delta x_{1,t+1}^{*} - \Delta \mathbb{F} x_{1,t+1}^{Rk}] > 0$, *then* $\mathbb{E}[G_{t+1}^{Rk}] > 0$.

The intuition is immediate. When the CFO underestimates the input growth required to meet the sales target, the firm must revise plans ex post. If adjustment is costly, revision is partial, and realized investment remains below the optimal level. Because incoherent rules such as (R2), and under some conditions (R1), underpredict CapEx growth relative to the cost-minimizing benchmark, they can yield lower investment than (R5). Implication 2 rationalizes our empirical results in Table 9; see derivations in Web Appendix A.5.

In sum, coherent forecasts require the forecaster to account for the structure of the firm's problem and the joint behavior of output and inputs. Incoherent RoTs can generate lower performance, overreaction, and underinvestment relative to coherent forecasts, consistent with our empirical results.

3.3 *A Model of Bracketing in Corporate Forecasts*

A natural way to interpret incoherent forecasting rules is through the lens of behavioral bracketing. We provide a formal model of bracketing-based forecasting building on Lian (2021), which shows how different informational frictions can rationalize the use of simple forecasting rules. Importantly, the empirical patterns documented above do not depend on this particular microfoundation.

Narrowly bracketed forecasts treat each decision in isolation, whereas broadly bracketed forecasts account for interactions across decisions. We capture this idea by introducing noisy signals and recasting the multidimensional forecasting problem as multiple selves playing an incomplete-information, common-interest game. The CFO "capital self" makes forecasts of capital expenditures' growth by observing noisy signals of output and labor growth. Conversely, the CFO "labor self" makes forecasts of labor expenditures' growth by observing noisy signals of output and capital growth. In equilibrium, each self does not observe the others' signals and therefore forms forecasts with imperfect knowledge of the others' forecasts.

We obtain the following results (full details and proofs in Web Appendix A.7).

Proposition 2 (Optimal forecasts with noisy signals). *The optimal forecast of* x_1*, given signals* $\eta_y = \log y + \epsilon_y$ *and* $\eta_2 = \log x_2 + \epsilon_2$*, with* $\epsilon_y \sim \mathcal{N}\left(\mu_y, s_y^2\right)$*,* $\epsilon_2 \sim \mathcal{N}\left(\mu_2, s_2^2\right)$ *is,*

$$\mathbb{E}\left[\log x_1 | \eta_y, \eta_2\right] = \mu_1 + \beta_y\left(\eta_y - \mu_y\right) + \beta_2\left(\eta_2 - \mu_2\right),$$

where $\beta_y = \frac{a\sigma_1^2}{a^2\sigma_1^2 + b^2\sigma_2^2 + s_y^2 - \frac{b^2\sigma_2^4}{\sigma_2^2 + s_2^2}}$, $\beta_2 = \frac{ab\sigma_1^2\sigma_2^2}{b^2\sigma_2^4 - \left(\sigma_2^2 + s_2^2\right)\left(a^2\sigma_1^2 + b^2\sigma_2^2 + s_y^2\right)}$.

Proposition 2 rationalizes a version of (R5) also as second-best optimal and clarifies that the accuracy of the linear projection depends on the signals' precision.

Corollary 2.1 (Narrow Bracketing). *When* s_y^2*,* $s_2^2 \to +\infty$*, the optimal forecast is* $\mathbb{E}\left[\log x_1 | \eta_y, \eta_2\right] = \mu_1$.

When all signals are infinitely noisy, the optimal forecast projects the prior mean of the input being forecasted into the future. Thus, Corollary 2.1 provides conditions under which a narrow-bracketing version of (R1) is second-best optimal.

Corollary 2.2 (Univariate Projections). *When* $s_2^2 \to +\infty$ *and* $0 < s_y^2 < +\infty$*, the* *optimal forecast is* $\mathbb{E}\left[\log x_1 | \eta_y, \eta_2\right] = \mu_1 + \beta_y \left(\eta_y - \mu_y\right)$, *where* $\beta_y = \frac{a\sigma_1^2}{a^2\sigma_1^2 + b^2\sigma_2^2 + s_y^2}$.

When the signal of the second input is infinitely noisy and that of the output is noisy but informative, the optimal forecast is a univariate linear projection, where the intercept is still the prior mean of the first input and the slope is a function of the fundamental uncertainty and precision of the output's signal. Thus, Corollary 2.2 provides conditions under which a univariate-projection version of (R3) is second-best optimal.

4. Related Literature

Our evidence on incoherence in firm plans provides a possible mechanism for understanding resource allocation within the firm (e.g., Maksimovic and Phillips (2002, 2013), Hoberg and Phillips (2025)). We discuss here our contributions to the study of survey expectations of firms and to behavioral corporate finance and economics.

Survey Expectations of Firms. Our paper is related to the recent and growing literature studying top executives' beliefs and forecasts about their firm's variables (Ben-David et al. (2013), Boutros et al. (2025), Campello et al. (2010), Campello et al. (2011, 2012)). This literature typically studies the accuracy of managerial forecasts, one variable at a time. Bloom et al. (2021) show that forecasting firms' own variables is harder than forecasting the aggregate economy. Gennaioli et al. (2016) show that both planned and actual investment are explained by CFOs' expectations of earnings growth. Graham (2022) documents that the sales revenue growth forecast is most important for firms' plans and consequences. We confirm that firms make on average accurate forecasts about sales, but we also show that expectations of other variables—most notably, capital expenditures—are much less accurate, and we link this to incoherence. We add to this literature by documenting wide heterogeneity in the ways managerial forecasts link together multiple related balance-sheet variables, some of which lead to incoherence.

Coherence and Accuracy Sides of Rationality in Behavioral Economics and Psychology. The psychology literature has long recognized that rationality of judgements involves both accuracy ('correspondence') and coherence ('consistency'), and has typically maintained that the two are separate properties (e.g., Hammond (1996, 2007), Osherson et al. (1994), Gigerenzer and Gaissmaier (2011), Lee and Zhang (2012), Arkes et al. (2016)). This literature has struggled to provide a formal framework or direct evidence to define coherence and disentangle it from accuracy.

Behavioral economics has focused on predicting systematic inaccuracy from specific violations of statistics and probability laws, starting from the coherence principle (de Finetti (1937)), which states that subjective probabilities must conform to the logical and probability axioms posited in Kolmogorov (1933). In a series of famous experiments, Tversky and Kahneman (1971, 1974, 1983) document systematic misconceptions of logic, statistics, and probability theory, including the law of large numbers, the conjunction rule, the law of total probability, and Bayes' theorem. Since then, a number of theories and experiments have shown how such misconceptions predict systematic inaccuracy in specific prediction tasks (e.g., Wright et al. (1994), Rabin (2002), Benjamin et al. (2016), Zhu et al. (2022)). Tversky and Kahneman (1974) have famously labeled this research program "heuristics and biases," whereby the use of heuristics generates systematic and predictable forecast errors (see also Thaler (2018)). Some scholars have pointed out how some of these results can also be cast in terms of the coherence-accuracy framework. For instance, Hammond (1996), Tentori et al. (2013), Arkes et al. (2016), Jönsson and Shogenji (2019), and others discuss how Tversky and Kahneman (1983)'s conjunction fallacy can be understood as a violation of coherence with respect to probability laws. In this literature, documenting a form of incoherence in one domain typically serves the purpose of predicting future systematic inaccuracy in another domain, rather than jointly measuring coherence and accuracy within the same forecasting setting.

Our main contribution to this literature is to provide a formal framework, in which coherence and accuracy are defined with respect to the same forecasting task, allowing us to jointly assess forecast accuracy and coherence. To do so one needs both theory and data. In terms of theory, an economic model provides a benchmark against which to judge coherence. This is similar to the use of probability theory for assessing coherence of probabilistic judgments, but crucially economic theory allows us to nest coherence and accuracy in a setting with optimizing agents. In terms of data, observing both forecasts and realizations—and, thus, FEs—for multiple variables simultaneously allows us to jointly assess accuracy and coherence, and to disentangle the two. Accuracy is assessed by testing whether FEs of each variable are 'sufficiently' close to zero. Coherence is assessed by testing whether FEs of different variables are 'sufficiently' close to one another. In both cases, the extent of 'sufficiently' is pinned down by economic theory.

Disentangling coherence and accuracy is important, since under some conditions coherence can be assessed ex ante. The intuition is similar to that of the debiasing program in psychology (e.g., Fischhoff (1982)), with the key difference that in our framework one can use economic theory and regression analysis to determine the ex ante optimal rule.

Behavioral Corporate Finance. Baker et al. (2007) organize the behavioral corporate finance literature under the two themes of irrational investors and irrational managers (Guenzel and Malmendier (2020) discuss also other biased parties). The more developed irrational-investors literature studies how rational managers respond to irrational fluctuations in capital market prices. Our study belongs to the less developed and growing irrational-managers literature, in which, while markets are arbitrage free, managerial behavior can be influenced by psychological biases. So far, this stream of research has focused mostly on how psychological traits such as optimism, overconfidence, or preference for skewness can distort managerial expectations about the future and, thus, investment and financing decisions (e.g., Malmendier and Tate (2005), Gervais et al. (2011), Schneider and Spalt (2014)). We study incoherence—a novel managerial trait—and, controlling directly for measures of optimism and overconfidence (Ben-David et al. (2013)), we show incoherence differs from optimism or overconfidence both in theory and data.

Behavioral Research on Bracketing. Experimental and behavioral research shows decision makers tend to narrowly bracket, making interrelated decisions in isolation (e.g., Tversky and Kahneman (1981), Read et al. (1999), Rabin and Weizsäcker (2009), Ellis and Freeman (2020)). Thaler (1985, 2018) and Heath and Soll (1996) propose individuals hold a mental account for each decision, thus failing to capture the interdependence among choice variables implied by their budget constraint and utility function.

Economic theories of narrow bracketing and mental accounting include Barberis et al.

(2006) and Rabin and Weizsäcker (2009) on monetary gambles and Hastings and Shapiro (2013) and Lian (2021) on consumption choices. Our theory is closest to Lian (2021)'s, which endogenizes bracketing by modeling a narrow thinker who makes consumption decisions about individual goods with imperfect knowledge of the others and thus faces difficulties at coordinating different decisions.

We add to this literature by developing the first model of bracketing in a firm setting and providing field evidence consistent with narrow bracketing among CFOs making internal plans. As in Lian (2021), there are no explicit mental budgets, and agents make decisions (forecasts) based on different non-nested information in the sense of Blackwell, i.e., neither input i's signal is more informative than $\neg i$'s signal, nor vice versa. These features differentiate Lian (2021) and us from rational inattention models (e.g., Sims (2003), Mackowiak and Wiederholt (2009), Matějka and McKay (2015), Kőszegi and Matějka (2020)), and Gabaix (2014, 2019)'s sparsity model. Unlike Lian (2021), we study a production setting where decisions are interconnected via the production technology.

5. Conclusion

In this paper we introduce forecast coherence—internal consistency across forecasts of jointly determined variables—as a benchmark for assessing firms' internal plans. We make three contributions. First, we document that over 80% of CFOs make joint forecasts of CapEx and sales closely aligned with those implied by the heuristics taught in MBA textbooks. Second, we show that forecasts aligned with restricted heuristics, (R1) and (R2), are systematically associated in the cross section with predictable reversals in forecast errors, and lower ROA and CapEx than forecasts aligned with a less restricted rule, (R5). Third, we build a model of multidimensional forecasting under a coherence constraint, illustrating how reliance on restricted forecasting rules can generate incoherence and resource misallocation within firms, consistent with our evidence.

Taken together, our findings suggest that, whether or not CFOs literally use them, the planning heuristics taught in MBA textbooks provide a useful framework for interpreting firm plans: they are closely aligned with the forecasts CFOs actually make, and they are systematically related to observable firm outcomes.

We view our results as reflecting both a lack of consensus in MBA textbooks and case studies, and a lack of theory and evidence to distinguish among alternative heuristics. Two takeaways follow. First, simple and widely used rules such as (R1) and (R2) are incoherent and should not be used to forecast CapEx. Second, more generally, not all heuristics are the same: some may be useful, others less so, and they should be evaluated with respect to their intended goals. Doing so typically requires both theory and data.

Our results also raise two broader questions. First, if some heuristics perform poorly, why might CFOs still rely on them? We discuss one possible formalization based on behavioral bracketing, which rationalizes incoherent heuristics as a second-best response to the extent of imprecision of signals about future firm variables. At the same time, our evidence does not identify the economic primitives generating incoherence in the data. Distinguishing among behavioral bracketing and other possible microfoundations is an important task for future work. Second, if some heuristics perform poorly, do firms or markets learn over time to make better plans? Answering this question requires going beyond the cross-sectional variation studied here.

We do not address the broader question of what the best possible forecasting method for internal plans should be, beyond the textbook heuristics studied here. Given the nonlinear nature of the firm's problem, richer data and more flexible forecasting methods, when available, may well yield superior plans.

References

Arkes, H. R., G. Gigerenzer, and R. Hertwig (2016). How Bad Is Incoherence? *Decision 3*(1), 20–39.

Baker, M., R. S. Ruback, and J. Wurgler (2007). *Handbook in Corporate Finance: Empirical Corporate Finance*, Chapter Behavioral Corporate Finance: A survey, pp. 145–188. Elsevier.

Barberis, N., M. Huang, and R. H. Thaler (2006). Individual Preferences, Monetary Gambles, and Stock Market Participation: A Case for Narrow Framing. *American Economic Review 96*(4), 1069–1090.

Becker, G. S. (1996). *Accounting for Tastes* (1st ed.). Harvard University Press.

Ben-David, I., J. R. Graham, and C. R. Harvey (2013). Managerial Miscalibration. *Quarterly Journal of Economics 128*(4), 1547–1584.

Benjamin, D. J. (2019). Errors in Probabilistic Reasoning and Judgment Biases. In B. D. Bernheim, Stefano DellaVigna, and David Laibson (Eds.), *Handbook of Behavioral Economics: Applications and Foundations 1*, Volume 2, Chapter 2, pp. 69–186. Elsevier.

Benjamin, D. J., M. Rabin, and C. Raymond (2016). A Model of Non-Belief in the Law of Large Numbers. *Journal of the European Economic Association 14*(2), 515–544.

Berndt, E. R. (1976). Reconciling Alternative Estimates of the Elasticity of Substitution. *Review of Economics and Statistics 58*(1), 59–68.

Bloom, N., T. Kawakubo, C. Meng, P. Mizen, R. Riley, T. Senga, and J. V. Reenen (2021). Do Well Managed Firms Make Better Forecasts? Working Paper 29591, NBER.

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020). Overreaction in Macroeconomic Expectations. *American Economic Review 110*(9), 2748–2782.

Boutros, M., I. Ben-David, J. Graham, C. Harvey, and J. Payne (2025). The Persistence of Miscalibration. Forthcoming, Review of Financial Studies.

Campello, M., E. Giambona, J. R. Graham, and C. R. Harvey (2011). Liquidity Management and Corporate Investment During a Financial Crisis. *Review of Financial Studies 24*(6), 1944–1979.

Campello, M., E. Giambona, J. R. Graham, and C. R. Harvey (2012). Access to Liquidity and Corporate Investment in Europe the Credit Crisis of 2009. *Review of Finance 16*(2), 323–346.

Campello, M., J. R. Graham, and C. R. Harvey (2010). The Real Effects of Financial Constraints: Evidence from a financial crisis. *Journal of Financial Economics 97*(3), 470–487.

Coibion, O. and Y. Gorodnichenko (2015). Information Rigidity and the Expectations Formation Process: A simple framework and New Facts. *American Economic Review 105*(8), 2644–2678.

Crossley, T., Y. Gong, T. Stinebrickner, and R. Stinebrickner (2024). Examining Income Expectations in the College and Early Post-college Periods: New Distributional Tests of Rational Expectations. *Journal of the European Economic Association 22*(6), 2700–2747.

de Finetti, B. (1937). Foresight: Its Logical Laws, Its Subjective Sourses. In S. Kotz,

N. J. Johnson, and H. E. Kyburg (Eds.), *Breakthroughs in Statistics*, pp. 134–174. Springer-Verlag.

D'Haultfoeuille, X., C. Gaillac, and A. Maurel (2021). Rationalizing Rational Expectations: Characterizations and Tests. *Quantitative Economics 12*(3), 817–842.

Ellis, A. and D. J. Freeman (2020). Revealing Choice Bracketing. Working Paper 14869, arXiv.org.

Fischhoff, B. (1982). Debiasing. In D. Kahneman, Slovic, and A. Tversky (Eds.), *Judgment Under Uncertainty: Heuristics and Biases*, Chapter 31, pp. 422–444. Cambridge University Press, New York.

Fogelin, R. (2003). *Walking the Tightrope of Reason: The Precarious Life of a Rational Animal*. Oxford University Press.

Gabaix, X. (2014). A Sparsity-Based Model of Bounded Rationality. *Quarterly Journal of Economics 129*(4), 1661–1710.

Gabaix, X. (2019). Behavioral Inattention. In B. D. Bernheim, Stefano DellaVigna, and David Laibson (Eds.), *Handbook of Behavioral Economics: Applications and Foundations 1*, Volume 2, Chapter 4, pp. 261–343. Elsevier.

Gennaioli, N., Y. Ma, and A. Shleifer (2016). Expectations and Investment. In M. Eichenbaum and J. A. Parker (Eds.), *NBER Macroeconomics Annual 15*, Volume 30. University of Chicago Press.

Gervais, S., J. B. Heaton, and T. Odean (2011). Overconfidence, Compensation Contracts, and Capital Budgeting. *Journal of Finance 66*(5), 1735–1777.

Gigerenzer, G. and W. Gaissmaier (2011). Heuristic Decision Making. *Annual Review of Psychology 62*, 451–482.

Giustinelli, P. and M. D. Shapiro (2024). SeaTE: Subjective ex ante Treatment Effect of Health on Retirement. *American Economic Journal: Applied Economics 16*(2), 278–317.

Graham, J. (2022). Presidential Address: Corporate Finance and Reality. *Journal of Finance 77*(4), 1975–2049.

Graham, J. R., C. R. Harvey, and M. Puri (2015). Capital Allocation and Delegation of Decision-Making Authority within Firms. *Journal of Financial Economics 115*(3), 449–470.

Guenzel, M. and U. Malmendier (2020). Behavioral Corporate Finance: Life Cycle of a CEO Career. In *Oxford Research Encyclopedia of Economics and Finance*. Oxford University Press.

Hammond, K. R. (1996). *Human Judgement and Social Policy: Irreducible Uncertainty, Inevitable Error, Unavoidable Injustice*. Oxford University Press.

Hammond, K. R. (2007). *Beyond Rationality: The Search for Wisdom in a Troubled Time*. Oxford University Press.

Hastings, J. S. and J. M. Shapiro (2013). Fungibility and Consumer Choice: Evidence from Commodity Price Shocks. *Quarterly Journal of Economics 128*(4), 1449–1498.

Heath, C. and J. B. Soll (1996). Mental Budgeting and Consumer Decisions. *Journal of Consumer Research 23*(1), 40–52.

Hoberg, G. and G. Phillips (2025). Scope, Scale, and Concentration: The 21st Century

Firm. *80*(1), 415–466.

Holthausen, R. W. and M. E. Zmijewski (2020). *Corporate Valuation: Theory, Evidence & Practice* (2nd ed.). Cambridge University Press.

Jönsson, M. L. and T. Shogenji (2019). A Unified Account of the Conjunction Fallacy by Coherence. *Synthese 196*, 221–237.

Kőszegi, B. and F. Matějka (2020). Choice Simplification: A Theory of Mental Budgeting and Naive Diversification. *Quarterly Journal of Economics 135*(2), 1153–1207.

Koller, T., M. Goedhart, and D. Wessels (2020). *Valuation: Measuring and Managing the Value of Companies* (7th ed.). John Wiley & Sons. The McKinsey Guide.

Kolmogorov, A. N. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Springer. English translation by N. Morrison, Foundations of the Theory of Probability, 1956.

Lee, M. D. and S. Zhang (2012). Evaluating the Coherence of Take-the-Best in Structured Environments. *Judgment and Decision Making 7*(4), 360–372.

Lian, C. (2021). A Theory of Narrow Thinking. *Review of Economic Studies 88*(5), 2344–2374.

Lovell, M. C. (1986). Tests of the Rational Expectations Hypothesis. *American Economic Review 76*(1), 110–124.

Luehrman, T. A. and J. L. Heilprin (2009). Mercury Athletic Footwear, Inc.: Valuing the Opportunity. Brief Case 4051, Harvard Business School.

Mackowiak, B. and M. Wiederholt (2009). Optimal Sticky Prices under Rational Inattention. *American Economic Review 99*(3), 769–803.

Maksimovic, V. and G. Phillips (2002). Do Conglomerate Firms Allocate Resources Inefficiently across Industries? Theory and Evidence. *Journal of Finance 57*(2), 721–767.

Maksimovic, V. and G. Phillips (2013). Conglomerate Firms, Internal Capital Markets, and the Theory of the Firm. *Annual Review of Financial Economics 5.*

Malmendier, U. and G. Tate (2005). CEO Overconfidence and Corporate Investment. *Journal of Finance 60*, 2661–2700.

Matějka, F. and A. McKay (2015). Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model. *American Economic Review 105*(1), 272–298.

Moysan, G. and M. Senouci (2016). A Note on 2-Input Neoclassical Production Functions. *Journal of Mathematical Economics 67*, 80–86.

Muth, J. F. (1961). Rational Expectations and the Theory of Price Movements. *Econometrica 29*(3), 315–335.

Oberfield, E. and D. Raval (2021). Micro Data and Macro Technology. *Econometrica 89*(2), 703–732.

Osherson, D., E. Shafir, and E. E. Smith (1994). Extracting the Coherent Core of Human Probability Judgement: A research program for cognitive psychology. *Cognition 50*(1-3), 299–313.

Pesaran, H. and M. Weale (2006). Survey expectations. In G. Elliott, C. Granger, and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Volume 1, Handbook chapter 14, pp. 715–776. Elsevier.

Posner, R. A. (2014). *Economic Analysis of Law* (9th ed.). Aspen Publishing.

Rabin, M. (2002). Inference by Believers in the Law of Small Numbers. *Quarterly Journal of Economics 117*(3), 775–816.

Rabin, M. and G. Weizsäcker (2009). Narrow Bracketing and Dominated Choices. *American Economic Review 99*(4), 1508–1543.

Read, D., G. Loewenstein, and M. Rabin (1999). Choice Bracketing. *Journal of Risk and Uncertainty 19*(1-3), 171–197.

Ruback, R. S. (2004). Know Your Worth: Critical Valuation Errors to Avoid. Transcript, Harvard University.

Schneider, C. and O. Spalt (2014). Conglomerate Investment, Skewness, and the CEO Long-Shot Bias. *Journal of Finance 71*(2), 635–671.

Sen, A. (1993). Internal Consistency of Choice. *Econometrica 61*(3), 495–521.

Sims, C. A. (2003). Implications of Rational Inattention. *Journal of Monetary Economics 50*(3), 665–690.

Stafford, E. and J. L. Heilprin (2011). Valuation of AirThread Connections. Brief Case 4263, Harvard Business School.

Syverson, C. (2011). What Determines Productivity? *Journal of Economic Literature 49*(2), 326–365.

Tentori, K., V. Crupi, and S. Russo (2013). On the Determinants of the Conjunction Fallacy: Probability Versus Inductive Confirmation. *Journal of Experimental Psychology: General 142*(1), 235–255.

Thaler, R. H. (1985). Mental Accounting and Consumer Choice. *Management Science 4*(3), 199–214.

Thaler, R. H. (2018). From Cashews to Nudges: The Evolution of Behavioral Economics. *American Economic Review 108*(6), 1265–1287.

Titman, S. and J. D. Martin (2016). *Valuation: The Art and Science of Corporate Investment Decisions* (3rd ed.). Pearson.

Tversky, A. and D. Kahneman (1971). Belief in the Law of Small Numbers. *Psychological Bulletin 76*(2), 105–110.

Tversky, A. and D. Kahneman (1974). Judgment Under Uncertainty: Heuristics and Biases. *Science 185*(4157), 1124–1131.

Tversky, A. and D. Kahneman (1981). The Framing of Decisions and the Psychology of Choice. *Science 211*(4481), 453–458.

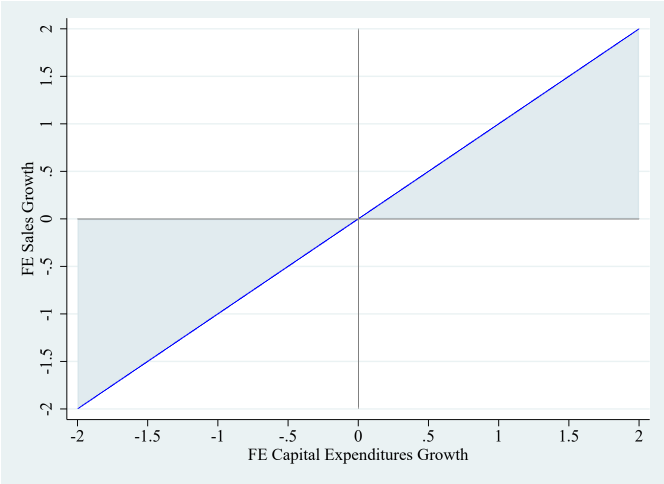
Tversky, A. and D. Kahneman (1983). Extensional Versus Intuitive Reasoning: The conjunction fallacy in probability judgment. *Psychological Review 90*(4), 293–315.

Welch, I. (2017). *Corporate Finance* (4th ed.). Iaw.

Wright, G., G. Rowe, F. Bolger, and J. Gammack (1994). Coherence, Calibration and Expertise in Judgmental Probability Forecasting. *Organizational Behavior and Human Decision Processes 57*(1), 1–25.

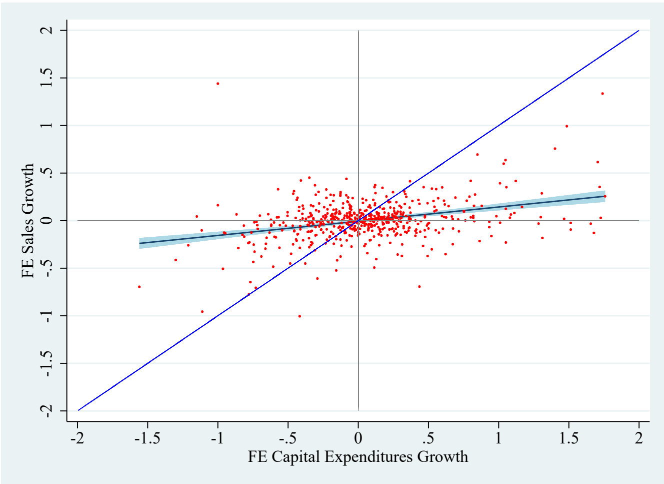
Zhu, J.-Q., P. W. Newall, J. Sundh, N. Chater, and A. N. Sanborn (2022). Clarifying the Relationship Between Coherence and Accuracy in Probability Judgments. *Cognition 223*, 105022.

Figure 1: **Coherence Restrictions on Contemporaneous Forecast Errors**



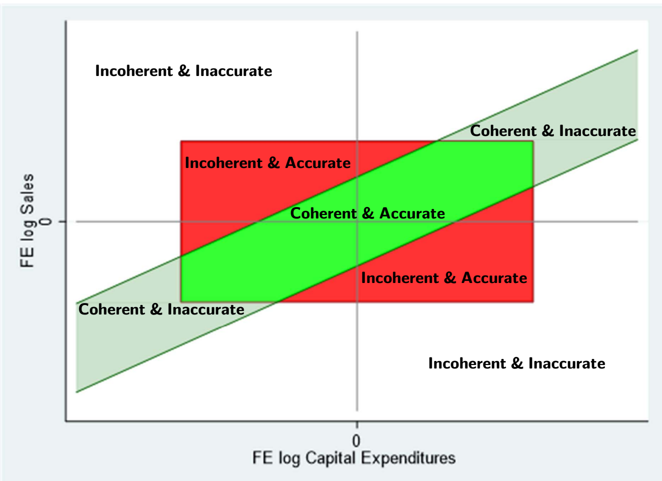
The figure depicts the joint space of forecast errors (FEs) of sales growth and capital expenditures (CapEx) growth, each computed as realization minus forecast. The shaded areas depict the coherence regions, in which FEs are positively associated with one another and they both lie between the 45-degree line and the horizontal axis.

Figure 2: **Contemporaneous Forecasts Errors of Output and Capital**



The figure plots the joint distribution of the forecast errors (FEs) of sales growth and CapEx growth in the data in the same space described in Figure 1 above.

Figure 3: **(In)Coherence and (In)Accuracy Areas**



The figure illustrates the theoretical connection between forecast accuracy and forecast coherence for a hypothetical firm whose capital input has a loading of 1/3. There are four conceptual areas, described in the figure.

Table 1: **Compustat Regressions for Rules of Thumb**

The table reports results from OLS estimations. The dependent variable is CapEx Growth, the annual percent growth rate of capital expenditures. The regressors are Past CapEx Growth, the average of the past two lagged annual percent growth rates of capital expenditures; Sales Growth, the annual percent growth rate of total revenues; and Earnings Growth, the annual percent growth rate of net income. In the first two rows and the last row the sample is the Compustat universe over 1999-2019. In the other rows the sample is the one-digit SIC industry indicated in the first column. CapEx Growth, Sales Growth, and Earnings Growth are winsorized at the 1st and 99th percentiles. Standard errors are clustered at the firm-year level and t-statistics are reported in parentheses. ***, **, and * indicate two-tailed significance at the 1%, 5%, and 10% level, respectively.

	Intercept	Past CapEx Growth	Sales Growth	Earnings Growth	R^2	N Obs
Rule 1 ('narrow bracketing')	0.316***	-0.089***			0.004	74,413
	(0.041)	(0.016)				
Rule 3 ('economies of scale')	0.217***		1.055***		0.081	100,441
	(0.025)		(0.036)			
Rule 4 ('industry based')						
SIC 0	0.330***		2.050***		0.097	358
	(0.156)		(0.344)			
SIC 1	0.243***		0.950***		0.115	8,983
	(0.045)		(0.072)			
SIC 2	0.243***		0.859***		0.050	14,777
	(0.026)		(0.054)			
SIC 3	0.186***		1.188***		0.104	24,852
	(0.024)		(0.058)			
SIC 4	0.180***		0.925***		0.064	14,398
	(0.022)		(0.091)			
SIC 5	0.163***		1.281***		0.081	10,266
	(0.027)		(0.121)			
SIC 6	0.402***		0.963***		0.036	7,477
	(0.041)		(0.062)			
SIC 7	0.202***		1.162***		0.105	14,673
	(0.039)		(0.090)			
SIC 8	0.198***		1.216***		0.088	3,911
	(0.023)		(0.128)			
SIC 9	0.222***		1.288***		0.123	746
	(0.058)		(0.182)			
Rule 5	0.217***		1.042***	0.018***	0.082	100,040
	(0.025)		(0.036)	(0.004)		

Table 2: **Minimum Distance of Sales Forecasts from Rules of Thumb**

The top part of Panel A reports selected statistics on the orthogonal distance between the CFO forecast of sales growth and the forecast of sales growth implied by the five rules of thumb. The first column reports distance statistics for the CFOs whose forecast is closest to the forecast implied by R1. The second column reports distance statistics for the CFOs whose forecast is closest to the forecast implied by R2. The third column reports distance statistics for the CFOs whose forecast is closest to the forecast implied by R3. The fourth column reports distance statistics for the CFOs whose forecast is closest to the forecast implied by R4. The fifth column reports statistics for the CFOs whose forecast is closest to the forecast implied by R5. The last column reports statistics for the full sample, whereby each CFO's distance is the minimum distance between that CFO's forecast and the forecasts implied by the five rules of thumb. The bottom part of Panel A reports the absolute and relative frequencies of CFOs whose forecast of sales growth is closest to the corresponding rule of thumb by column. Panel B reports the same statistics, assigning to a residual category, R6, the CFOs whose actual forecasts are outside of a $+/-0.05$ interval around the forecasts implied by R1, R2, R3, R4, and R5.

Panel A – Main Categorization R1-R5

	R1	R2	R3	R4	R5	All
Mean	0.058	0.030	0.019	0.031	0.043	0.033
Std. Dev.	0.100	0.064	0.017	0.035	0.069	0.059
Frac. Zeros	0.000	0.268	0.000	0.000	0.000	0.106
P10	0.008	0.000	0.005	0.002	0.003	0.000
P25	0.015	0.000	0.006	0.007	0.008	0.007
P50	0.028	0.014	0.010	0.023	0.023	0.019
P75	0.064	0.035	0.028	0.048	0.043	0.036
P90	0.114	0.071	0.048	0.072	0.089	0.071
N of Observations	30	157	43	107	59	396
Fraction	0.076	0.396	0.109	0.270	0.149	1.000

Panel B – Alternative Categorization Allowing for Residual Category R6

	R1	R2	R3	R4	R5	R6	All
Mean	0.020	0.014	0.016	0.017	0.020	0.107	0.033
Std. Dev.	0.012	0.012	0.015	0.014	0.014	0.106	0.059
Frac. Zeros	0.000	0.321	0.000	0.000	0.000	0.000	0.106
P10	0.006	0.000	0.005	0.002	0.001	0.050	0.000
P25	0.009	0.000	0.006	0.005	0.007	0.063	0.007
P50	0.018	0.014	0.010	0.012	0.018	0.075	0.019
P75	0.028	0.021	0.022	0.027	0.034	0.106	0.036
P90	0.037	0.035	0.042	0.040	0.039	0.141	0.071
N of Observations	20	131	41	83	47	74	396
Fraction (of R1-R6)	0.050	0.331	0.103	0.210	0.119	0.187	1.000
Fraction (of R1-R5)	0.062	0.407	0.127	0.258	0.146	.	1.000

Table 3: **Incoherence and Rules of Thumb Indicators**

The top part of the table reports results from OLS estimations. The dependent variable is our measure of incoherence, the minimum orthogonal distance between the CFO forecast of sales growth and the forecast of sales growth implied by Rule 5. The regressors are indicator variables. Rule 1 ('narrow bracketing') equals one if the CFO's sales growth forecast is closest to the forecast implied by R1. Rule 2 ('sales anchoring') equals one if the CFO's sales growth forecast is closest to the forecast implied by R2. Rule 3 ('economies of scale') equals one if the CFO's sales growth forecast is closest to the forecast implied by R3. Rule 4 ('industry based') equals one if the CFO's sales growth forecast is closest to the forecast implied by R4. Rule 5 is the excluded category. t-statistics are reported in parentheses. ***, **, and * indicate two-tailed significance at the 1%, 5%, and 10% level, respectively. The bottom part of the table reports selected summary statistics of the dependent variable.

	(1)	(2)	(3)	(4)	(5)
Rule 1 ('narrow bracketing')	0.081***				0.104***
	(0.014)				(0.016)
Rule 2 ('sales anchoring')		0.039***			0.053***
		(0.008)			(0.011)
Rule 3 ('economies of scale')			-0.055***		-0.020
			(0.012)		(0.014)
Rule 4 ('industry based')				-0.027***	0.010
				(0.009)	(0.012)
Constant	0.066***	0.057***	0.079***	0.080***	0.043***
	(0.004)	(0.005)	(0.004)	(0.005)	(0.009)
R^2	0.071	0.057	0.045	0.023	0.175
N observations	396	396	396	396	396

Summary Statistics of the dependent variable	
Mean	0.073
Std. Dev.	0.079
P10	0.012
Median	0.059
P90	0.139

Table 4: **Incoherence and CFO & Firm Characteristics**

The top part of the table (Panel A) reports descriptive statistics on selected CFO and firm characteristics. The bottom part of the table (Panel B) reports results from OLS estimations showing how our measure of incoherence relates to CFO and firm characteristics. The dependent variable is our measure of incoherence, the minimum orthogonal distance between the CFO forecast of sales growth and the forecast of sales growth implied by Rule 5. CFO has MBA is an indicator equal to one if the CFO has an MBA. Age is the age of CFOs expressed in years. Age 40- is an indicator equal to one if age is lower than 40. Age 41-50 is an indicator equal to one if age is between 41 and 50. Age 51-60 is an indicator equal to one if age is between 51 and 60. Tenure is the number of years the CFO has been with their current firm. Tenure > Median is an indicator equal to one if the CFO's tenure is above the sample median. Gender is an indicator equal to one if the CFO is female. Miscalibration is the negative imputed individual volatility, computed as the mean of the imputed individual volatility within each survey date subtracted from the imputed individual volatility; the result is then multiplied by -1 and divided the standard deviation of the imputed individual volatilities within the survey date. Optimism is a standardized expected return measure, computed by subtracting the mean within each survey date from the expected return estimate and dividing by the standard deviation within the survey date. Both miscalibration and optimism apply the same procedure to Short-Term (one-year) and Long-Term (10-year) forecasts. Firm size is the log of total assets. Market-to-book is the book value of assets plus the market value of common equity minus the book value of common equity and deferred taxes divided by total assets. Dividends is an indicator equal to one if common dividend is strictly larger than zero. One-digit SIC industry fixed effects and survey fixed effects are included but their coefficients are not reported. t-statistics are reported in parentheses. ***, **, and * indicate two-tailed significance at the 1%, 5%, and 10% level, respectively.

Panel A – Summary Statistics

	Mean	Std. Dev.	Q10	Median	Q90	N Obs.
CFO has MBA	0.452	0.498	0.000	0.000	1.000	396
Age	50.43	6.937	42.00	50.00	60.00	396
Tenure	4.273	4.099	0.000	3.000	9.000	396
Gender	0.088	0.284	0.000	0.000	0.000	396
Miscalibration ST	0.035	0.920	-1.166	0.329	0.985	360
Optimism ST	0.052	0.981	-0.918	-0.077	1.285	373
Miscalibration LT	0.039	0.979	-1.095	0.262	0.917	362
Optimism LT	0.033	1.088	-1.008	-0.078	1.077	374
Firm size	7.829	2.296	4.898	7.805	10.61	396
Market-to-Book	1.685	0.900	0.998	1.393	2.731	364
Dividends	0.636	0.482	0.000	1.000	1.000	396

Table 4: **Incoherence and CFO & Firm Characteristics** (Continued)

Panel B – Incoherence and CFO Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
CFO has MBA	0.005					0.007
	(0.009)					(0.011)
Tenure > Median	0.008					0.006
	(0.008)					(0.011)
Age 40-		-0.011				-0.025
		(0.022)				(0.029)
Age 41-50		-0.027*				-0.038**
		(0.016)				(0.019)
Age 51-60		-0.024				-0.030*
		(0.017)				(0.017)
Gender		0.002				0.005
		(0.010)				(0.012)
Miscalibration ST			-0.012			-0.014
			(0.008)			(0.010)
Optimism ST			-0.012			-0.010
			(0.007)			(0.009)
Miscalibration LT				-0.005		-0.002
				(0.004)		(0.006)
Optimism LT				0.001		0.005
				(0.004)		(0.007)
Firm size					-0.006*	-0.005*
					(0.003)	(0.003)
Market-to-Book					0.011	0.014
					(0.013)	(0.014)
Dividends					-0.015	-0.020
					(0.012)	(0.013)
Constant	0.043*	0.078***	0.052*	0.046**	0.137**	0.159***
	(0.022)	(0.025)	(0.027)	(0.021)	(0.036)	(0.049)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.111	0.112	0.154	0.126	0.162	0.227
N of Observations	396	396	360	362	364	332

Table 5: **Regression Tests of Accuracy and Coherence**

The table reports results from OLS estimations. The dependent variable in column (1) is the forecast error of log capex growth. The dependent variable in columns (2), (3), and (4) is the forecast error of log sales growth. The forecast error of log capex (re. sales) growth is constructed using the variance-adjusted approximation, $\mathbb{E}_t \log x_{t+1} \approx \log \mathbb{E}_t x_{t+1} - \frac{1}{2} V_t \log x_{t+1}$. In turn, the conditional variance is computed under an AR1 regression of capex (re. sales) growth at the one-digit SIC industry level under stationarity. Hence, the dependent variable equals the realized log growth minus this variance-adjusted forecast. The forecast error of log wages growth is computed using a similar procedure. Robust standard errors are reported in parentheses. ***, **, and * indicate two-tailed significance at the 1%, 5%, and 10% level, respectively.

	FE of Log CapEx Growth	**FE of Log Sales Growth**			
	(1)	(2)	(3)	(4)	(5)
FE of Log CapEx Growth			0.113*	0.135***	
			(0.063)	(0.032)	
FE of Log Wages Growth			0.023		0.019
			(0.309)		(0.321)
Constant	-0.042*	-0.009	0.046**	-0.004	0.033
	(0.025)	(0.009)	(0.023)	(0.009)	(0.022)
R^2	0.000	0.000	0.108	0.127	0.018
N observations	359	359	51	359	52

Table 6: **Rules of Thumb and Predictable Forecast Errors**

The table reports results from OLS estimations. The dependent variable is the forecast error of log capex growth, constructed using the variance-adjusted approximation, $\mathbb{E}_t \log x_{t+1} \approx \log \mathbb{E}_t x_{t+1} - \frac{1}{2} V_t \log x_{t+1}$. In turn, the conditional variance of capex growth is computed under an AR1 regression of capex growth at the one-digit SIC industry level under stationarity. Hence, the dependent variable equals realized log capex growth minus the corresponding variance-adjusted forecast. The main independent variables are Capex Growth, the annual percent growth rate of capital expenditures; Sales Growth, the annual percent growth rate of total revenues; Earnings Growth, the annual percent growth rate of net income; and four indicator variables. Rule 1 ('narrow bracketing') equals one if the CFO's sales growth forecast is closest to the forecast implied by R1. Rule 2 ('sales anchoring') equals one if the CFO's sales growth forecast is closest to the forecast implied by R2. Rule 3 ('economies of scale') equals one if the CFO's sales growth forecast is closest to the forecast implied by R3. Rule 4 ('industry based') equals one if the CFO's sales growth forecast is closest to the forecast implied by R4. Rule 5 is the excluded category. Robust standard errors are reported in parentheses. ***, **, and * indicate two-tailed significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)	(4)
CapEx Growth	-0.202***	-0.227***	0.023	0.005
	(0.075)	(0.077)	(0.108)	(0.111)
Sales Growth		0.268		0.243
		(0.210)		(0.210)
Earnings Growth		0.003		0.003
		(0.005)		(0.005)
CapEx Growth $*$ Rule 1			-0.300	-0.300
			(0.327)	(0.309)
CapEx Growth $*$ Rule 2			-0.352**	-0.354**
			(0.147)	(0.141)
CapEx Growth $*$ Rule 3			-0.070	-0.105
			(0.148)	(0.165)
CapEx Growth $*$ Rule 4			-0.283	-0.275
			(0.205)	(0.193)
Rule 1			0.043	0.059
			(0.107)	(0.106)
Rule 2			0.043	0.060
			(0.068)	(0.070)
Rule 3			0.156*	0.151*
			(0.086)	(0.085)
Rule 4			0.013	0.008
			(0.079)	(0.078)
Constant	-0.003	-0.055	-0.047	-0.074
	(0.023)	(0.030)	(0.063)	(0.066)
R^2	0.090	0.114	0.142	0.161
N of Observations	359	359	359	359

Table 7: **Incoherence and Corporate Performance (Return on Assets)**

The table reports results from OLS estimations. The dependent variable is return on assets, operating income divided by beginning-of-period total assets. The main regressor is our measure of incoherence, the minimum orthogonal distance between the CFO forecast of sales growth and the forecast of sales growth implied by Rule 5. Miscalibration is the negative imputed individual volatility, computed as the mean of the imputed individual volatility within each survey date subtracted from the imputed individual volatility; the result is then multiplied by -1 and divided the standard deviation of the imputed individual volatilities within the survey date. Optimism is a standardized expected return measure, computed by subtracting the mean within each survey date from the expected return estimate and dividing by the standard deviation within the survey date. Both miscalibration and optimism apply the same procedure to Short-Term (one-year) and Long-Term (10-year) forecasts. Firm size is the log of total assets. Market-to-book is the book value of assets plus the market value of common equity minus the book value of common equity and deferred taxes divided by total assets. Dividends is an indicator equal to one if common dividend is strictly larger than zero. One-digit SIC industry fixed effects and survey fixed effects are included but their coefficients are not reported. Standard errors are clustered by firm and bootstrapped following Cameron, Gelbach, and Miller (2009). t-statistics are reported in parentheses. ***, **, and * indicate two-tailed significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Incoherence	-0.377**	-0.378**	-0.360**	-0.396**	-0.399**	-0.386**	-0.317*	-0.307*
	(0.157)	(0.179)	(0.162)	(0.162)	(0.186)	(0.169)	(0.192)	(0.181)
Miscalibration ST		0.003			0.001		-0.001	
		(0.005)			(0.005)		(0.004)	
Optimism ST		0.000			0.000		0.001	
		(0.006)			(0.006)		(0.005)	
Miscalibration LT			0.004			0.002		0.001
			(0.005)			(0.005)		(0.005)
Optimism LT			0.008			0.007		0.009
			(0.006)			(0.006)		(0.006)
Firm size							0.009***	0.009**
							(0.003)	(0.003)
Market-to-Book							0.028**	0.027*
							(0.014)	(0.015)
Dividends							0.022*	0.023*
							(0.012)	(0.013)
Constant	0.069***	0.069***	0.068***	0.054***	0.056***	0.057***	-0.131***	-0.123***
	(0.011)	(0.011)	(0.011)	(0.014)	(0.020)	(0.019)	(0.047)	(0.0471)
Industry FE	No	No	No	Yes	Yes	Yes	Yes	Yes
Survey FE	No	No	No	Yes	Yes	Yes	Yes	Yes
R^2	0.046	0.042	0.047	0.071	0.064	0.068	0.177	0.185
N of CFOs	311	282	284	311	282	284	263	265
N of Firms	277	252	254	277	252	254	235	237
N of Observations	468	423	428	468	423	428	396	401

Table 8: **Rules of Thumb and Corporate Performance (Return on Assets)**

The table reports results from OLS estimations. The dependent variable is return on assets, operating income divided by beginning-of-period total assets. The main regressors are four indicator variables. Rule 1 ('narrow bracketing') equals one if the CFO's sales growth forecast is closest to the forecast implied by R1. Rule 2 ('sales anchoring') equals one if the CFO's sales growth forecast is closest to the forecast implied by R2. Rule 3 ('economies of scale') equals one if the CFO's sales growth forecast is closest to the forecast implied by R3. Rule 4 ('industry based') equals one if the CFO's sales growth forecast is closest to the forecast implied by R4. Rule 5 is the excluded category. Miscalibration is the negative imputed individual volatility, computed as the mean of the imputed individual volatility within each survey date subtracted from the imputed individual volatility; the result is then multiplied by -1 and divided the standard deviation of the imputed individual volatilities within the survey date. Optimism is a standardized expected return measure, computed by subtracting the mean within each survey date from the expected return estimate and dividing by the standard deviation within the survey date. Both miscalibration and optimism apply the same procedure to Short-Term (one-year) and Long-Term (10-year) forecasts. Firm size is the log of total assets. Market-to-book is the book value of assets plus the market value of common equity minus the book value of common equity and deferred taxes divided by total assets. Dividends is an indicator equal to one if common dividend is strictly larger than zero. One-digit SIC industry fixed effects and survey fixed effects are included but their coefficients are not reported. Standard errors are clustered by firm and bootstrapped following Cameron, Gelbach, and Miller (2009). t-statistics are reported in parentheses. ***, **, and * indicate two-tailed significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rule 1 ('narrow bracketing')	-0.057**	-0.061**	-0.059**	-0.051**	-0.059**	-0.055**	-0.053**	-0.051**
	(0.022)	(0.025)	(0.024)	(0.023)	(0.025)	(0.025)	(0.026)	(0.025)
Rule 2 ('sales anchoring')	-0.026*	-0.027*	-0.023	-0.023	-0.028*	-0.024	-0.034	-0.031
	(0.0138)	(0.015)	(0.015)	(0.015)	(0.017)	(0.016)	(0.021)	(0.019)
Rule 3 ('economies of scale')	-0.031*	-0.036*	-0.034*	-0.027	-0.037*	-0.034	-0.047**	-0.045**
	(0.017)	(0.019)	(0.019)	(0.019)	(0.020)	(0.021)	(0.023)	(0.022)
Rule 4 ('industry based')	-0.012	-0.010	-0.010	-0.008	-0.008	-0.007	-0.012	-0.011
	(0.012)	(0.014)	(0.014)	(0.013)	(0.014)	(0.015)	(0.015)	(0.015)
Miscalibration ST		0.001			-0.001		-0.002	
		(0.005)			(0.005)		(0.004)	
Optimism ST		0.001			0.000		0.001	
		(0.006)			(0.005)		(0.005)	
Miscalibration LT			0.003			0.002		0.001
			(0.006)			(0.005)		(0.004)
Optimism LT			0.007			0.006		0.008
			(0.006)			(0.006)		(0.005)
Firm size							0.010***	0.009***
							(0.004)	(0.004)
Market-to-Book							0.028**	0.028*
							(0.014)	(0.015)
Dividends							0.029**	0.030**
							(0.013)	(0.014)
Constant	0.065***	0.066***	0.064***	0.040***	0.045**	0.046*	-0.147***	-0.137***
	(0.011)	(0.012)	(0.013)	(0.015)	(0.019)	(0.028)	(0.046)	(0.050)
Industry FE	No	No	No	Yes	Yes	Yes	Yes	Yes
Survey FE	No	No	No	Yes	Yes	Yes	Yes	Yes
R^2	0.014	0.014	0.019	0.033	0.031	0.034	0.165	0.170
N of CFOs	311	282	284	311	282	284	263	265
N of Firms	277	252	254	277	252	254	235	237
N of Observations	468	423	428	468	423	428	396	401

Table 9: **Rules of Thumb and Corporate Investment**

The table reports results from OLS estimations. The dependent variable is investment, capital expenditures divided by beginning-of-period assets. The main regressors are four indicator variables. Rule 1 ('narrow bracketing') equals one if the CFO's sales growth forecast is closest to the forecast implied by R1. Rule 2 ('sales anchoring') equals one if the CFO's sales growth forecast is closest to the forecast implied by R2. Rule 3 ('economies of scale') equals one if the CFO's sales growth forecast is closest to the forecast implied by R3. Rule 4 ('industry based') equals one if the CFO's sales growth forecast is closest to the forecast implied by R4. Rule 5 is the excluded category. Miscalibration is the negative imputed individual volatility, computed as the mean of the imputed individual volatility within each survey date subtracted from the imputed individual volatility; the result is then multiplied by -1 and divided the standard deviation of the imputed individual volatilities within the survey date. Optimism is a standardized expected return measure, computed by subtracting the mean within each survey date from the expected return estimate and dividing by the standard deviation within the survey date. Both miscalibration and optimism apply the same procedure to Short-Term (one-year) and Long-Term (10-year) forecasts. Firm size is the log of total assets. Market-to-book is the book value of assets plus the market value of common equity minus the book value of common equity and deferred taxes divided by total assets. Dividends is an indicator equal to one if common dividend is strictly larger than zero. One-digit SIC industry fixed effects and survey fixed effects are included but their coefficients are not reported. Standard errors are clustered by firm and bootstrapped following Cameron, Gelbach, and Miller (2009). t-statistics are reported in parentheses. ***, **, and * indicate two-tailed significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)
Rule 1 ('narrow bracketing')	-0.016	-0.014	-0.015
	(0.011)	(0.011)	(0.012)
Rule 2 ('sales anchoring')	-0.013**	-0.015**	-0.012
	(0.006)	(0.007)	(0.008)
Rule 3 ('economies of scale')	-0.007	-0.011	-0.010
	(0.008)	(0.010)	(0.010)
Rule 4 ('industry based')	-0.003	-0.003	-0.003
	(0.007)	(0.008)	(0.008)
Miscalibration ST		0.001	
		(0.003)	
Optimism ST		0.002	
		(0.003)	
Miscalibration LT			0.002
			(0.002)
Optimism LT			0.004*
			(0.002)
Firm size	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)
Market-to-Book	0.006**	0.006*	0.006*
	(0.003)	(0.003)	(0.003)
Dividends	0.000	0.004	0.004
	(0.009)	(0.008)	(0.008)
Constant	0.044*	0.043*	0.050**
	(0.025)	(0.023)	(0.023)
Industry FE	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes
R^2	0.210	0.223	0.230
N of Observations	437	397	402

Table 10: **Change in Corporate Performance and Investment when New CFOs Take Office**
The table reports results from OLS estimations. The dependent variable in columns (1) and (2) is change in ROA, and in columns (3) and (4) is change in investment. ROA is return on assets, operating income divided by beginning-of-period total assets. Investment is capital expenditures divided by beginning-of-period assets. The changes in the dependent variables are computed as the average in the two years after the CFO's taking office minus the average in the two years before the CFO's taking office. The main regressors are incoherence and four indicator variables. Incoherence is the minimum orthogonal distance between the CFO forecast of sales growth and the forecast of sales growth implied by Rule 5. Rule 1 ('narrow bracketing') equals one if the CFO's sales growth forecast is closest to the forecast implied by R1. Rule 2 ('sales anchoring') equals one if the CFO's sales growth forecast is closest to the forecast implied by R2. Rule 3 ('economies of scale') equals one if the CFO's sales growth forecast is closest to the forecast implied by R3. Rule 4 ('industry based') equals one if the CFO's sales growth forecast is closest to the forecast implied by R4. Rule 5 is the excluded category. Firm characteristics (firm size, market-to-book, and dividends), one-digit SIC industry fixed effects, and survey fixed effects are included but their coefficients are not reported. Standard errors are clustered by firm and bootstrapped following Cameron, Gelbach, and Miller (2009). t-statistics are reported in parentheses. ***, **, and * indicate two-tailed significance at the 1%, 5%, and 10% level, respectively.

	Change in ROA		**Change in Investment**	
	(1)	(2)	(3)	(4)
Incoherence	-1.633*		-0.049	
	(0.989)		(0.045)	
Rule 1 ('narrow bracketing')		-0.274		-0.022*
		(0.213)		(0.012)
Rule 2 ('sales anchoring')		-0.000		-0.003
		(0.036)		(0.008)
Rule 3 ('economies of scale')		-0.057		-0.008
		(0.051)		(0.012)
Rule 4 ('industry based')		0.019		0.001
		(0.048)		(0.009)
Firm characteristics	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes
R^2	0.391	0.192	0.024	0.017
N of Observations	142	142	140	140

Web Appendix
With Supplementary Material for
The Coherence Side of Rationality
Theory and evidence from firm plans

Pamela Giustinelli and Stefano Rossi

A. Theoretical Proofs and Extensions

In this section we present our theoretical proofs. In subsection A.1, we report the theoretical justification for our empirical specification of (R5). In subsection A.2, we derive the incoherence measure in equation (1) in the main text. In subsection A.3, we present the full stochastic setting and a set of auxiliary results used in proving Proposition 1 and Implication 1. In subsection A.4, we prove the benchmark-model results stated in subsection 3.1 of the main text. In subsection A.5, we derive the reduced-form underinvestment mechanism in subsection 3.2 of the main text from a structural ex post adjustment problem. In subsection A.6 we develop the CES extension. In subsection A.7, we present our narrow bracketing microfoundation.

A.1 Derivation of Empirical Specification for (R5)

The derivation below follows closely Hayashi (1982)[15] and Abel and Eberly (1994).[16]

Setup. Consider a Cobb-Douglas production function, $Y_t = K_t^{\alpha} L_t^{1-\alpha}$, in which profits are $\Psi_t = Y_t - \omega_t L_t - I_t - \Phi\left(\frac{I_t}{K_t}\right) K_t$, where $\Phi\left(\frac{I_t}{K_t}\right)$ denotes adjustment costs, and the output price is normalized to one ($p_y = 1$). Under perfect competition in factor markets, the labor share satisfies,

$$\omega_t L_t = (1-\alpha) Y_t. \tag{11}$$

Define the first difference in logs as,

$$\Delta \log Z_t = \log Z_{t+1} - \log Z_t, \tag{12}$$

and denote by $dZ_t = \bar{Z}\,\Delta \log Z_t$ the corresponding deviation from the steady state, where $\bar{Z}$ is the non-stochastic steady-state value.

Statement. Around the non-stochastic steady state $(\bar{Y}, \bar{K}, \bar{I}, \bar{L}, \bar{\Psi})$, the following log-linear identity holds:

$$\Delta \log I_t = \beta_Y\, \Delta \log Y_t + \beta_\Psi\, \Delta \log \Psi_t + \beta_L\, \Delta \log L_t, \tag{13}$$

[15] Hayashi, Fumio, 1982, Tobin's Marginal q and Average q: A Neoclassical Interpretation, Econometrica 50(1), 213-224.

[16] Abel, Andrew B., and Janice C. Eberly, 1994, A Unified Model of Investment Under Uncertainty, American Economic Review 84(5), 1369-1384.

where

$$\beta_Y = \frac{1}{\alpha}\left[1 - (1-\alpha)\frac{1}{1+\Phi'\left(\frac{\bar{I}}{\bar{K}}\right)}\frac{\alpha\bar{Y}}{\bar{I}} + \frac{1}{1+\Phi'\left(\frac{\bar{I}}{\bar{K}}\right)}\frac{\bar{\Psi}}{\bar{I}}\right],$$

$$\beta_\Psi = -\frac{1}{1+\Phi'\left(\frac{\bar{I}}{\bar{K}}\right)}\frac{\bar{\Psi}}{\bar{I}}, \tag{14}$$

$$\beta_L = \frac{1-\alpha}{\alpha}\left[-1 + \frac{1}{1+\Phi'\left(\frac{\bar{I}}{\bar{K}}\right)}\left(\frac{\alpha\bar{Y}}{\bar{I}} - \frac{\bar{\Psi}}{\bar{I}}\right)\right].$$

Proof. Start by expressing profits per unit of capital and using perfect competition in labor markets,

$$\frac{\Psi_t}{K_t} = \alpha\frac{Y_t}{K_t} - \frac{I_t}{K_t} - \Phi\left(\frac{I_t}{K_t}\right). \tag{15}$$

Totally differentiating profits per unit of capital around the steady state gives,

$$\Delta\log\left(\frac{I_t}{K_t}\right) = \frac{1}{1+\Phi'\left(\frac{\bar{I}}{\bar{K}}\right)}\left[\frac{\alpha\bar{Y}}{\bar{I}}\Delta\log\left(\frac{Y_t}{K_t}\right) - \frac{\bar{\Psi}}{\bar{I}}\Delta\log\left(\frac{\Psi_t}{K_t}\right)\right]. \tag{16}$$

Under Cobb-Douglas production,

$$\Delta\log K_t = \frac{1}{\alpha}\Delta\log Y_t - \frac{1-\alpha}{\alpha}\Delta\log L_t. \tag{17}$$

Using expression (17) to substitute for $\Delta\log K_t$ in both (16) and (15) yields

$$\Delta\log\left(\frac{Y_t}{K_t}\right) = \left(1-\frac{1}{\alpha}\right)\Delta\log Y_t + \frac{1-\alpha}{\alpha}\Delta\log L_t = \frac{1-\alpha}{\alpha}\left(\Delta\log L_t - \Delta\log Y_t\right), \tag{18}$$

and

$$\Delta\log\left(\frac{\Psi_t}{K_t}\right) = \Delta\log\Psi_t - \frac{1}{\alpha}\Delta\log Y_t + \frac{1-\alpha}{\alpha}\Delta\log L_t. \tag{19}$$

Using these two expressions for $\Delta\log\left(\frac{Y_t}{K_t}\right)$ and $\Delta\log\left(\frac{\Psi_t}{K_t}\right)$ delivers the coefficients β_Y, β_Ψ, and β_L as above and thus proves the statement.
QED

Empirical Implementation. If labor is constant ($\Delta\log L_t = 0$) or grows at a constant rate ($\Delta\log L_t = c$), the labor term is either absent or absorbed by the intercept. Alternatively, one can linearize around a point such that $\beta_L = 0$, which occurs, for example, when adjustment costs are quadratic,

$$\Phi\left(\frac{I_t}{K_t}\right) = \frac{\varphi}{2}\left(\frac{I_t}{K_t} - \delta\right)^2. \tag{20}$$

In that case, at the steady state $\frac{\bar{I}}{\bar{K}} = \delta$,

$$\Phi'\left(\frac{\bar{I}}{\bar{K}}\right) = \left.\frac{\partial\Phi}{\partial(I/K)}\right|_{\bar{I}/\bar{K}} = 0, \tag{21}$$

so that the adjustment-cost function satisfies a local homogeneity condition, $\Phi'(\bar{I}/\bar{K}) = 0$. Consequently,

$$\beta_L = \frac{1-\alpha}{\alpha}\left[-1 + \left(\frac{\alpha\bar{Y} - \bar{\Psi}}{\bar{I}}\right)\right] = 0, \tag{22}$$

and the empirical regression model simplifies to

$$\Delta \log I_t = \beta_0 + \beta_Y \,\Delta \log Y_t + \beta_\Psi \,\Delta \log \Psi_t + \varepsilon_t. \tag{23}$$

A.2 Orthogonal Distance to the Coherence Hyperplane (R5)

In the main text, incoherence is defined as the orthogonal distance between the CFO's forecast of sales growth and the sales-growth forecast implied by the multivariate benchmark rule (R5). We begin from the empirical implementation of (R5) in the investment dimension:

$$\Delta \log I_t = \alpha + \beta \cdot \Delta \log Y_t + \theta \cdot \Delta \log \Psi_t + \varepsilon_t, \tag{24}$$

where $\Delta \log I_t$, $\Delta \log Y_t$, and $\Delta \log \Psi_t$ denote growth rates of capital expenditures, sales, and earnings, respectively.

To define distance in the sales-forecast dimension, we rewrite the same hyperplane with sales growth on the left-hand side. Solving equation (24) for $\Delta \log Y_t$ yields

$$\Delta \log Y_t = -\frac{\alpha}{\beta} + \frac{1}{\beta}\Delta \log I_t - \frac{\theta}{\beta}\Delta \log \Psi_t + \frac{1}{\beta}\varepsilon_t. \tag{25}$$

Hence, the sales-side representation of the same hyperplane is

$$\Delta \log Y_t = \alpha^Y + \beta^Y \cdot \Delta \log I_t + \theta^Y \cdot \Delta \log \Psi_t + \varepsilon_t^Y, \tag{26}$$

where

$$\alpha^Y = -\frac{\alpha}{\beta}, \qquad \beta^Y = \frac{1}{\beta}, \qquad \theta^Y = -\frac{\theta}{\beta}. \tag{27}$$

Analogously, for the estimated coefficients,

$$\widehat{\alpha}^Y = -\frac{\widehat{\alpha}}{\widehat{\beta}}, \qquad \widehat{\beta}^Y = \frac{1}{\widehat{\beta}}, \qquad \widehat{\theta}^Y = -\frac{\widehat{\theta}}{\widehat{\beta}}. \tag{28}$$

Equation (26) can be rewritten in implicit form as

$$y - \beta^Y i - \theta^Y \psi - \alpha^Y = 0, \tag{29}$$

where, to lighten notation, we write

$$y \equiv \Delta \log Y_t, \qquad i \equiv \Delta \log I_t, \qquad \psi \equiv \Delta \log \Psi_t. \tag{30}$$

Let (i_0, ψ_0, y_0) denote a generic forecast point. The standard formula for the orthogonal distance from a point (x_1, x_2, x_3) to a hyperplane $a_1x_1 + a_2x_2 + a_3x_3 + c = 0$ is

$$d = \frac{|a_1x_1 + a_2x_2 + a_3x_3 + c|}{\sqrt{a_1^2 + a_2^2 + a_3^2}}. \tag{31}$$

Applying this formula to equation (A.2.6), with coefficients

$$a_1 = -\beta^Y, \qquad a_2 = -\theta^Y, \qquad a_3 = 1, \qquad c = -\alpha^Y, \tag{32}$$

yields

$$d = \frac{\left| y_0 - \beta^Y i_0 - \theta^Y \psi_0 - \alpha^Y \right|}{\sqrt{1 + (\beta^Y)^2 + (\theta^Y)^2}}. \tag{33}$$

Replacing the generic point (i_0, ψ_0, y_0) with the CFO's forecast vector

$$\left(\Delta\mathbb{F}_{i,t}[I_{i,t+1}], \Delta\mathbb{F}_{i,t}[\Psi_{i,t+1}], \Delta\mathbb{F}_{i,t}[Y_{i,t+1}]\right), \tag{34}$$

and replacing population coefficients with their estimates, we obtain equation (1) in the main text:

$$\text{Incoherence}_{i,t} = \frac{\left| \Delta\mathbb{F}_{i,t}[Y_{i,t+1}] - \widehat{\beta}^Y \Delta\mathbb{F}_{i,t}[I_{i,t+1}] - \widehat{\theta}^Y \Delta\mathbb{F}_{i,t}[\Psi_{i,t+1}] - \widehat{\alpha}^Y \right|}{\sqrt{1 + \left(\widehat{\beta}^Y\right)^2 + \left(\widehat{\theta}^Y\right)^2}}. \tag{35}$$

Hence, the normalization in equation (1) is a standard geometric transformation that converts the numerator into the shortest orthogonal distance from the CFO's forecast point to the hyperplane implied by the multivariate benchmark rule. Importantly, this is exactly the same fitted hyperplane as in equation (24), now expressed in the sales-forecast dimension.

A.3 Full Stochastic Setting and Auxiliary Results

In this section we present our full stochastic setting and establish a number of auxiliary results to help derive Proposition 1 and Implication 1 in the main paper.

Full Setting. Output is produced by a general CES technology (Moysan and Senouci, 2016),

$$y = f(x_1, x_2) = \left(\frac{a}{a+b} x_1^{\xi} + \frac{b}{a+b} x_2^{\xi} \right)^{\frac{a+b}{\xi}}, \tag{36}$$

where $\nu \equiv a + b > 0$ captures returns to scale (constant for $\nu = 1$, increasing for $\nu > 1$, and decreasing for $\nu < 1$), and the elasticity of substitution between x_1 and x_2 is $\chi = \frac{1}{1-\xi}$.[17]

We consider the problem of a CFO who takes $\Delta\mathbb{F}y_{t+1}$, set by the CEO and board of directors, as given, and forecasts the growth rate of input expenditures, $\Delta\mathbb{F}x_{1,t+1}$ and $\Delta\mathbb{F}x_{2,t+1}$, under square loss,

$$\min_{\Delta\mathbb{F}x_{1,t+1},\, \Delta\mathbb{F}x_{2,t+1}} \mathbb{E}\left[(\Delta x^*_{1,t+1} - \Delta\mathbb{F}x_{1,t+1})^2 + (\Delta x^*_{2,t+1} - \Delta\mathbb{F}x_{2,t+1})^2 \,\middle|\, \Omega_t \right],$$

where Ω_t denotes the information set at t, which includes knowledge of the production function in (36), and $\{\Delta x^*_{1,t+1}, \Delta x^*_{2,t+1}\}$ denote the cost-minimizing inputs growth needed to obtain the

[17] This formulation nests a number of widely used specifications. For $\chi \to +\infty$, the inputs are perfect substitutes and the production function is linear; for $\chi \to 0$, there is no substitution and the production function is Leontief; for $\chi = 1$, we have a Cobb-Douglas. The empirical literature suggests as plausible a range $\chi \in (0.5, 1]$ (e.g., Berndt (1976), Oberfield and Raval (2021)), implying $\xi \in (-1, 0]$.

sales growth target, $\Delta \mathbb{F} y_{t+1}$, that is,

$$\{\Delta x^*_{1,t+1}, \Delta x^*_{2,t+1}\} = \underset{\{\Delta x_{1,t+1}, \Delta x_{2,t+1}\}}{\text{argmin}} \quad p_{1,t+1}\bar{x}_{1,t}e^{\Delta x_{1,t+1}} + p_{2,t+1}\bar{x}_{2,t}e^{\Delta x_{2,t+1}} \tag{37}$$

$$\text{s.t.} \quad f(\bar{x}_{1,t}e^{\Delta x_{1,t+1}}, \bar{x}_{2,t}e^{\Delta x_{2,t+1}}) \geq \bar{y}_t e^{\Delta \mathbb{F} y_{t+1}}, \tag{38}$$

where $\bar{\eta}_t \cdot e^{\Delta \eta_{t+1}} \equiv \mathbb{F}\eta_{t+1}$ for a generic $\eta \in \{y, x_1, x_2\}$.

At solution, $\mathbb{F}\eta^*_{t+1} = \mathbb{E}\left[\eta_{t+1}|\Omega_t\right] \equiv \mathbb{E}_t\left[\eta_{t+1}\right]$ for $\eta \in \{y, x_1, x_2\}$.

We assume the vector-valued stochastic process $\{\log \mathbb{F} y_{t+1}, \log p_{1,t}, \log p_{2,t}\}_{t\geq 1}$ is a stationary Gaussian process. In particular, we assume that log prices follow an AR(1) process, $\log p_{i,t} = \bar{p}_i + \gamma_i \log p_{i,t-1} + \epsilon_{i,t}$, with $|\gamma_i| < 1$, $\epsilon_{i,t} \sim \mathcal{N}\left(0, \sigma_i^2\left(1-\gamma_i^2\right)\right)$, and $\text{Cov}(\epsilon_{1,t}, \epsilon_{2,t}) = \rho_{1,2}\left(1-\gamma_1\gamma_2\right)$. We also assume that $e_t = \log \mathbb{F} y_{t+1} - \mathbb{E}\left[\log \mathbb{F} y_{t+1} | \log p_{1,t}, \log p_{2,t}\right]$ is an AR(1) process, $e_t = \iota e_{t-1} + \varepsilon_t$, with $|\iota| < 1$, $\varepsilon_t \overset{i.i.d.}{\sim} N(0, \sigma^2(1-\iota^2))$, $\sigma^2 \geq 0$, and $\varepsilon_t \perp (\eta_{t'})_{t'\leq t}$ $\forall \eta$ and t, t'.[18]

Henceforth, let $p^F_{i,t}$ refer to input prices faced by a randomly sampled firm from the cross-sectional dataset available to the CFO/forecaster. This is in contrast to $p_{i,t}$, which reflects the input price faced by the forecasting firm. In general, variables with an F superscript refer to firms in the cross-sectional dataset. Assume that firms in the cross-section face (log) input prices following an AR(1) process $\log p^F_{i,t} = \lambda_i + \rho_i \log p^F_{i,t-1} + \xi_{i,t}$, where $(\xi_{1,t}, \xi_{2,t})$ are *i.i.d* Gaussian with $\text{Var}\left(\xi_{i,t}\right) = \eta_i^2(1-\rho_i^2)$ and $\text{Cov}\left(\xi_{1,t}, \xi_{2,t}\right) = \eta_{1,2}(1-\rho_1\rho_2)$.

To parameterize the degree of input price co-movement, we model the short-run shock, $\epsilon_{2,t}$, affecting the price of input 2 as a weighted sum of the shocks that affect the price of input 1, $\epsilon_{1,t}$, and an independent Gaussian shock, φ_t:[19]

$$\epsilon_{2t}(\tau) = \epsilon_{1t}\sqrt{\frac{\tau^2\sigma_2^2(\tau)(1-\gamma_2(\tau)^2)}{\sigma_1^2(1-\gamma_1^2)}} + \varphi_t\sqrt{(1-\tau^2)\sigma_2^2(\tau)(1-\gamma_2(\tau)^2)} \quad \varphi_t \overset{i.i.d}{\sim} N(0,1) \tag{39}$$

$$\sigma_2^2(\tau) \overset{\tau\to 1}{\to} \sigma_1^2 \quad \gamma_2(\tau) \overset{\tau\to 1}{\to} \gamma_1 \quad \bar{p}_2(\tau) \overset{\tau\to 1}{\to} \bar{p}_1,$$

where the weights depend on a parameter $\tau \in [0,1]$ that captures how closely the input prices of the forecasting firm co-move. If $\tau = 1$, then the input prices perfectly co-move and are observationally equal, while $\tau = 0$ implies that input prices evolve independently. For $\tau \in (0,1)$, input prices are dependent but do not perfectly co-move. Note that τ simultaneously captures the correlation of input prices and the similarity in their variance. It is possible for input prices to be perfectly correlated but to have imperfect co-movement because their variances differ. To facilitate our analysis, we assume the regularity conditions that, for each t,

[18] These assumptions imply stationarity and are empirically plausible, as we implement most of our analysis over 2001-2007 at the peak of the 'great moderation,' when aggregate volatility was not a concern.

[19] Under this weighted sum approach, the (log) of both input prices still evolve as a VAR(1) process. As $\tau \to 1$, more weight is placed on ϵ_{1t} and thus we allow greater dependence between the shocks that generate both log-prices. This ensures that the log-price of input 2 will more tightly co-move with price of input 1. In contrast, if we let $\tau \to 0$, more weight is placed on the independent shocks φ_t, and the prices co-move to a lesser extent. Under this framework, the parameters governing the variance of input 2, $\sigma_2^2(\tau)$ and $\gamma_2(\tau)$, are arbitrary functions of τ that converge to σ_1^2 and γ_1, respectively, as $\tau \to 1$. The covariance between both input prices is also a function of τ, with:

$$\sigma_{12}(\tau) = \frac{\sqrt{\sigma_1^2(1-\gamma_1^2)\tau^2\sigma_2^2(\tau)(1-\gamma_2(\tau)^2)}}{1-\gamma_1\gamma_2(\tau)}.$$

$\lim_{\tau\to 1,\tau\in[0,1)} \mathbb{E}\left[\log \mathbb{F} y_{t+1} | \log p_{1,t}, \log p_{2,t}\right]$ converges to a random variable Y almost surely, and that $e_t = \log \mathbb{F} y_{t+1} - \mathbb{E}\left[\log \mathbb{F} y_{t+1} | \log p_{1,t}, \log p_{2,t}\right]$ does not depend on τ. Lastly, using a similar construction, we introduce a parameter $\kappa \in [0,1]$, capturing the degree of co-movement between the input prices of a randomly sampled firm from the cross-section.[20]

Auxiliary Results. In this section we derive a number of auxiliary results to help establish Proposition 1 and Implication 1 in the paper.

Lemma 4. *Let Y be a random variable and X a random vector. Then β is the linear projection coefficient of Y on X if and only if there exists ε with $\mathbb{E}[X\varepsilon] = 0$ such that $Y = X^T\beta + \varepsilon$.*

Proof. Suppose β is the linear projection coefficient of Y on X. Then $\beta = \mathbb{E}[XX^{\mathrm{T}}]^{-1}\mathbb{E}[XY]$. Define $\varepsilon = Y - X^{\mathrm{T}}\beta$. Notice that $\mathbb{E}[X\varepsilon] = \mathbb{E}[X(Y - X^{\mathrm{T}}\beta)] = 0$. This proves the "only if" part. Now, suppose that that there exists ε with $\mathbb{E}[X\varepsilon] = 0$ and $Y = X^{\mathrm{T}}\beta + \varepsilon$. Then $\mathbb{E}[XY] = \mathbb{E}[XX^{\mathrm{T}}]\beta$ and $\beta = \mathbb{E}[XX^{\mathrm{T}}]^{-1}\mathbb{E}[XY]$. This proves the "if" part. $\square$

Lemma 5. *Let (τ_n) be an arbitrary sequence satisfying $\tau_n \overset{n\to\infty}{\to} 1$ and $\tau_n \in [0,1)$ for all $n \in \mathbb{N}$. Then for each $t = 1, ..., T$:*

$$\lim_{n\to\infty} \log p_{2,t} = \log p_{1,t} \quad a.s.$$

Proof. The log price of input 2 for the forecaster is the AR(1) process with parameters $\gamma_2(\tau_n), \sigma_2^2(\tau_n), \sigma_{12}(\tau_n)$ and shocks $\epsilon_{2t}(\tau_n)$ given by (39). Let

$$\Pi_n = \sqrt{\frac{\tau^2\sigma_2^2(\tau)(1-\gamma_2(\tau)^2)}{\sigma_1^2(1-\gamma_1^2)}}, \qquad \zeta_n = \sqrt{(1-\tau^2)\sigma_2^2(\tau)(1-\gamma_2(\tau)^2)},$$

with $\sigma_2^2(\tau) \overset{\tau\to 1}{\to} \sigma_1^2 \quad \gamma_2(\tau) \overset{\tau\to 1}{\to} \gamma_1, \bar{p}_2(\tau) \overset{\tau\to 1}{\to} \bar{p}_1$. Then $\Pi_n \to 1$ and $\zeta_n \to 0$ as $n \to \infty$. Next, since $\gamma_2^2(\tau_n) < 1$ for all $n \in \mathbb{N}$, by the $MA(\infty)$ representation of an AR(1) process, we have for each $n \in \mathbb{N}$ and $t = 1, ..., T$,

$$\begin{aligned}
\log p_{2,t} - \log p_{1,t} &= \bar{p}_2(\tau) - \bar{p}_1 + \sum_{j=0}^{\infty} \gamma_2(\tau_n)^j \epsilon_{2,t-j}(\tau_n) - \sum_{j=0}^{\infty} \gamma_1^j \epsilon_{1,t-j} \\
&= \bar{p}_2(\tau) - \bar{p}_1 + \sum_{j=0}^{\infty} \gamma_1^j \tau_n^j \left[\epsilon_{1,t-j}\Pi_n + \varphi_{T-j}\zeta_n\right] - \sum_{j=0}^{\infty} \gamma_1^j \epsilon_{1,t-j} \\
&= \bar{p}_2(\tau) - \bar{p}_1 + \sum_{j=0}^{\infty} \gamma_1^j \epsilon_{1,t-j}[\tau_n^j \Pi_n - 1] + \zeta_n \sum_{j=0}^{\infty} \tau_n^j \gamma_1^j \varphi_{t-j}.
\end{aligned}$$

[20] Formally, we rewrite (39) as follows:

$$\xi_{2t}(\kappa) = \xi_{1t}\sqrt{\frac{\kappa^2\eta_2^2(\kappa)(1-\rho_2(\kappa)^2)}{\eta_1^2(1-\rho_1^2)}} + \varphi_t\sqrt{(1-\kappa^2)\eta_2^2(\kappa)(1-\rho_2(\kappa)^2)}$$

$$\rho_2(\kappa) \overset{\kappa\to 1}{\to} \rho_1, \quad \eta_2^2(\kappa) \overset{\kappa\to 1}{\to} \eta_1^2, \quad \lambda_2(\kappa) \overset{\kappa\to 1}{\to} \lambda_1, \quad \varphi_t \overset{\text{i.i.d.}}{\sim} N(0,1)$$

.

Note that since $\sum_{j=0}^{\infty}|\gamma_1^j| < \infty$, $\sum_{j=0}^{\infty}|\tau_n^j\gamma_1^j| < \infty$, and both $\epsilon_{t-j}, \varphi_{t-j}$ have finite expectation, the infinite sums converge absolutely, almost surely. Next, since $\sum_{j-0}^{\infty}|\tau_n\gamma_1^j\varphi_{t-j}| < \infty$, we have by the triangle inequality that:

$$\begin{aligned}\left|\zeta_n\sum_{j=0}^{\infty}\tau_n^j\gamma_1^j\varphi_{t-j}\right| &\leq |\zeta_n|\sum_{j=0}^{\infty}|\tau_n^j\gamma_1^j||\varphi_{T-j}| \\ &\leq |\zeta_n|\sum_{j=0}^{\infty}|\gamma_1^j||\varphi_{t-j}|.\end{aligned}$$

Thus, since $\zeta_n \to 0$ as $n \to \infty$, it holds that $\lim_{n\to\infty}\zeta_n\sum_{j=0}^{\infty}\tau_n^j\gamma_1^j\varphi_{t-j} = 0$ almost surely. Since $\sum_{j=0}^{\infty}|\gamma_1^j||\epsilon_{t-j}| < \infty$ almost surely and $|\gamma_1^j\epsilon_{1,t-j}[\tau_n^j\Pi_n - 1]| \leq |\gamma_1^j||\epsilon_{t-j}|$ for all j, an application of the dominated convergence theorem (for an integral w.r.t. counting measure) allows us to interchange limit and infinite sum, as follows:

$$\begin{aligned}\lim_{n\to\infty}\sum_{j=0}^{\infty}\gamma_1^j\epsilon_{1,t-j}[\tau_n^j\Pi_n - 1] &= \sum_{j=0}^{\infty}\gamma_1^j\epsilon_{1,t-j}\lim_{n\to\infty}[\tau_n^j\Pi_n - 1] \\ &= 0.\end{aligned}$$

Thus, $\lim_{n\to\infty}|\log p_{2,t} - \log p_{1,t}| = 0$ almost surely for all $t = 1, ..., T$. □

Lemma 6. *Let (c_0, c_1) be the coefficients from a linear projection of $\Delta\bar{x}_{1,T}^F$ on a constant and $\Delta\bar{y}_T^F$. Then, $c_0 = 0$ and*

$$\begin{aligned}c_1 &= \frac{1}{a+b} - \frac{b}{a+b}\frac{\mathrm{Cov}\left(\Delta p_{1,t}^F - \Delta p_{2,t}^F, \Delta\bar{y}_t^F\right)}{\mathrm{Var}\left(\Delta\bar{y}_t^F\right)} \\ &= \frac{1}{a+b} - \frac{b(1-a-b)}{2(a+b)}\frac{\rho_1 a[\eta_{12}(1-\rho_2)^2 - \eta_1^2(1-\rho_1)^2] + \rho_2 b[\eta_2^2(1-\rho_2)^2 - \eta_{12}(1-\rho_1)^2]}{a^2\rho_1^2(1-\rho_1)\eta_1^2 + b^2\rho_2^2(1-\rho_2)\eta_2^2 + ab\rho_1\rho_2(2-\rho_1-\rho_2)}.\end{aligned}$$

We can write $\eta_2 = \eta_2(\kappa), \eta_{12} = \eta_{12}(\kappa), \rho_2 = \rho_2(\kappa)$ and thus $c_1 = c_1(\kappa)$. For any arbitrary sequence (κ_n) satisfying $\kappa_n \overset{n\to\infty}{\to} 1$ and $\kappa_n \in [0,1)$ for all $n \in \mathbb{N}$, it holds that:

$$\lim_{n\to\infty} c_1(\kappa_n) = \frac{1}{a+b}.$$

Proof. Since we assumed that CFOs in the cross-section are rational, optimizing agents, for each t, $\bar{x}_{1,t}^F$ is the level of inputs that matches the cost minimizing level needed to achieve their respective sales target, which is assumed to equal the optimal forecast for the growth rate needed to take current sales revenue to the profit maximizing level in the next period. Thus,

$$\begin{aligned}\Delta\bar{x}_{1,T}^F &= \log\bar{x}_{1,T}^F - \log\bar{x}_{1,T-1}^F \\ &= \frac{1}{a+b}\Delta\bar{y}_{1,T}^F + \frac{b}{a+b}(\Delta p_{2,T}^F - \Delta p_{1,T}^F) \\ &= c_0 + c_1\Delta\bar{y}_T^F + \left(\frac{1}{a+b} - c_1\right)\Delta\bar{y}_{1,T}^F + \frac{b}{a+b}(\Delta p_{2,T}^F - \Delta p_{1,T}^F) - c_0.\end{aligned}$$

Since (c_0, c_1) are the coefficients from a linear projection of $\Delta\bar{x}_{1,T}^F$ on a constant and $\Delta\bar{y}_T^F$, by

Lemma 4, the following hold:

$$\mathbb{E}\left[\left(\frac{1}{a+b}-c_1\right)\Delta\bar{y}^F_{1,T}+\frac{b}{a+b}(\Delta p^F_{2,T}-\Delta p^F_{1,T})-c_0\right]=0, \tag{40}$$

$$\mathbb{E}\left[\left(\left(\frac{1}{a+b}-c_1\right)\Delta\bar{y}^F_{1,T}+\left(\frac{b}{a+b}(\Delta p^F_{2,T}-\Delta p^F_{1,T})\right)-c_0\right)\Delta\bar{y}^F_T\right]=0. \tag{41}$$

Since the firms in the cross section issue sales growth forecasts that match the optimal, profit maximizing forecast, and since the firm implements a growth rate in inputs to achieve this target at minimum cost, it follows that the implemented growth rate in output at time t, $\Delta\bar{y}^F_t=\log\bar{y}^F_t-\log\bar{y}^F_{t-1}$, satisfies:

$$\log\bar{y}^F_t=\mathbb{E}_{t-1}\left[\log y^F_t\right]=-\frac{a}{1-a-b}[\lambda_1+\rho_1\log p^F_{1,t-1}]-\frac{b}{1-a-b}[\lambda_2+\rho_2\log p^F_{2,t-1}]$$

Thus,

$$\Delta\bar{y}^F_t=-\frac{a}{1-a-b}[\rho_1\Delta p^F_{1,t-1}]-\frac{b}{1-a-b}[\rho_2\Delta p^F_{2,t-1}].$$

Hence, $\mathbb{E}\left[\Delta\bar{y}^F_t\right]=0$, and by (40), it holds that $c_0=0$. Rearranging (41), we obtain

$$\begin{aligned}c_1&=\frac{1}{a+b}-\frac{b}{a+b}\frac{\mathrm{Cov}\left(\Delta p^F_{1,t}-\Delta p^F_{2,t},\Delta\bar{y}^F_t\right)}{\mathrm{Var}\left(\Delta\bar{y}_t\right)}\\&=\frac{1}{a+b}-\frac{b(1-a-b)}{2(a+b)}\frac{\rho_1a[\eta_{12}(1-\rho_2)^2-\eta_1^2(1-\rho_1)^2]+\rho_2b[\eta_2^2(1-\rho_2)^2-\eta_{12}(1-\rho_1)^2]}{a^2\rho_1^2(1-\rho_1)\eta_1^2+b^2\rho_2^2(1-\rho_2)\eta_2^2+ab\rho_1\rho_2(2-\rho_1-\rho_2)}.\end{aligned}$$

Now, let (κ) be an arbitrary sequence satisfying $\kappa_n\overset{n\to\infty}{\to}1$ and $\tau_n\in[0,1)$ for all $n\in\mathbb{N}$. As a result, $\rho_2(\kappa_n)\overset{n\to\infty}{\to}\rho_1,\eta_2^2(\kappa_n)\overset{n\to\infty}{\to}\eta_1^2,\eta_{12}(\kappa_n)\overset{n\to\infty}{\to}\eta_1^2$. It follows that $c_1(\kappa_n)\overset{n\to\infty}{\to}\dfrac{1}{a+b}$. This establishes the result. □

Lemma 7. *Suppose that for $t=1,...,T$, CFO forecasts were correct, $\Delta\mathbb{F}x_{i,t+1}=\Delta x_{i,t+1}$ $\forall i$, and the firm implemented the cost minimizing level of inputs needed to achieve the sales target ($\bar{x}_{i,t}=x^*_{i,t},t\leq T,i=1,2$). Since $\Delta x^*_{i,t}=\log x^*_{i,t+1}-\log\bar{x}_{i,t}$, with $\log x_{i,t+1}$ being the cost-minimizing (log) level of input i at time $t+1$ needed to achieve the sales target $\log\mathbb{F}y_{t+1}$,* *we have for each $t = 1, ..., T$:*

$$\Delta x^*_{1,t} = \frac{1}{a+b}\Delta \mathbb{F} y_t + \frac{b}{a+b}(\Delta p_{2,t} - \Delta p_{1,t}) \tag{42}$$

$$\Delta x^*_{2,t} = \frac{1}{a+b}\Delta \mathbb{F} y_t + \frac{b}{a+b}(\Delta p_{1,t} - \Delta p_{2,t}) \tag{43}$$

$$\Delta p_{i,t} = \bar{p}_i - (1-\gamma_i)\log p_{i,t-1} + \epsilon_{i,t} \tag{44}$$

$$\log p_{i,t} = \bar{p}_i \sum_{j=0}^{k-1} \gamma_i^j + \gamma_i^k \log p_{i,t-k} + \epsilon_{i,t} + \gamma_i \epsilon_{i,t-1} + ... + \gamma_i^{k-1}\epsilon_{i,t-k+1} \tag{45}$$

$$\Delta x^*_{1,t+1} - \mathbb{E}_t\left[\Delta x_{1,t+1}\right] = \frac{b}{a+b}(\epsilon_{2,t+1} - \epsilon_{1,t+1}) \tag{46}$$

$$\Delta x^*_{1,t+1} - c\Delta \mathbb{F} y_{t+1} = \left(\frac{1}{a+b} - c\right)\Delta \mathbb{F} y_{t+1} + \frac{b}{a+b}(\Delta p_{2,t+1} - \Delta p_{1,t+1}) \tag{47}$$

$$\begin{aligned} \Delta x^*_{1,t+1} - \Delta \mathbb{F} x^{R1}_{1,t+1} = &\frac{1}{a+b}\left[\Delta \mathbb{F} y_{t+1} - \frac{1}{k}\sum_{j=1}^{k}\Delta \mathbb{F} y_{t+1-j}\right] \\ &+ \frac{b}{a+b}\left[\Delta p_{2,t+1} - \Delta p_{1,t+1} - \frac{1}{k}\sum_{j=1}^{k}(\Delta p_{2,t+1-j} - \Delta p_{1,t+1-j})\right]. \end{aligned} \tag{48}$$

Proof. These identities follow directly from the variables definitions. □

Lemma 8. *For each $t = 1, ..., T$, there exists $(\theta_{0,t}(\tau), \theta_{1,t}(\tau), \theta_{2,t}(\tau))$ such that*

$$\mathbb{E}\left[\log \mathbb{F} y_{t+1} | \log p_{1,t}, \log p_{2,t}\right] = \theta_{0,t}(\tau) + \theta_{1,t}(\tau)\log p_{1,t} + \theta_{2,t}(\tau)\log p_{2,t}.$$

Moreover, it can be shown that:

1. *The coefficients are time invariant. That is, $(\theta_{0,t}(\tau), \theta_{1,t}(\tau), \theta_{2,t}(\tau)) = (\theta_{0,t'}(\tau), \theta_{1,t'}(\tau), \theta_{2,t'}(\tau))$ for all $t, t' = 1, ..., T$. In this case, denote $(\theta_0(\tau), \theta_1(\tau), \theta_2(\tau))$ the time-invariant coefficients.*

2. *Defining $e_t = \log \mathbb{F} y_{t+1} - \mathbb{E}\left[\log \mathbb{F} y_{t+1} | \log p_{1,t}, \log p_{2,t}\right]$ for all $t = 1, ..., T$, we obtain $Cov(e_t, \log p_{i,t+j}) = 0$ for all $i = 1, 2, t = 1, ..., T, j \geq 0$.*

3. *For all $t = 1, ..., T$:*

$$\log \mathbb{F} y_{t+1} - \log \mathbb{F} y_t = \theta_1(\tau)\Delta p_{1,t} + \theta_2(\tau)\Delta p_{2,t} + e_t - e_{t-1}.$$

4. *$Cov(e_t, \log p_{i,t\pm k}) = 0$ for all $i = 1, 2, k \geq 0, t = 1, ..., T$.*

5. *$\lim_{\tau\to 1, \tau\in[0,1)}(\theta_0(\tau), \theta_1(\tau), \theta_2(\tau))$ exists. Denote $(\tilde{\theta}_0, \tilde{\theta}_1, \tilde{\theta}_2)$ this limit.*

Proof. By assumption, the vector valued stochastic process $(\log \mathbb{F} y_{t+1}, \log p_{1,t}, \log p_{2,t})$ is a jointly stationary Gaussian process. Thus, the expectation of $\log \mathbb{F} x_{1,t+1}$ conditional on log prices is linear in $(\log p_{1,t}, \log p_{2,t})$:

$$\mathbb{E}\left[\log \mathbb{F} y_{t+1} | \log p_{1,t}, \log p_{2,t}\right] = \text{Cov}\left(\log \Delta \mathbb{F} y_{t+1}, \begin{bmatrix}\log p_{1,t} \\ \log p_{2,t}\end{bmatrix}\right)\text{Var}\left(\begin{bmatrix}\log p_{1,t} \\ \log p_{2,t}\end{bmatrix}\right)^{-1}\begin{bmatrix}\log p_{1,t} \\ \log p_{2,t}\end{bmatrix}.$$

Due to joint stationarity of $(\log \mathbb{F} y_{t+1}, \log p_{1,t}, \log p_{2,t})$, the matrix premultiplying $(\log p_{1,t}, \log p_{2,t})^{\mathrm{T}}$ is time invariant. This proves point (1) of Lemma 8. Next, since $e_t = \log \mathbb{F} x_{1,t+1} - \mathbb{E}\left[\log \mathbb{F} y_{t+1} | \log p_{1,t}, \log p_{2,t}\right]$, we have for $i = 1, 2$:

$$\mathrm{Cov}\left(e_t, \log p_{i,t}\right) = \mathbb{E}\left[\left(\log p_{i,t} - \frac{\bar{p}_i}{1-\gamma_i}\right) e_t\right] = \mathbb{E}\left[\left(\log p_{i,t} - \frac{\bar{p}_i}{1-\gamma_i}\right) \mathbb{E}\left[e_t | \log p_{1,t}, \log p_{2,t}\right]\right] = 0.$$

Next, notice that for $k \geq 0$:

$$\mathrm{Cov}\left(e_t, \log p_{i,t+k}\right) = \gamma_i^k \,\mathrm{Cov}\left(e_t, \log p_{i,t}\right) + \sum_{j=0}^{k-1} \gamma_i^j \,\mathrm{Cov}\left(e_t, \epsilon_{i,t+k-j}\right) = 0.$$

This establishes point (2) of Lemma 8. Next, for each t we can write:

$$\log \mathbb{F} x_{1,t+1} = \mathbb{E}\left[\log \mathbb{F} y_{t+1} | \log p_{1,t}, \log p_{2,t}\right] + e_t = \theta_{0,t} + \theta_{1,t} \log p_{1,t} + \theta_{2,t} \log p_{2,t} + e_t.$$

First-differencing establishes point (3) of Lemma 8. Next, for $k \geq 0$

$$\mathrm{Cov}\left(e_t, \log p_{i,t-k}\right) = \iota^j \,\mathrm{Cov}\left(e_{t-k}, \log p_{i,t-k}\right) + \sum_{j=0}^{k-1} \iota^j \,\mathrm{Cov}\left(\varepsilon_{t-j}, \log p_{i,t-k}\right) = 0.$$

This establishes point (4) of Lemma 8. Next, let (τ_n) be an arbitrary sequence satisfying $\tau_n \overset{n\to\infty}{\to} 1$ and $\tau_n \in [0, 1)$ for all $n \in \mathbb{N}$. Define $\theta_n = (\theta_0(\tau_n), \theta_1(\tau_n), \theta_2(\tau_n))$. Suppose by contradiction that θ_n is unbounded. Thus, there exists a subsequence $\theta_{\phi(j)}$ whose norm tends to infinity $||\theta_{\phi(j)}|| \to \infty$ as $j \to \infty$. Define $v_{\phi(j)} = \theta_{\phi(j)}/||\theta_{\phi(j)}||$. Then $v_{\phi(j)}$ is bounded and there exists a sub-subsequence $v_{\phi(\psi(k))}$ such that $v_{\phi(\psi(k))} \to v$ for some $v \in \mathbb{R}$. Note that $\theta_{\phi(\psi(k))}$ is a subsequence of $\theta_{\phi(j)}$. Thus, since $||\theta_{\phi(j)}|| \to \infty$ as $j \to \infty$, it also follows that $||\theta_{\phi(\psi(k))}|| \to \infty$ as $k \to \infty$. This implies that:

$$\begin{aligned}
\mathbb{E}\left[\log \mathbb{F} y_{t+1} | \log p_{1,t}, \log p_{2,t}\right] &= \theta_{\phi(\psi(k))}^{\mathrm{T}} \begin{bmatrix} 1 \\ \log p_{1,t} \\ \log p_{2,t} \end{bmatrix} \\
&= ||\theta_{\phi(\psi(k))}|| v_{\theta_{\phi(\psi(k))}}^{\mathrm{T}} \begin{bmatrix} 1 \\ \log p_{1,t} \\ \log p_{2,t} \end{bmatrix} \overset{k\to\infty}{\to} \infty \cdot \mathrm{sign}\left(v^{\mathrm{T}} \begin{bmatrix} 1 \\ \log p_{1,t} \\ \log p_{1,t} \end{bmatrix}\right).
\end{aligned}$$

By assumption, $\lim_{n\to\infty} \mathbb{E}\left[\log \mathbb{F} y_{t+1} | \log p_{1,t}, \log p_{2,t}\right]$ converges to a real valued random variable Y almost surely. Thus, every subsequence should also converge to the same random variable almost surely. This yields a contradiction. Thus, it must be that θ_n is bounded. Now suppose that θ_n does not converge. Since it is bounded, there exists two subsequences $\theta_{\phi(j)}$ and $\theta_{\psi(j)}$ that converge to two different limits $\theta_{\phi(j)} \to \xi_1$ and $\theta_{\psi(j)} \to \xi_2$ as $j \to \infty$ with $\xi_1 \neq \xi_2$. By the assumption that $\lim_{n\to\infty} \mathbb{E}\left[\log \mathbb{F}_{y,t+1} | \log p_{1,t}, \log p_{2,t}\right]$ converges to a real valued random variable, every subsequence should converge to the same random variable almost surely. Thus, it must be that the following holds almost surely:

$$(\theta_{\phi(j)} - \theta_{\psi(j)})^{\mathrm{T}} \begin{bmatrix} 1 \\ \log p_{1,t} \\ \log p_{2,t} \end{bmatrix} \overset{j\to\infty}{\to} 0.$$

Since

$$\theta_{\phi(j)}^{\mathrm{T}} \begin{bmatrix} 1 \\ \log p_{1,t} \\ \log p_{2,t} \end{bmatrix} \to \xi_1^{\mathrm{T}} \begin{bmatrix} 1 \\ \log p_{1,t} \\ \log p_{1,t} \end{bmatrix} \qquad \theta_{\psi(j)}^{\mathrm{T}} \begin{bmatrix} 1 \\ \log p_{1,t} \\ \log p_{2,t} \end{bmatrix} \to \xi_2^{\mathrm{T}} \begin{bmatrix} 1 \\ \log p_{1,t} \\ \log p_{1,t}, \end{bmatrix}$$

the following holds almost surely:

$$(\xi_1 - \xi_2)^{\mathrm{T}} \begin{bmatrix} 1 \\ \log p_{1,t} \\ \log p_{1,t} \end{bmatrix} = 0.$$

Assuming $\xi_1 \neq \xi_2 \neq 0$ would imply that $\log p_{1,t} = 0$ almost surely, which is a contradiction since $\log p_{1,t}$ has a Gaussian distribution with non-zero variance. This implies that θ_n is a convergent sequence. Now, suppose that there exists two sequences, τ_n and $\tilde{\tau}_n$, satisfying $\tau_n, \tilde{\tau}_n \to 1$ and $\tilde{\tau}_n, \tau_n \in [0,1)$ for all $n \in \mathbb{N}$. From the above discussion, $\theta_n = (\theta_0(\tau_n), \theta_1(\tau_n), \theta_2(\tau_n))$ and $\tilde{\theta}_n = (\theta_0(\tilde{\tau}_n), \theta_1(\tilde{\tau}_n), \theta_2(\tilde{\tau}_n))$ are both convergent sequences. Suppose by contradiction that $\theta_n \to \zeta_1, \tilde{\theta}_n \to \zeta_2$ with $\zeta_1 \neq \zeta_2$. By assumption, there exists a real valued random variable Y such that $\lim_{\tau \to 1, \tau \in [0,1)} \mathbb{E}\left[\log \mathbb{F} y_{t+1} | \log p_{1,t}, \log p_{2,t}\right] = Y$ almost surely. Thus, it must be that the following holds almost surely:

$$\zeta_1^{\mathrm{T}} \begin{bmatrix} 1 \\ \log p_{1,t} \\ \log p_{1,t} \end{bmatrix} = Y = \zeta_2^{\mathrm{T}} \begin{bmatrix} 1 \\ \log p_{1,t} \\ \log p_{1,t} \end{bmatrix} \longleftrightarrow (\zeta_1 - \zeta_2)^{\mathrm{T}} \begin{bmatrix} 1 \\ \log p_{1,t} \\ \log p_{1,t} \end{bmatrix} = 0.$$

Assuming $\zeta_1 \neq \zeta_2$ would imply that $\log p_{1,t} = 0$ almost surely, which is a contradiction since it has a Gaussian distribution with non-zero variance. Thus, it must be that $\zeta_1 = \zeta_2$. Thus, for every sequence (τ_n) satisfying $\tau_n \overset{n\to\infty}{\to} 1$ and $\tau_n \in [0,1)$ for all $n \in \mathbb{N}$, the sequence $\theta_n = (\theta_0(\tau_n), \theta_1(\tau_n), \theta_2(\tau_n))$ converges to the same limit. This implies that $\lim_{\tau \to 1, \tau \in [0,1)} (\theta_0(\tau), \theta_1(\tau), \theta_2(\tau))$ exists. This proves point (5) of Lemma 8. ☐

Lemma 9. *Suppose that for $t = 1, ..., T$, CFO forecasts were correct, $\Delta \mathbb{F} x_{i,t+1} = \Delta x_{i,t+1}\ \forall i$, and the firm implemented the cost minimizing level of inputs needed to achieve the sales target ($\bar{x}_{i,t} = x_{i,t}^*, t \leq T, i = 1,2$). Then, for each $t = 1, ..., T$ and $j \geq 0$:*

$$\begin{aligned} \mathrm{Var}(\Delta \mathbb{F} y_{t+1}) = {} & 2\sigma_1^2 \theta_1(\tau)^2 (1-\gamma_1) + 2\sigma_2^2(\tau)\theta_2(\tau)^2(1-\gamma_2(\tau)) + 2\theta_1(\tau)\theta_2(\tau)\sigma_{12}(2-\gamma_1-\gamma_2(\tau)) \\ & + 2(1-\iota)\sigma^2 \end{aligned} \tag{49}$$

$$\begin{aligned} \mathrm{Cov}\left(\Delta \mathbb{F} y_{t+1}, \Delta \mathbb{F} y_{t-j}\right) = {} & \theta_1(\tau)^2[-\sigma_1^2\gamma_1^j(1-\gamma_1)^2] + \theta_1(\tau)\theta_2(\tau)[-\sigma_{12}(\tau)\gamma_1^j(1-\gamma_1)^2 \\ & - \sigma_{12}(\tau)\gamma_2(\tau)^j(1-\gamma_2(\tau))^2] + \theta_2(\tau)^2[-\sigma_2^2(\tau)\gamma_2(\tau)^j(1-\gamma_2(\tau))^2] \\ & - \sigma^2\iota^j(1-\iota)^2 \end{aligned} \tag{50}$$

$$\mathrm{Cov}\left(\Delta p_{1,t+1}, \Delta \mathbb{F} y_{t+1}\right) = \theta_1(\tau)[-\sigma_1^2(1-\gamma_1)^2] + \theta_2(\tau)[-\sigma_{12}(\tau)(1-\gamma_1)^2] \tag{51}$$

$$\mathrm{Cov}\left(\Delta p_{1,t}, \Delta \mathbb{F} y_{t+1}\right) = \theta_1(\tau)[2\sigma_1^2(1-\gamma_1)] + \theta_2(\tau)[\sigma_{12}(2-\gamma_1-\gamma_2)] \tag{52}$$

$$\mathrm{Cov}\left(\Delta p_{1,t-1-j}, \Delta \mathbb{F} y_{t+1}\right) = \theta_1(\tau)[-\sigma_1^2\gamma_1^j(1-\gamma_1)^2] + \theta_2(\tau)[-\sigma_{12}(\tau)\gamma_2(\tau)^j(1-\gamma_2(\tau))^2] \tag{53}$$

$$\mathrm{Cov}\left(\Delta p_{2,t+1}, \Delta \mathbb{F} y_{t+1}\right) = \theta_1(\tau)[-\sigma_{12}(\tau)(1-\gamma_2)^2] + \theta_2(\tau)[-\sigma_2^2(1-\gamma_2(\tau)^2)]. \tag{54}$$

Proof. These identities follow directly from the variables definitions. ☐

Lemma 10. *Suppose that for $t = 1, ..., T$, CFO forecasts were correct, $\Delta \mathbb{F} x_{i,t+1} = \Delta x_{i,t+1}$ $\forall i$, and the firm implemented the cost minimizing level of inputs needed to achieve the sales target ($\bar{x}_{i,t} = x^*_{i,t}, t \leq T, i = 1, 2$). Then for each $\Delta x \in \mathbb{R}, j \geq 0, t = 1, ..., T$:*

$$\lim_{\tau \to 1, \tau \in [0,1)} \operatorname{Var}\left(\Delta x^*_{1,t}\right) = \frac{1}{(a+b)^2}[2\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1 + \tilde{\theta}_2)^2 + 2(1-\iota)\sigma^2] \tag{55}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \operatorname{Cov}\left(\epsilon_{1,t} - \epsilon_{2,t}, \Delta x^*_{1,t}\right) = 0 \tag{56}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \operatorname{Cov}\left(\epsilon_{1,t} - \epsilon_{2,t}, \Delta x^*_{1,t-1}\right) = 0 \tag{57}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \mathbb{E}\left[\epsilon_{i,t} | \Delta x^*_{1,t-1} = \Delta x\right] = 0 \qquad i = 1, 2 \tag{58}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \mathbb{E}\left[\epsilon_{i,t} | \Delta x^*_{1,t} = \Delta x\right] = 0 \qquad i = 1, 2 \tag{59}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \mathbb{E}\left[\Delta p_{i,t} | \Delta x^*_{1,t-1} = \Delta x\right] = \frac{(a+b)(\tilde{\theta}_1 + \tilde{\theta}_2)[-\sigma_1^2\gamma_1(1-\gamma_1)^2]}{2\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1 + \tilde{\theta}_2)^2 + 2(1-\iota)\sigma^2}\Delta x \qquad i = 1, 2 \tag{60}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \mathbb{E}\left[\Delta p_{i,t} | \Delta x^*_{1,t} = \Delta x\right] = \frac{(a+b)[-\sigma_1^2(1-\gamma_1)^2](\tilde{\theta}_1 + \tilde{\theta}_2)}{2\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1 + \tilde{\theta}_2)^2 + 2(1-\iota)\sigma^2}\Delta x \qquad i = 1, 2 \tag{61}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \mathbb{E}\left[\Delta p_{i,t-1} | \Delta x^*_{1,t} = \Delta x\right] = \frac{(a+b)2\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1 + \tilde{\theta}_2)}{2\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1 + \tilde{\theta}_2)^2 + 2(1-\iota)\sigma^2}\Delta x \qquad i = 1, 2 \tag{62}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \mathbb{E}\left[\Delta p_{i,t-2-j} | \Delta x^*_{1,t} = \Delta x\right] = \frac{(a+b)[-\sigma_1^2\gamma_1^j(1-\gamma_1)^2](\tilde{\theta}_1 + \tilde{\theta}_2)}{2\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1 + \tilde{\theta}_2)^2 + 2(1-\iota)\sigma^2}\Delta x \qquad i = 1, 2 \tag{63}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \mathbb{E}\left[\Delta \mathbb{F} y_t | \Delta x^*_{1,t-1} = \Delta x\right] = \frac{(a+b)[-(\tilde{\theta}_1 + \tilde{\theta}_2)^2\sigma_1^2(1-\gamma_1)^2 - \sigma^2(1-\iota)^2]}{2\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1 + \tilde{\theta}_2)^2 + 2(1-\iota)\sigma^2}\Delta x \tag{64}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \mathbb{E}\left[\Delta \mathbb{F} y_t | \Delta x^*_{1,t} = \Delta x\right] = (a+b)\Delta x \tag{65}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \mathbb{E}\left[\Delta \mathbb{F} y_{t-1-j} | \Delta x^*_{1,t} = \Delta x\right] = (a+b)\frac{-(\tilde{\theta}_1 + \tilde{\theta}_2)^2\sigma_1^2\gamma_1^j(1-\gamma_1)^2 - \sigma^2\iota^j(1-\iota)^2}{2\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1 + \tilde{\theta}_2)^2 + 2(1-\iota)\sigma^2}\Delta x \tag{66}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \mathbb{E}\left[\log \mathbb{F} y_{t-1} - \log \mathbb{F} y_{t-k-1} | \Delta x^*_{1,t} = \Delta x\right] =$$

$$= -(a+b)\frac{(\tilde{\theta}_1 + \tilde{\theta}_2)^2\sigma_1^2(1-\gamma_1^k)(1-\gamma_1) + \sigma^2(1-\iota)(1-\iota^k)}{2\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1 + \tilde{\theta}_2)^2 + 2\sigma^2(1-\iota)}\Delta. \tag{67}$$

Proof. These identities follow directly from the variables definitions. □

Lemma 11. *Suppose that for $t = 1, ..., T$, CFO forecasts were correct, $\Delta \mathbb{F} x_{i,t+1} = \Delta x_{i,t+1}$ $\forall i$, and the firm implemented the cost minimizing level of inputs needed to achieve the sales target*

$(\bar{x}_{i,t} = x^*_{i,t}, t \le T, i = 1, 2)$. *Then for each* $\Delta x \in \mathbb{R}, j \ge 0, t = 1, ..., T, i = 1, 2$*:*

$$\lim_{\tau \to 1, \tau \in [0,1)} \mathbb{E}\left[\log p_{2,t-j} - \log p_{1,t-j} | \Delta x^*_{1,t} = \Delta x\right] = 0 \tag{68}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \text{Var}\left(\epsilon_{i,t} | \Delta x^*_{1,t} = \Delta x\right) = \sigma_1^2(1-\gamma_1^2) \tag{69}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \text{Cov}\left(\epsilon_{1,t}, \epsilon_{2,t} \middle| \Delta x^*_{1,t} = \Delta x\right) = \sigma_1^2(1-\gamma_1^2) \tag{70}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \text{Var}\left(\Delta p_{i,t} | \Delta x^*_{1,t} = \Delta x\right) = 2\sigma_1^2(1-\gamma_1) - \frac{(\tilde{\theta}_1 + \tilde{\theta}_2)^2 \sigma_1^4 (1-\gamma_1)^4}{2\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1 + \tilde{\theta}_2)^2 + 2\sigma^2(1-\iota)} \tag{71}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \text{Var}\left(\Delta p_{i,t-1} | \Delta x^*_{1,t} = \Delta x\right) = 2\sigma_1^2(1-\gamma_1) - \frac{(\tilde{\theta}_1 + \tilde{\theta}_2)^2 4\sigma_1^4 (1-\gamma_1)^2}{2\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1 + \tilde{\theta}_2)^2 + 2\sigma^2(1-\iota)} \tag{72}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \text{Var}\left(\Delta p_{i,t-2-j} | \Delta x^*_{1,t} = \Delta x\right) = 2\sigma_1^2(1-\gamma_1) - \frac{(\tilde{\theta}_1 + \tilde{\theta}_2)^2 \sigma_1^4 (1-\gamma_1)^4 \gamma_1^{2j}}{2\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1 + \tilde{\theta}_2)^2 + 2\sigma^2(1-\iota)} \tag{73}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \text{Var}\left(\Delta p_{2,t} - \Delta p_{1,t} \middle| \Delta x^*_{1,t} = \Delta x\right) = 0 \tag{74}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \text{Cov}\left(\Delta p_{2,t} - \Delta p_{1,t}, \Delta \mathbb{F}_{y,t} \middle| \Delta x^*_{1,t} = \Delta x\right) = 0 \tag{75}$$

$$\lim_{\tau \to 1, \tau \in [0,1)} \text{Var}\left(\Delta \mathbb{F}_{y,t} \middle| \Delta x^*_{1,t} = \Delta x\right) = 0. \tag{76}$$

Proof. These identities follow directly from the variables definitions. □

Lemma 12. *Suppose we have large (but not necessarily perfect) co-movement of the forecaster's input prices,* $\tau \to 1$. *For simplicity, assume that in previous time periods* $t \le T$*, the firm implemented the cost minimizing level of inputs needed to achieve the sales target* $(\bar{x}_{i,t} = x^*_{i,t}, t \le T, i = 1, 2)$*, and that the forecast for input 2 is correct,* $\Delta \mathbb{F} x_{2,t+1} = \Delta x_{2,t+1}$. *Let* $\theta \in (0, 1)$ *be an arbitrary constant. Define* $q^{R5}(\theta)$ *as,*

$$q^{R5}(\theta) = \lim_{\tau \to 1, \tau \in [0,1)} \mathbb{P}\left[\Delta x^*_{1,t+1} - \Delta \mathbb{F} x^{R5}_{1,t+1} - \mathbb{A}_1^{R5} > \theta(1 - R_1(m))\Delta x \middle| \Delta x^*_{1,t+1} = \Delta x\right].$$

where $\mathbb{A}_1^{R5}$ *denotes the ex-post adjustment for input 1 when the ex ante forecast used rule R5. That is,* $q^{R5}(\theta)$ *is the probability that a firm using rule R5 will underinvest by more than* $\theta(1 - R_1(m))\Delta x$ *percentage points, when the optimal growth rate in input 1 is* Δx. *Define* $q^{R1}(\theta), q^{R2}(\theta), q^{R3}(\theta)$ *similarly. Henceforth, let* $z \sim \Phi(z)$*, where* $\Phi(\cdot)$ *denotes the CDF of the standard normal distribution. If the optimal growth rate in input 1 is positive* $\Delta x^*_{1,t+1} = \Delta x > 0$*, the following holds:*

1. *(R5): It holds that* $q^{R5}(\theta) = 0$ *for all* $\theta \in (0, 1)$.
2. *(R3): Suppose that* $Corr(\Delta p^F_{1,t} - \Delta p^F_{2,t}, \Delta \bar{y}^F_t) > 0$. *Let* c_1 *be the coefficient on* $\Delta \bar{y}^F_t$ *from a linear projection of* $\Delta \bar{x}^F_{1,t}$ *on a constant and* $\Delta \bar{y}^F_t$. *We have three cases:*
   (a) *If* $a + b < 1$*, then:*

$$q^{R3}(\theta) = \begin{cases} 0 & \text{if } \theta > 1 - c_1(a+b) \\ 1 & \text{if } \theta \le 1 - c_1(a+b). \end{cases}$$

(b) *If $a+b=1$, then for all $\theta \in (0,1)$:*

$$q^{R3}(\theta) = \begin{cases} 0 & \textit{if } \theta > 1 - c_1(a+b) \\ \frac{1}{2} & \textit{if } \theta = 1 - c_1(a+b) \\ 1 & \textit{if } \theta < 1 - c_1(a+b). \end{cases}$$

(c) *If $a+b>1$, then for all $\theta \in (0,1)$:*

$$q^{R3}(\theta) = \begin{cases} 0 & \textit{if } \theta \geq 1 - c_1(a+b) \\ 1 & \textit{if } \theta < 1 - c_1(a+b). \end{cases}$$

3. *(R2): Suppose we have $a+b<1$. Then:*

$$q^{R2}(\theta) = \begin{cases} 0 & \textit{if } \theta > 1 - a - b \\ 1 & \textit{if } \theta \leq 1 - a - b \end{cases}$$

4. *(R1): For all $\theta \in (0,1)$:*

$$q^{R1}(\theta) = \Phi\left(\frac{k\Delta x}{\sqrt{C}}\left[1 - \theta + \frac{1}{k}\frac{\sigma_1^2(1-\gamma_1^k)(1-\gamma_1)(\tilde{\theta}_1+\tilde{\theta}_2)^2 + \sigma^2(1-\iota^k)(1-\iota)}{2\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1+\tilde{\theta}_2)^2 + 2\sigma^2(1-\iota)}\right]\right),$$

where:

$$\begin{aligned} C = {} & (\sigma_1^2)^2(\tilde{\theta}_1+\tilde{\theta}_2)^4(1-\gamma_1)(1-\gamma_1^k)[4-(1-\gamma_1)(1-\gamma_1^k)] \\ & + \sigma_1^2\sigma^2(\tilde{\theta}_1+\tilde{\theta}_2)^2(1-\gamma_1^k)(1-\iota)[4-2(1-\gamma_1)(1-\iota^k)] \\ & + 4\sigma_1^2\sigma^2(\tilde{\theta}_1+\tilde{\theta}_2)^2(1-\iota^k)(1-\gamma_1) + (\sigma^2)(1-\iota^k)(1-\iota)[4-(1-\iota)(1-\iota^k)]. \end{aligned}$$

Proof. Point (1) of Lemma 12. Using the fact that $(\epsilon_{2,t+1}, \epsilon_{1,t+1}, \Delta x_{1,t+1}^*)$ has joint Gaussian distribution,

$$\begin{aligned} \mathbb{P}\left[\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{t+1}^{R5} - \mathbb{A}_1^{R5} > \theta(1-R_1(m))\Delta x_{1,t+1}^* \middle| \Delta x_{1,t+1}^*\right] \\ &= \mathbb{P}\left[(1-R_1(m))[\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1}^{R5}] > \theta(1-R_1(m))\Delta x_{1,t+1}^* \middle| \Delta x_{1,t+1}^*\right] \\ &= \mathbb{P}\left[\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1}^{R5} > \theta\Delta x_{1,t+1}^* \middle| \Delta x_{1,t+1}^*\right] \\ &= \mathbb{P}\left[\frac{b}{a+b}[\epsilon_{2,t+1} - \epsilon_{1,t+1}] > \theta\Delta x_{1,t+1}^* \middle| \Delta x_{1,t+1}^*\right] \\ &= \Phi\left(\frac{\mathbb{E}\left[\epsilon_{2,t+1} - \epsilon_{1,t+1} | \Delta x_{1,t+1}^*\right] - \frac{a+b}{b}\theta\Delta x_{1,t+1}^*}{\mathrm{Var}\left(\epsilon_{2,t+1} - \epsilon_{1,t+1} \middle| \Delta x_{1,t+1}^*\right)^{1/2}}\right). \end{aligned}$$

Next, by (59),

$$\lim_{\tau \to 1, \tau \in [0,1)} \mathbb{E}\left[\epsilon_{2,t+1} - \epsilon_{1,t+1} | \Delta x_{1,t+1}^* = \Delta x\right] = 0.$$

Next, notice that:

$$
\begin{aligned}
\operatorname{Var}\left(\epsilon_{2,t+1}-\epsilon_{1,t+1}\middle|\Delta x^*_{1,t+1}=\Delta x\right) = \operatorname{Var}\left(\epsilon_{1,t+1}\middle|\Delta x^*_{1,t+1}=\Delta x\right) + \operatorname{Var}\left(\epsilon_{2,t+1}\middle|\Delta x^*_{1,t+1}=\Delta x\right)\\
-2\operatorname{Cov}\left(\epsilon_{2,t+1},\epsilon_{1,t+1}\middle|\Delta x^*_{1,t+1}=\Delta x\right).
\end{aligned}
$$

Thus, by (69) and (70), it holds that:

$$
\lim_{\tau\to 1,\tau\in[0,1)}\operatorname{Var}\left(\epsilon_{2,t+1}-\epsilon_{1,t+1}\middle|\Delta x^*_{1,t+1}=\Delta x\right)=0.
$$

For $\Delta x>0$ and $\theta\in(0,1)$, it holds that:

$$
\lim_{\tau\to 1,\tau\in[0,1)}\frac{\mathbb{E}\left[\epsilon_{2,t+1}-\epsilon_{1,t+1}|\Delta x^*_{1,t+1}=\Delta x\right]-\dfrac{a+b}{b}\theta\Delta x}{\operatorname{Var}\left(\epsilon_{2,t+1}-\epsilon_{1,t+1}\middle|\Delta x^*_{1,t+1}=\Delta x\right)^{1/2}}=-\infty.
$$

Since $\lim_{x\to-\infty}\Phi(x)=0$, we have for each $\theta\in(0,1)$ and $\Delta x>0$:

$$
q^{R5}(\theta)=\lim_{\tau\to 1,\tau\in[0,1)}\mathbb{P}\left[\Delta x^*_{1,t+1}-\Delta\mathbb{F}x^{R5}_{1,t+1}-\mathbb{A}^{R5}_1>\theta(1-R_1(m))\Delta x^*_{1,t+1}\middle|\Delta x^*_{1,t+1}=\Delta x\right]=0.
$$

This establishes the result for point (1) of Lemma 12.
Point (2) of Lemma 12. Let $\theta\in(0,1)$. First, note that by Lemma 6, $\operatorname{Corr}(\Delta p^F_{1,t}-\Delta p^F_{2,t},\bar{y}^F_t)>0$ implies $1-c_1(a+b)>0$. Then,

$$
\begin{aligned}
&\mathbb{P}\left[\Delta x^*_{1,t+1}-\Delta\mathbb{F}x^{R3}_{1,t+1}-\mathbb{A}^{R3}_1>\theta(1-R_1(m))\Delta x^*_{1,t+1}\middle|\Delta x^*_{1,t+1}\right]\\
&\quad=\mathbb{P}\left[(1-R_1(m))[\Delta x^*_{1,t+1}-\Delta\mathbb{F}x^{R3}_{1,t+1}]>\theta(1-R_1(m))\Delta x^*_{1,t+1}\middle|\Delta x^*_{1,t+1}\right]\\
&\quad=\mathbb{P}\left[\Delta x^*_{1,t+1}-\Delta\mathbb{F}x^{R3}_{1,t+1}>\theta\Delta x^*_{1,t+1}\middle|\Delta x^*_{1,t+1}\right]\\
&\quad=\Phi\left(\frac{\mathbb{E}\left[\Delta x^*_{1,t+1}-\Delta\mathbb{F}x^{R3}_{1,t+1}|\Delta x^*_{1,t+1}\right]-\theta\Delta x^*_{1,t+1}}{\operatorname{Var}\left(\Delta x^*_{1,t+1}-\Delta\mathbb{F}x^{R3}_{1,t+1}|\Delta x^*_{1,t+1}\right)^{1/2}}\right),
\end{aligned}
$$

where the last equality follows from (47), (42), and the assumption that $(\log\mathbb{F}_{y,t+1},\log p_{1,t},\log p_{2,t})_{t\geq 1}$ is a Gaussian process. Next, by Corollary 1.1, for all Δx:

$$
\lim_{\tau\to 1,\tau\in[0,1)}\mathbb{E}\left[\Delta x^*_{1,t+1}-\Delta\mathbb{F}x^{R3}_{1,t+1}\middle|\Delta x^*_{1,t+1}=\Delta x\right]=(1-c_1(a+b))\Delta x. \tag{77}
$$

Next, by (47),

$$
\begin{aligned}
\operatorname{Var}\left(\Delta x^*_{1,t+1}-\Delta\mathbb{F}x^{R3}_{1,t+1}|\Delta x^*_{1,t+1}\right)&=\left(\frac{1}{a+b}-c\right)^2\operatorname{Var}\left(\Delta\mathbb{F}_{y,t+1}|\Delta x^*_{1,t+1}\right)\\
&+\left(\frac{b}{a+b}\right)^2\operatorname{Var}\left(\Delta p_{2,t+1}-\Delta p_{1,t+1}|\Delta x^*_{1,t+1}\right)\\
&+2\left(\frac{1}{a+b}-c\right)\left(\frac{b}{a+b}\right)\operatorname{Cov}\left(\Delta p_{2,t+1}-\Delta p_{1,t+1},\Delta\mathbb{F}_{y,t+1}\middle|\Delta x^*_{1,t+1}\right).
\end{aligned}
$$

Thus, by (74), (75), and (76), it holds that:

$$\lim_{\tau\to 1,\tau\in[0,1)} \operatorname{Var}\left(\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R3}_{1,t+1} | \Delta x^*_{1,t+1} = \Delta x\right) = 0. \tag{78}$$

There are three cases. First, suppose $\theta < 1 - c_1(a+b)$. Then,

$$\lim_{\tau\to 1,\tau\in[0,1)} \frac{\mathbb{E}\left[\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R3}_{1,t+1} | \Delta x^*_{1,t+1} = \Delta x\right] - \theta\Delta x}{\operatorname{Var}\left(\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R3}_{1,t+1} | \Delta x^*_{1,t+1} = \Delta x\right)^{1/2}} = +\infty.$$

Thus, since $\lim_{x\to\infty}\Phi(x) = 1$, for all $\theta < 1 - c_1(a+b)$:

$$q^{R3}(\theta) = \lim_{\tau\to 1,\tau\in[0,1)} \mathbb{P}\left[\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R3}_{1,t+1} - \mathbb{A}^{R3}_1 > \theta(1 - R_1(m))\Delta x^*_{1,t+1} \middle| \Delta x^*_{1,t+1} = \Delta x\right] = 1.$$

Second, suppose $\theta > 1 - c_1(a+b)$. Then,

$$\lim_{\tau\to 1,\tau\in[0,1)} \frac{\mathbb{E}\left[\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R3}_{1,t+1} | \Delta x^*_{1,t+1} = \Delta x\right] - \theta\Delta x}{\operatorname{Var}\left(\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R3}_{1,t+1} | \Delta x^*_{1,t+1} = \Delta x\right)^{1/2}} = -\infty.$$

Thus, since $\lim_{x\to-\infty}\Phi(x) = 0$, for all $\theta > 1 - c_1(a+b)$:

$$q^{R3}(\theta) = \lim_{\tau\to 1,\tau\in[0,1)} \mathbb{P}\left[\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R3}_{1,t+1} - \mathbb{A}^{R3}_1 > \theta(1 - R_1(m))\Delta x^*_{1,t+1} \middle| \Delta x^*_{1,t+1} = \Delta x\right] = 0.$$

Third and finally, suppose $\theta = 1 - c_1(a+b)$. Notice that by (47) and (42), it holds that:

$$\begin{aligned}
\mathbb{P}\Big[\Delta x^*_{1,t+1} & - \Delta\mathbb{F}x^{R3}_{1,t+1} - \mathbb{A}^{R3}_1 > \theta(1 - R_1(m))\Delta x^*_{1,t+1} \Big| \Delta x^*_{1,t+1}\Big] \\
&= \mathbb{P}\left[\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R3}_{1,t+1} > \theta\Delta x^*_{1,t+1} \middle| \Delta x^*_{1,t+1}\right] \\
&= \mathbb{P}\left[\left(\frac{1}{a+b} - c_1\right)\Delta\mathbb{F}_{y,t+1} + \frac{b}{a+b}(\Delta p_{2,t+1} - \Delta p_{1,t+1}) > \frac{1}{a+b}\left(\frac{1}{a+b} - c\right)\Delta x^*_{1,t+1} \middle| \Delta x^*_{1,t+1}\right] \\
&= \mathbb{P}\left[\left(1 - \frac{1}{a+b}\right)\left[\left(\frac{1}{a+b} - c_1\right)\Delta\mathbb{F}_{y,t+1} + \frac{b}{a+b}(\Delta p_{2,t+1} - \Delta p_{1,t+1})\right] > 0 \middle| \Delta x^*_{1,t+1}\right] \\
&= \Phi\left(\frac{\frac{1}{a+b}(1 - a - b)\mathbb{E}\left[\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R3}_{1,t+1} | \Delta x^*_{1,t+1}\right]}{\operatorname{Var}\left(\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R3}_{1,t+1} | \Delta x^*_{1,t+1}\right)^{1/2}}\right).
\end{aligned}$$

We now consider three subcases when $\theta = 1 - c_1(a+b)$:

1. If $a + b < 1$, then by (77), (78) and the result that $\lim_{x\to+\infty}\Phi(x) = 1$:

$$\lim_{\tau\to 1,\tau\in[0,1)} \Phi\left(\frac{\frac{1}{a+b}(1 - a - b)\mathbb{E}\left[\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R3}_{1,t+1} | \Delta x^*_{1,t+1} = \Delta x\right]}{\operatorname{Var}\left(\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R3}_{1,t+1} | \Delta x^*_{1,t+1} = \Delta x\right)^{1/2}}\right) = 1.$$

   Thus, if $\theta = 1 - c_1(a+b)$ and $a + b < 1$, then,

$$q^{R3}(\theta) = \lim_{\tau\to 1,\tau\in[0,1)} \mathbb{P}\left[\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R3}_{1,t+1} - \mathbb{A}^{R3}_1 > \theta(1 - R_1(m))\Delta x^*_{1,t+1} \middle| \Delta x^*_{1,t+1} = \Delta x\right] = 1.$$

2. If $a+b=1$, then by $\Phi(0)=1/2$:

$$\lim_{\tau\to 1,\tau\in[0,1)} \Phi\left(\frac{\frac{1}{a+b}(1-a-b)\mathbb{E}\left[\Delta x^*_{1,t+1}-\Delta\mathbb{F}x^{R3}_{1,t+1}|\Delta x^*_{1,t+1}=\Delta x\right]}{\mathrm{Var}\left(\Delta x^*_{1,t+1}-\Delta\mathbb{F}x^{R3}_{1,t+1}|\Delta x^*_{1,t+1}=\Delta x\right)^{1/2}}\right)=1/2.$$

Thus, if $\theta=1-c_1(a+b)$ and $a+b=1$, then,

$$q^{R3}(\theta)=\lim_{\tau\to 1,\tau\in[0,1)}\mathbb{P}\left[\Delta x^*_{1,t+1}-\Delta\mathbb{F}x^{R3}_{1,t+1}-\mathbb{A}^{R3}_1>\theta(1-R_1(m))\Delta x^*_{1,t+1}\middle|\Delta x^*_{1,t+1}=\Delta x\right]=1/2.$$

3. If $a+b>1$, then by (77), (78) and $\lim_{x\to-\infty}\Phi(x)=0$:

$$\lim_{\tau\to 1,\tau\in[0,1)} \Phi\left(\frac{\frac{1}{a+b}(1-a-b)\mathbb{E}\left[\Delta x^*_{1,t+1}-\Delta\mathbb{F}x^{R3}_{1,t+1}|\Delta x^*_{1,t+1}=\Delta x\right]}{\mathrm{Var}\left(\Delta x^*_{1,t+1}-\Delta\mathbb{F}x^{R3}_{1,t+1}|\Delta x^*_{1,t+1}=\Delta x\right)^{1/2}}\right)=0.$$

Thus, if $\theta=1-c_1(a+b)$ and $a+b>1$, we have:

$$q^{R3}(\theta)=\lim_{\tau\to 1,\tau\in[0,1)}\mathbb{P}\left[\Delta x^*_{1,t+1}-\Delta\mathbb{F}x^{R3}_{1,t+1}-\mathbb{A}^{R3}_1>\theta(1-R_1(m))\Delta x^*_{1,t+1}\middle|\Delta x^*_{1,t+1}=\Delta x\right]=0$$

This establishes the result for point (2) of Lemma 12.
The proof for point (3) of Lemma 12 is identical to that of point (2), just replace c_1 with 1.

Point (4) of Lemma 12. Let $\theta\in(0,1)$ be arbitrary. By assumption, $\log\bar{x}_{1,t}=\log x_{1,t}$ for all $t\leq T$. Thus,

$$\Delta\mathbb{F}x^{R1}_{1,t+1}=\frac{1}{k}\sum_{j=1}^{k}\Delta\bar{x}_{1,t+1-j}=\frac{1}{k}\sum_{j=1}^{k}\Delta x_{1,t+1-j}=\frac{1}{k}[\log x_{1,t}-\log x_{t-k}].$$

By (48) and the fact that $(\log x_{1,t},\log x_{1,t-k},\Delta x^*_{1,t+1})$ is jointly Gaussian, thus,

$$\begin{aligned}\mathbb{P}&\left[\Delta x^*_{1,t+1}-\Delta\mathbb{F}x^{R1}_{1,t+1}-\mathbb{A}^{R1}_1>\theta(1-R_1(m))\Delta x^*_{1,t+1}\middle|\Delta x^*_{1,t+1}\right]\\
&=\mathbb{P}\left[(1-R_1(m))[\Delta x^*_{1,t+1}-\Delta\mathbb{F}x^{R1}_{1,t+1}]>\theta(1-R_1(m))\Delta x^*_{1,t+1}\middle|\Delta x^*_{1,t+1}\right]\\
&=\mathbb{P}\left[\Delta x^*_{1,t+1}-\Delta\mathbb{F}x^{R1}_{1,t+1}>\theta\Delta x^*_{1,t+1}\middle|\Delta x^*_{1,t+1}\right]\\
&=\mathbb{P}\left[\Delta\mathbb{F}x^{R1}_{1,t+1}<(1-\theta)\Delta x^*_{1,t+1}\right]\\
&=\mathbb{P}\left[\frac{1}{k}(\log x_{1,t}-\log x_{1,t-k})<(1-\theta)\Delta x^*_{1,t+1}\middle|\Delta x^*_{1,t+1}\right]\\
&=\Phi\left(\frac{(1-\theta)\Delta x^*_{1,t+1}-\frac{1}{k}\mathbb{E}\left[\log x_{1,t}-\log x_{1,t-k}|\Delta x^*_{1,t+1}\right]}{\frac{1}{k^2}\mathrm{Var}\left(\log x_{1,t}-\log x_{1,t-k}\middle|\Delta x^*_{1,t+1}\right)^{1/2}}\right).\end{aligned}$$

First, since $\log x_{1,t}$ is the cost-minimizing (log) level of input 1 needed to achieve output target $\log \mathbb{F}_{y,t}$, it holds that:

$$\begin{aligned}\mathbb{E}\left[\log x_{1,t} - \log x_{1,t-k} | \Delta x^*_{1,t+1}\right] = &\frac{1}{a+b}\mathbb{E}\left[\log \mathbb{F}_{y,t} - \log \mathbb{F}_{y,t-k} \middle| \Delta x^*_{1,t+1}\right] \\ &+ \frac{b}{a+b}\mathbb{E}\left[\log p_{2,t} - \log p_{1,t} | \Delta x^*_{1,t+1}\right] \\ &- \frac{b}{a+b}\mathbb{E}\left[\log p_{2,t-k} - \log p_{1,t-k} | \Delta x^*_{1,t+1}\right].\end{aligned}$$

By (67) and (68),

$$\begin{aligned}\lim_{\tau \to 1, \tau \in [0,1)} &\mathbb{E}\left[\log x_{1,t} - \log x_{1,t-k} | \Delta x^*_{1,t+1} = \Delta x\right] \\ &= -\frac{(\tilde{\theta}_1 + \tilde{\theta}_2)^2 \sigma_1^2 (1-\gamma_1^k)(1-\gamma_1) + \sigma^2 (1-\iota)(1-\iota^k)}{2\sigma_1^2 (1-\gamma_1)(\tilde{\theta}_1 + \tilde{\theta}_2) + 2\sigma^2 (1-\iota)} \Delta x.\end{aligned} \tag{79}$$

Moreover, using the result that $\operatorname{Var}(Y|X) = \operatorname{Var}(Y) - \operatorname{Cov}(Y,X)^2 \operatorname{Var}(X)^{-1}$ if (X,Y) are jointly Gaussian, it can be shown that:

$$\begin{aligned}\lim_{\tau \to 1, \tau \in [0,1)} &\operatorname{Var}\left(\log x_{1,t} - \log x_{1,t-k} \middle| \Delta x^*_{1,t+1} = \Delta x\right) \\ &= (\sigma_1^2)^2 (\tilde{\theta}_1 + \tilde{\theta}_2)^4 (1-\gamma_1)(1-\gamma_1^k)[4 - (1-\gamma_1)(1-\gamma_1^k)] \\ &+ \sigma_1^2 \sigma^2 (\tilde{\theta}_1 + \tilde{\theta}_2)^2 (1-\gamma_1^k)(1-\iota)[4 - 2(1-\gamma_1)(1-\iota^k)] \\ &+ 4\sigma_1^2 \sigma^2 (\tilde{\theta}_1 + \tilde{\theta}_2)^2 (1-\iota^k)(1-\gamma_1) + (\sigma^2)(1-\iota^k)(1-\iota)[4 - (1-\iota)(1-\iota^k)].\end{aligned}$$

Thus, by continuity of the CDF of a standard normal distribution $x \mapsto \Phi(x)$, we have for each $\theta \in (0,1)$:

$$q^{R1}(\theta) = \Phi\left(\frac{k\Delta x}{\sqrt{C}}\left[1 - \theta + \frac{1}{k}\frac{\sigma_1^2(1-\gamma_1^k)(1-\gamma_1)(\tilde{\theta}_1 + \tilde{\theta}_2)^2 + \sigma^2(1-\iota^k)(1-\iota)}{2\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1 + \tilde{\theta}_2)^2 + 2\sigma^2(1-\iota)}\right]\right),$$

where

$$\begin{aligned}C = &(\sigma_1^2)^2 (\tilde{\theta}_1 + \tilde{\theta}_2)^4 (1-\gamma_1)(1-\gamma_1^k)[4 - (1-\gamma_1)(1-\gamma_1^k)] \\ &+ \sigma_1^2 \sigma^2 (\tilde{\theta}_1 + \tilde{\theta}_2)^2 (1-\gamma_1^k)(1-\iota)[4 - 2(1-\gamma_1)(1-\iota^k)] \\ &+ 4\sigma_1^2 \sigma^2 (\tilde{\theta}_1 + \tilde{\theta}_2)^2 (1-\iota^k)(1-\gamma_1) + (\sigma^2)(1-\iota^k)(1-\iota)[4 - (1-\iota)(1-\iota^k)].\end{aligned}$$

This establishes the result for point (4) of Lemma 12. □

Lemma 13. *The expected loss associated with using forecast rules (R3) and (R2), respectively, is given by,*

$$\begin{aligned}\mathbb{E}_t\left[L^{R3}_{t+1}\right] = \mathbb{E}_t\left[L^O_{t+1}\right] + \Bigg[&\left(\frac{1}{a+b} - c_1\right)\Delta \mathbb{F}y_{t+1} + \frac{b}{a+b}\mathbb{E}_t\left[\Delta p_{2,t+1} - \Delta p_{1,t+1}\right] \\ &+ \frac{b}{a+b}[\Delta \bar{x}_{2,t} - \Delta x^*_{2,t} - (\Delta \bar{x}_{1,t} - \Delta x^*_{1,t})]\Bigg]^2,\end{aligned} \tag{80}$$

$$\mathbb{E}_t\left[L^{R2}_{t+1}\right] = \mathbb{E}_t\left[L^{O}_{t+1}\right] + \left[\left(\frac{1}{a+b} - 1\right)\Delta\mathbb{F}_{y,t+1} + \frac{b}{a+b}\mathbb{E}_t\left[\Delta p_{2,t+1} - \Delta p_{1,t+1}\right] + \frac{b}{a+b}[\Delta\bar{x}_{2,t} - \Delta x^*_{2,t} - (\Delta\bar{x}_{1,t} - \Delta x^*_{1,t})]\right]^2, \tag{81}$$

where c_1 is the coefficient on $\Delta\bar{y}^F_t$ from a linear projection of $\Delta\bar{x}^F_{1,t}$ on a constant and $\Delta\bar{y}^F_t$, and $\mathbb{E}_t\left[L^O_{t+1}\right]$ denotes expected loss under the optimal forecasting rule. In particular, $\mathbb{E}_t\left[L^{R3}_{t+1}\right] \geq \mathbb{E}_t\left[L^O_{t+1}\right]$ and $\mathbb{E}_t\left[L^{R2}_{t+1}\right] \geq \mathbb{E}_t\left[L^O_{t+1}\right]$.

Proof. First, since $\Delta x^*_{1,t+1}$ is the cost-minimizing growth rate in inputs needed to meet the sales growth target, it holds that:

$$\begin{aligned}\Delta x^*_{1,t+1} &= \frac{1}{a+b}[\log\bar{y}_t + \Delta\mathbb{F}y_{t+1}] + \frac{b}{a+b}[\log p_{2,t+1} - \log p_{1,t+1}] + \frac{b}{a+b}\log\left(\frac{a}{b}\right) - \log\bar{x}_{1,t} \\ &= \frac{1}{a+b}\Delta\mathbb{F}y_{t+1} + \frac{b}{a+b}[\log p_{2,t+1} - \log p_{1,t+1} + \log\bar{x}_{2,t} - \log\bar{x}_{1,t}] + \frac{b}{a+b}\log\left(\frac{a}{b}\right),\end{aligned}$$

where the second inequality follows because $\bar{y}_t = (x_{1,t})^a(x_{2,t})^b$. Thus,

$$\mathbb{E}_t\left[\Delta x^*_{1,t+1}\right] = \frac{1}{a+b}\Delta\mathbb{F}y_{t+1} + \frac{b}{a+b}[\mathbb{E}_t[\log p_{2,t+1}] - \mathbb{E}_t[\log p_{1,t+1}] + \log\bar{x}_{2,t} - \log\bar{x}_{1,t}] + \frac{b}{a+b}\log\left(\frac{a}{b}\right).$$

Next, since $x^*_{i,t}$ is the cost-minimizing level of input i needed to attain the sales growth target, we have,

$$-\log x^*_{2,t} + \log x^*_{1,t} = \log p_{2,t} - \log p_{1,t} + \log\left(\frac{a}{b}\right).$$

Plugging in for $\log(a/b)$, we obtain:

$$\mathbb{E}_t\left[\Delta x^*_{1,t+1}\right] = \frac{1}{a+b}\Delta\mathbb{F}y_{t+1} + \frac{b}{a+b}\left[\mathbb{E}_t[\Delta p_{2,t+1}] - \mathbb{E}_t[\Delta p_{1,t+1}] + \log\left(\frac{\bar{x}_{2,t}}{x^*_{2,t}}\right) - \log\left(\frac{\bar{x}_{1,t}}{x^*_{1,t}}\right)\right].$$

Thus, for all $c \in \mathbb{R}$:

$$\mathbb{E}_t\left[\Delta x^*_{1,t+1}\right] - c\Delta\mathbb{F}y_{t+1} = \left(\frac{1}{a+b} - c\right)\Delta\mathbb{F}y_{t+1} + \frac{b}{a+b}\left[\mathbb{E}_t[\Delta p_{2,t+1}] - \mathbb{E}_t[\Delta p_{1,t+1}] + \log\left(\frac{\bar{x}_{2,t}}{x^*_{2,t}}\right) - \log\left(\frac{\bar{x}_{1,t}}{x^*_{1,t}}\right)\right] \tag{82}$$

Next, notice that:

$$\begin{aligned}\mathbb{E}_t\left[(\Delta x^*_{1,t+1} - c\Delta\mathbb{F}y_{t+1})^2\right] &= \mathbb{E}_t\left[(\Delta x^*_{1,t+1} - \mathbb{E}_t\left[\Delta x^*_{1,t+1}\right] + \mathbb{E}_t\left[\Delta x^*_{1,t+1}\right] - c\Delta\mathbb{F}y_{t+1})^2\right] \\ &= \mathbb{E}\left[L^O_{t+1}\right] + \mathbb{E}\left[(\mathbb{E}_t\left[\Delta x^*_{1,t+1}\right] - c\Delta\mathbb{F}y_{t+1})^2\right],\end{aligned}$$

where the second equality follows because:

$$\begin{aligned}&\mathbb{E}_t\left[(\Delta x^*_{1,t+1} - \mathbb{E}_t\left[\Delta x^*_{1,t+1}\right])\ (\mathbb{E}_t\left[\Delta x^*_{1,t+1}\right] - c\Delta\mathbb{F}y_{t+1})\right] \\ &\quad = (\mathbb{E}_t\left[\Delta x^*_{1,t+1}\right] - c\Delta\mathbb{F}y_{t+1})\mathbb{E}_t\left[\Delta x^*_{1,t+1} - \mathbb{E}_t\left[\Delta x^*_{1,t+1}\right]\right] \\ &\quad = 0.\end{aligned}$$

Thus, by (82),

$$\begin{aligned}&\mathbb{E}_t\left[(\Delta x^*_{1,t+1} - c\Delta\mathbb{F}y_{t+1})^2\right]\\&\quad= \mathbb{E}_t\left[L^O_{t+1}\right] + \left[\left(\frac{1}{a+b} - c_1\right)\Delta\mathbb{F}y_{t+1} + \frac{b}{a+b}\left(\mathbb{E}_t\left[\Delta p_{2,t+1} - \Delta p_{2,t+1}\right] + \log\frac{\bar{x}_{2,t}}{x^*_{2,t}} - \log\frac{\bar{x}_{1,t}}{x^*_{1,t}}\right)\right]^2.\end{aligned}$$

Thus, (80) follows by replacing c with c_1, the coefficient on $\Delta\bar{y}_t$ from a linear projection of $\Delta\bar{x}_{1,t}$ on a constant and $\Delta\bar{y}_t$. The result for R2, (81), follows by replacing c with 1. □

Lemma 14. *The expected loss associated with using forecast rule R1 is given by,*

$$\begin{aligned}\mathbb{E}_t\left[L^{R1}_{t+1}\right] = \mathbb{E}_t\left[L^O_{t+1}\right] + \Bigg(&\frac{1}{a+b}\left[\Delta\mathbb{F}y_{t+1} - \frac{1}{k}\sum_{j=1}^{k}\Delta\mathbb{F}y_{t+1-j}\right] \qquad (83)\\&+ \frac{b}{a+b}\left[\mathbb{E}_t[\Delta p_{2,t+1} - \Delta p_{1,t+1}] - \frac{1}{k}\sum_{j=1}^{k}(\Delta p_{2,t+1-j} - \Delta p_{1,t+1-j})\right]\\&- \frac{b}{a+b}\left[\Delta\bar{x}_{1,t} - \Delta x^*_{1,t} - \frac{1}{k}\sum_{j=1}^{k}(\Delta\bar{x}_{1,t-j} - \Delta x^*_{1,t-j})\right]\\&+ \frac{b}{a+b}\left[\Delta\bar{x}_{2,t} - \Delta x^*_{2,t} - \frac{1}{k}\sum_{j=1}^{k}(\Delta\bar{x}_{2,t-j} - \Delta x^*_{2,t-j})\right]\\&- \frac{1}{k}\sum_{j=1}^{k}(\Delta\bar{x}_{1,t+1-j} - \Delta x^*_{1,t+1-j})\Bigg)^2.\end{aligned}$$

Thus, $\mathbb{E}_t\left[L^{R1}_{t+1}\right] \geq \mathbb{E}\left[L^O_{t+1}\right]$.

Proof. First, notice that:

$$\begin{aligned}\mathbb{E}_t\left[(\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R1}_{1,t+1})^2\right] &= \mathbb{E}_t\left[(\Delta x^*_{1,t+1} - \mathbb{E}_t\left[\Delta x^*_{1,t+1}\right] + \mathbb{E}_t\left[\Delta x^*_{1,t+1}\right] - \Delta\mathbb{F}x^{R1}_{1,t+1})^2\right] \qquad (84)\\&= \mathbb{E}\left[L^O_{t+1}\right] + \mathbb{E}\left[(\mathbb{E}_t\left[\Delta x^*_{1,t+1}\right] - \Delta\mathbb{F}x^{R1}_{1,t+1})^2\right], \qquad (85)\end{aligned}$$

where the second equality follows because

$$\begin{aligned}&\mathbb{E}_t\left[(\Delta x^*_{1,t+1} - \mathbb{E}_t\left[\Delta x^*_{1,t+1}\right])\left(\mathbb{E}_t\left[\Delta x^*_{1,t+1}\right] - \Delta\mathbb{F}x^{R1}_{1,t+1})\right]\right.\\&\qquad= (\mathbb{E}_t\left[\Delta x^*_{1,t+1}\right] - \Delta\mathbb{F}x^{R1}_{1,t+1})\mathbb{E}_t\left[\Delta x^*_{1,t+1} - \mathbb{E}_t\left[\Delta x^*_{1,t+1}\right]\right]\\&\qquad= 0.\end{aligned}$$

Next, notice that for each t, it holds that:

$$\begin{aligned}
\Delta x^*_{1,t+1} &= \log x^*_{1,t+1} - \log \bar{x}_{1,t} \\
&= \frac{1}{a+b}\log \mathbb{F}y_{t+1} + \frac{b}{a+b}[\log p_{2,t+1} - \log p_{1,t+1}] - \log \bar{x}_{1,t} + \frac{b}{a+b}\log\left(\frac{a}{b}\right) \\
&= \frac{1}{a+b}[\Delta \mathbb{F}y_{t+1} + \log \bar{y}_t] + \frac{b}{a+b}[\log p_{2,t+1} - \log p_{1,t+1} + \log\left(\frac{a}{b}\right)] - \log \bar{x}_{1,t} \\
&= \frac{1}{a+b}\Delta \mathbb{F}y_{t+1} + \frac{1}{a+b}[\log \bar{y}_t - \log \mathbb{F}y_t] + \frac{1}{a+b}\log \mathbb{F}y_t \\
&\quad + \frac{b}{a+b}[\log p_{2,t+1} - \log p_{1,t+1} + \log\left(\frac{a}{b}\right)] - \log \bar{x}_{1,t}.
\end{aligned}$$

First, notice that $\log \bar{y}_t = a\log \bar{x}_{1,t} + b\log \bar{x}_{2,t}$ and $\log \mathbb{F}y_t = a\log x^*_{1,t} + b\log x^*_{2,t}$. Thus,

$$\log \bar{y}_{1,t} - \log \mathbb{F}y_t = a\log\left(\frac{\bar{x}_{1,t}}{x^*_{1,t}}\right) + b\log\left(\frac{\bar{x}_{2,t}}{x^*_{2,t}}\right). \tag{86}$$

Moreover, since $x^*_{1,t}$ is the cost-minimizing level of input 1 needed to achieve sales revenue $\mathbb{F}y_t = \bar{y}_{t-1}\exp(\Delta \mathbb{F}y_t)$, it holds that:

$$\log x^*_{1,t} = \frac{1}{a+b}\log \mathbb{F}y_t + \frac{b}{a+b}[\log p_{2,t} - \log p_{1,t}] + \frac{b}{a+b}\log\left(\frac{a}{b}\right). \tag{87}$$

Combining (86) and (87), we obtain:

$$\Delta x^*_{1,t+1} = \frac{1}{a+b}\Delta \mathbb{F}y_{t+1} + \frac{b}{a+b}[\Delta p_{2,t+1} - \Delta p_{1,t+1}] + \frac{b}{a+b}\left[\log\left(\frac{\bar{x}_{2,t}}{x^*_{2,t}}\right) - \log\left(\frac{\bar{x}_{1,t}}{x^*_{1,t}}\right)\right] \quad \forall t. \tag{88}$$

Next, notice that:

$$\begin{aligned}
\Delta \mathbb{F}x^{R1}_{1,t+1} &= \frac{1}{k}\sum_{j=1}^{k}\Delta \bar{x}_{1,t+1-j} \qquad (89)\\
&= \frac{1}{k}\sum_{j=1}^{k}(\log x^*_{1,t+1-j} - \log \bar{x}_{1,t-j} + \log \bar{x}_{1,t+1-j} - \log x^*_{1,t+1-j}) \\
&= \frac{1}{k}\sum_{j=1}^{k}\Delta x^*_{1,t+1-j} + \frac{1}{k}\sum_{j=1}^{k}\log\frac{\bar{x}_{1,t+1-j}}{x^*_{1,t+1-j}}.
\end{aligned}$$

Combining (88) and (89), we obtain:

$$
\begin{aligned}
\Delta x^*_{1,t+1} - \Delta \mathbb{F} x^{R1}_{1,t+1} = & \frac{1}{a+b}\left[\Delta \mathbb{F} y_{t+1} - \frac{1}{k}\sum_{j=1}^{k} \Delta \mathbb{F} y_{t+1-j}\right] \\
& + \frac{b}{a+b}\left[\Delta p_{2,t+1} - \Delta p_{1,t+1} - \frac{1}{k}\sum_{j=1}^{k}(\Delta p_{2,t+1-j} - \Delta p_{1,t+1-j})\right] \\
& - \frac{b}{a+b}\left[\Delta \bar{x}_{1,t} - \Delta x^*_{1,t} - \frac{1}{k}\sum_{j=1}^{k}(\Delta \bar{x}_{1,t-j} - \Delta x^*_{1,t-j})\right] \\
& + \frac{b}{a+b}\left[\Delta \bar{x}_{2,t} - \Delta x_{2,t} - \frac{1}{k}\sum_{j=1}^{k}(\Delta \bar{x}_{2,t-j} - \Delta x^*_{2,t-j})\right] \\
& - \frac{1}{k}\sum_{j=1}^{k}(\Delta \bar{x}_{1,t+1-j} - \Delta x^*_{1,t+1-j}).
\end{aligned}
$$

Plugging into (84) obtains the result. □

A.4 Proofs of Main Results

Throughout the Web Appendix, for any variable η, define its forecast error as $\mathbb{F}\mathbb{E}_t[\eta_{t+1}] \equiv \eta_{t+1} - \mathbb{F}_t[\eta_{t+1}]$. We use this notation consistently for output, inputs, and their logarithmic transformations.

Proof of Lemma 1. Given our assumptions, $y = x_1^a \cdot x_2^b$. Therefore, taking the log-difference,

$$
\begin{aligned}
\log\left[\frac{y_{t+1}}{y_t}\right] &= a \cdot \log\left[\frac{x_{1,t+1}}{x_{1,t}}\right] + b \cdot \log\left[\frac{x_{2,t+1}}{x_{2,t}}\right] \\
\mathbb{E}_t \log\left[\frac{y_{t+1}}{y_t}\right] &= a \cdot \mathbb{E}_t \log\left[\frac{x_{1,t+1}}{x_{1,t}}\right] + b \cdot \mathbb{E}_t \log\left[\frac{x_{2,t+1}}{x_{2,t}}\right].
\end{aligned}
$$

QED

Proof of Lemma 2. The Proof follows immediately from,

$$
\begin{aligned}
\mathbb{F}\mathbb{E}_t[\Delta y_{t+1}] \equiv \Delta \bar{y}_{t+1} - \Delta \mathbb{F}_t[y_{t+1}] &= \Delta \bar{y}_{t+1} - \Delta \mathbb{F} y_{t+1} \\
&= \log \bar{y}_{t+1} - \log \bar{y}_t - \log \mathbb{F} y_{t+1} + \log \bar{y}_t \\
&= \mathbb{F}\mathbb{E}_t[\log y_{t+1}] \\
&= a \cdot \log x_{1,t+1} + b \cdot \log x_{2,t+1} - a \cdot \mathbb{F}_t[\log x_{1,t+1}] - b \cdot \mathbb{F}_t[\log x_{2,t+1}] \\
&= a \cdot \mathbb{F}\mathbb{E}_t[\log x_{1,t+1}] + b \cdot \mathbb{F}\mathbb{E}_t[\log x_{2,t+1}].
\end{aligned}
$$

QED

Proof of Lemma 3. A forecaster can estimate (a, b) using a linear projection of $\Delta \bar{y}_{1,t}$ on a constant, $\Delta \bar{x}_{1,t}$ and $\Delta \bar{x}_{2,t}$. First, since firms in the cross-section are assumed to have the same production technology as the forecasting firm, it follows that:

$$\log \bar{y}_t = a \log \bar{x}_{1,t} + b \log \bar{x}_{2,t}$$

An immediate application of Lemma 4 obtains the result. In particular, the linear projection coefficient on $\bar{x}_{1,t}$ is a and the linear projection coefficient on $\bar{x}_{2,t}$ is b.
QED

Proof of Proposition 1. We first establish that forecast rule (R5) coincides with the optimal forecast. Let (c_0, c_1, c_2) be the coefficients from a linear projection of $\Delta\bar{x}_{1,t}$ on a constant, $\Delta\bar{y}_t$, and $\Delta\bar{x}_{2,t}$. First, since firms in the cross-section have Cobb–Douglas production technology with input elasticities a and b, it must be that,

$$\log \bar{y}_t = a \log \bar{x}_{1,t} + b \log \bar{x}_{2,t}.$$

Thus, we have $\Delta\bar{y}_t = a\Delta\bar{x}_{1,t} + b\Delta\bar{x}_{2,t}$. Rearranging, we obtain,

$$\Delta\bar{x}_{1,t} = \frac{1}{a}\Delta\bar{y}_t - \frac{b}{a}\Delta\bar{x}_{2,t}.$$

By Lemma 4, $c_0 = 0, c_1 = 1/a, c_2 = -b/a$ and, thus,

$$\Delta\mathbb{F}x^{R5}_{1,t+1} = \frac{1}{a}\Delta\mathbb{F}y_{t+1} - \frac{b}{a}\mathbb{E}_t\left[\Delta x_{2,t+1}\right].$$

Next, since the forecasting firm has Cobb–Douglas technology,

$$\begin{aligned}
&\bar{y}_t \exp(\Delta\mathbb{F}y_{t+1}) = (\bar{x}_{1,t}\exp(\Delta x^*_{1,t+1}))^a(\bar{x}_{2,t}\exp(\Delta x^*_{2,t+1}))^b \\
&a\Delta x^*_{1,t+1} + b\Delta x^*_{2,t+1} = \Delta\mathbb{F}y_{t+1},
\end{aligned}$$

where the second equality follows because $(\bar{x}_{1,t})^a(\bar{x}_{2,t})^b = \bar{y}_t$. Thus,

$$\mathbb{E}_t\left[\Delta x^*_{1,t+1}\right] = \frac{1}{a}\Delta\mathbb{F}y_{t+1} - \frac{b}{a}\mathbb{E}_t\left[\Delta x^*_{2,t+1}\right] = \Delta\mathbb{F}x^{R5}_{1,t+1}.$$

This establishes that (R5) coincides with the optimal forecast. We now prove that (R5) is unbiased in the sense that,

$$\lim_{\tau\to 1, \tau\in[0,1)} \mathbb{E}\left[\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R5}_{1,t+1}\right] = 0.$$

First, since $\Delta x^*_{i,t+1}$ is the cost-minimizing growth in inputs needed to attain the sales growth target,

$$\begin{aligned}
\Delta x^*_{1,t+1} &= \frac{1}{a+b}[\log\bar{y}_t + \Delta\mathbb{F}y_{t+1}] + \frac{b}{a+b}[\log p_{2,t+1} - \log p_{1,t+1}] + \frac{b}{a+b}\log\left(\frac{a}{b}\right) - \log\bar{x}_{1,t} \\
&= \frac{1}{a+b}\Delta\mathbb{F}y_{t+1} + \frac{b}{a+b}[\log p_{2,t+1} - \log p_{1,t+1} + \log\bar{x}_{2,t} - \log\bar{x}_{1,t}] + \frac{b}{a+b}\log\left(\frac{a}{b}\right),
\end{aligned}$$

where the second inequality follows because $(\bar{x}_{1,t})^a(\bar{x}_{2,t})^b = \bar{y}_t$. Next, since $x^*_{i,t}$ is the cost-minimizing level of input i needed to attain the sales target,

$$-\log x^*_{2,t} + \log x^*_{1,t} = \log p_{2,t} - \log p_{1,t} + \log\left(\frac{a}{b}\right).$$

Plugging in for $\log(a/b)$, we obtain,

$$\Delta x^*_{1,t+1} = \frac{1}{a+b}\Delta\mathbb{F}y_{t+1} + \frac{b}{a+b}\left[\Delta p_{2,t+1} - \Delta p_{1,t+1} + \log\frac{\bar{x}_{2,t}}{x^*_{2,t}} - \log\frac{\bar{x}_{1,t}}{x^*_{1,t}}\right].$$

Thus,

$$\mathbb{E}_t\left[\Delta x^*_{1,t+1}\right] = \frac{1}{a+b}\Delta\mathbb{F}y_{t+1} + \frac{b}{a+b}\left[\mathbb{E}_t[\Delta p_{2,t+1}] - \mathbb{E}_t[\Delta p_{1,t+1}] + \log\frac{\bar{x}_{2,t}}{x^*_{2,t}} - \log\frac{\bar{x}_{1,t}}{x^*_{1,t}}\right].$$

As a result,

$$\begin{aligned}\Delta x^*_{1,t+1} - \mathbb{E}_t\left[\Delta x^*_{1,t+1}\right] &= \frac{b}{a+b}[\Delta p_{2,t+1} - \mathbb{E}_t[\Delta p_{2,t+1}] - \Delta p_{1,t+1} + \mathbb{E}_t[\Delta p_{1,t+1}]] \\ &= \frac{b}{a+b}[\varepsilon_{2,t+1} - \varepsilon_{1,t+1}].\end{aligned}$$

Next, notice that,

$$\begin{aligned}\mathrm{Cov}\left(\varepsilon_{1,t+1}, \Delta x^*_{1,t+1}\right) &= \frac{b}{a+b}[\sigma_{12}(\tau)(1-\gamma_1\gamma_2(\tau)) - \sigma_1^2(1-\gamma_1^2)], \\ \mathrm{Cov}\left(\varepsilon_{2,t+1}, \Delta x^*_{1,t+1}\right) &= \frac{b}{a+b}[\sigma_2^2(\tau)(1-\gamma_2(\tau)^2) - \sigma_{12}(1-\gamma_1\gamma_2(\tau))].\end{aligned}$$

Since,

$$\mathbb{E}\left[\varepsilon_{i,t+1}\middle|\Delta x^*_{1,t+1} = \Delta x\right] = \frac{\mathrm{Cov}\left(\varepsilon_{i,t+1}, \Delta x^*_{1,t+1}\right)}{\mathrm{Var}\left(\Delta x^*_{1,t+1}\right)}\Delta x \qquad i = 1, 2,$$

it can be verified that,

$$\lim_{\tau\to 1, \tau\in[0,1)} \mathbb{E}\left[\varepsilon_{1,t+1}\middle|\Delta x^*_{1,t+1} = \Delta x\right] = \lim_{\tau\to 1} \mathbb{E}\left[\varepsilon_{2,t+1}\middle|\Delta x^*_{1,t+1} = \Delta x\right].$$

As a result,

$$\lim_{\tau\to 1, \tau\in[0,1)} \mathbb{E}\left[\Delta x^*_{1,t+1} - \mathbb{E}_t\left[\Delta x^*_{1,t+1}\right]\right] = \lim_{\tau\to 1, \tau\in[0,1)} \mathbb{E}\left[\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R5}_{1,t+1}\right] = 0.$$

Now, we can prove that,

$$\begin{aligned}\lim_{\substack{\kappa,\tau\to 1 \\ \tau,\kappa\in[0,1)}} \mathbb{P}\left[\mathbb{E}_t\left[L^{R3}_{t+1}\right] < \mathbb{E}_t\left[L^{R1}_{t+1}\right]\right] = 1, \\ \lim_{\substack{\kappa,\tau\to 1 \\ \tau,\kappa\in[0,1)}} \mathbb{P}\left[\mathbb{E}_t\left[L^{R3}_{t+1}\right] < \mathbb{E}_t\left[L^{R2}_{t+1}\right]\right] = 1.\end{aligned}$$

Let c_1 be the coefficient from a linear projection of $\Delta\bar{x}_{1,t}$ on a constant and $\Delta\bar{y}_{1,t}$. Let (τ_n, κ_n)

be an arbitrary sequence satisfying $\tau_n, \kappa_n \overset{n\to\infty}{\to} 1$ and $\kappa_n, \tau_n \in [0,1)$ for all $n \in \mathbb{N}$. Using Lemmas 5 and 8, we have for each t:

$$\lim_{n\to\infty} \Delta \mathbb{F} y_{t+1} = (\tilde{\theta}_1 + \tilde{\theta}_2)\Delta p_{1,t} + e_t - e_{t-1}. \tag{90}$$

Next, notice that $\mathbb{E}_t[\Delta p_{i,t+1}] = \bar{p}_i - (1-\gamma_i)\log p_{i,t}$. Thus, using Lemma 5,

$$\lim_{n\to\infty} \mathbb{E}_t[\Delta p_{2,t+1} - \Delta p_{1,t+1}] = 0. \tag{91}$$

Using (90) and (91), the assumption that $x^*_{i,t} = \bar{x}_{i,t}$ for all $t \le T$ and $i = 1, 2$, Lemma 6, and (80), (81), and (83), we obtain:

$$\begin{aligned}
&\lim_{n\to\infty}(\mathbb{E}_t\left[L^{R3}_{t+1}\right] - \mathbb{E}_T\left[L^{O}_{t+1}\right]) = 0 \\
&\lim_{n\to\infty}(\mathbb{E}_t\left[L^{R2}_{t+1}\right] - \mathbb{E}_T\left[L^{O}_{t+1}\right]) = \left[\left(1 - \frac{1}{a+b}\right)[(\tilde{\theta}_1+\tilde{\theta}_2)\Delta p_{1,t} + e_t - e_{t-1}]\right]^2 \\
&\lim_{n\to\infty}(\mathbb{E}_t\left[L^{R1}_{t+1}\right] - \mathbb{E}_t\left[L^{O}_{t+1}\right]) = \left[\frac{(\tilde{\theta}_1+\tilde{\theta}_2)}{a+b}\left[\Delta p_{1,t} - \frac{1}{k}\sum_{j=1}^{k}\Delta p_{1,t-j}\right]\right. \\
&\qquad\qquad\left. + \frac{1}{a+b}\left[e_t - e_{t-1} - \frac{1}{k}\sum_{j=1}^{k}(e_{t-j} - e_{t-j-1})\right]\right]^2.
\end{aligned}$$

Since $\lim_{n\to\infty}(\mathbb{E}_t\left[L^{R2}_{t+1}\right] - \mathbb{E}_t\left[L^{O}_{t+1}\right])$ and $\lim_{n\to\infty}(\mathbb{E}_t\left[L^{R1}_{t+1}\right] - \mathbb{E}_t\left[L^{O}_{t+1}\right])$ are strictly positive almost surely (equal to 0 happens with probability 0 since the limits are Gaussian),

$$\begin{aligned}
&\liminf_{n\to\infty} 1(\mathbb{E}_t\left[L^{R3}_{t+1}\right] < \mathbb{E}_t\left[L^{R1}_{t+1}\right]) = 1, \\
&\liminf_{n\to\infty} 1(\mathbb{E}_t\left[L^{R3}_{t+1}\right] < \mathbb{E}_t\left[L^{R2}_{t+1}\right]) = 1.
\end{aligned}$$

Next, by Fatou's lemma,

$$\begin{aligned}
&\liminf_{n\to\infty} \mathbb{P}\left[\mathbb{E}_t\left[L^{R3}_{t+1}\right] < \mathbb{E}_t\left[L^{R1}_{t+1}\right]\right] \ge \mathbb{E}\left[\liminf_{n\to\infty} 1(\mathbb{E}_t\left[L^{R3}_{t+1}\right] < \mathbb{E}_t\left[L^{R1}_{t+1}\right])\right] = 1, \\
&\liminf_{n\to\infty} \mathbb{P}\left[\mathbb{E}_t\left[L^{R3}_{t+1}\right] < \mathbb{E}_t\left[L^{R2}_{t+1}\right]\right] \ge \mathbb{E}\left[\liminf_{n\to\infty} 1(\mathbb{E}_t\left[L^{R3}_{t+1}\right] < \mathbb{E}_t\left[L^{R2}_{t+1}\right])\right] = 1.
\end{aligned}$$

Thus,

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{P}\left[\mathbb{E}_t\left[L^{R3}_{t+1}\right] < \mathbb{E}_t\left[L^{R1}_{t+1}\right]\right] = 1, \\
&\lim_{n\to\infty} \mathbb{P}\left[\mathbb{E}_t\left[L^{R3}_{t+1}\right] < \mathbb{E}_t\left[L^{R2}_{t+1}\right]\right] = 1.
\end{aligned}$$

Since (τ_n, κ_n) are arbitrary sequences satisfying $\tau_n, \kappa_n \overset{n\to\infty}{\to} 1$ and $\kappa_n, \tau_n \in [0,1)$ for all $n \in \mathbb{N}$, it follows that:

$$\begin{aligned}
&\lim_{\substack{\kappa,\tau\to 1 \\ \tau,\kappa\in[0,1)}} \mathbb{P}\left[\mathbb{E}_t\left[L^{R3}_{t+1}\right] < \mathbb{E}_t\left[L^{R1}_{t+1}\right]\right] = 1, \\
&\lim_{\substack{\kappa,\tau\to 1 \\ \tau,\kappa\in[0,1)}} \mathbb{P}\left[\mathbb{E}_t\left[L^{R3}_{t+1}\right] < \mathbb{E}_t\left[L^{R2}_{t+1}\right]\right] = 1.
\end{aligned}$$

This establishes the result. Now we establish that:

$$\lim_{\substack{\tau\to 1, a+b\to 1\\ \tau,\in[0,1)a,b\in(0,1),a+b<1}} \mathbb{P}\left[\mathbb{E}_t\left[L^{R2}_{t+1}\right] < \mathbb{E}_t\left[L^{R1}_{t+1}\right]\right] = 1.$$

Let $(\tau_n),(a_n),(b_n)$ be arbitrary sequences satisfying $\tau_n \overset{n\to\infty}{\to} 1, a_n+b_n \to 1$ and $\tau_n \in [0,1), a_n \in (0,1), b_n \in (0,1), a_n+b_n \in (0,1)$ for all $n \in \mathbb{N}$. Using Lemmas 5 and 8, we have for each t:

$$\lim_{n\to\infty} \Delta\mathbb{F}_{y,t+1} = (\tilde{\theta}_1+\tilde{\theta}_2)\Delta p_{1,t} + e_t - e_{t-1}. \tag{92}$$

Next, notice that $\mathbb{E}_t\left[\Delta p_{i,t+1}\right] = \bar{p}_i - (1-\gamma_i)\log p_{i,t}$. Thus, using Lemma 5,

$$\lim_{n\to\infty} \mathbb{E}_t\left[\Delta p_{2,t+1} - \Delta p_{1,t+1}\right] = 0. \tag{93}$$

Using (92) and (93), the assumption that $x^*_{i,t} = \bar{x}_{i,t}$ for all $t \le T$ and $i = 1,2$, Lemma 6, and (81) and (83), we obtain:

$$\begin{aligned}
&\lim_{n\to\infty}(\mathbb{E}_t\left[L^{R2}_{t+1}\right] - \mathbb{E}_t\left[L^{O}_{t+1}\right]) = 0,\\
&\lim_{n\to\infty}(\mathbb{E}_t\left[L^{R1}_{t+1}\right] - \mathbb{E}_t\left[L^{O}_{t+1}\right]) = \Bigg[(\tilde{\theta}_1+\tilde{\theta}_2)\left[\Delta p_{1,t} - \frac{1}{k}\sum_{j=1}^{k}\Delta p_{1,t-j}\right]\\
&\qquad\qquad + e_t - e_{t-1} - \frac{1}{k}\sum_{j=1}^{k}(e_{t-j} - e_{t-j-1})\Bigg]^2.
\end{aligned}$$

Since $\lim_{n\to\infty}(\mathbb{E}_t\left[L^{R2}_{t+1}\right] - \mathbb{E}_t\left[L^{O}_{t+1}\right])$ is strictly positive almost surely (equal to 0 happens with probability 0 since the limits are Gaussian),

$$\liminf_{n\to\infty} 1(\mathbb{E}_t\left[L^{R2}_{t+1}\right] < \mathbb{E}_t\left[L^{R1}_{t+1}\right]) = 1.$$

Next, by Fatou's lemma,

$$\liminf_{n\to\infty} \mathbb{P}\left[\mathbb{E}_t\left[L^{R2}_{t+1}\right] < \mathbb{E}_t\left[L^{R1}_{t+1}\right]\right] \ge \mathbb{E}\left[\liminf_{n\to\infty} 1(\mathbb{E}_t\left[L^{R2}_{t+1}\right] < \mathbb{E}_t\left[L^{R1}_{t+1}\right])\right] = 1.$$

Thus,

$$\lim_{n\to\infty} \mathbb{P}\left[\mathbb{E}_t\left[L^{R2}_{t+1}\right] < \mathbb{E}_t\left[L^{R1}_{t+1}\right]\right] = 1.$$

Since $(\tau_n),(a_n),(b_n)$ are arbitrary sequences satisfying $\tau_n \overset{n\to\infty}{\to} 1, a_n+b_n \to 1$ and $\tau_n \in [0,1), a_n \in (0,1), b_n \in (0,1), a_n+b_n \in (0,1)$ for all $n \in \mathbb{N}$, it follows that:

$$\lim_{\substack{\tau\to 1, a+b\to 1\\ \tau,\in[0,1)a,b\in(0,1),a+b<1}} \mathbb{P}\left[\mathbb{E}_t\left[L^{R2}_{t+1}\right] < \mathbb{E}_t\left[L^{R1}_{t+1}\right]\right] = 1.$$

This completes the proof.
QED

We now examine under which conditions, if any, RoTs (R1), (R2), and (R3) may yield unbiased estimates and thus approach the optimal forecast. We provide and prove three ancillary

corollaries.
Corollary 1.1 (Rationalizing R3). *Under (R3), it holds that,*

$$\lim_{\tau\to 1,\tau\in[0,1)} \mathbb{E}[\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R3}_{1,t+1}|\Delta x^*_{1,t+1} = \Delta x] = [1 - d_1(a+b)]\,\Delta x,$$
$$\lim_{\substack{\kappa\to 1,\tau\to 1\\ \tau,\kappa\in[0,1)}} \mathbb{E}[\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R3}_{1,t+1}|\Delta x^*_{1,t+1} = \Delta x] = 0.$$

Corollary 1.1 provides conditions under which (R3) produces unbiased forecasts. The bias of (R3) tends to 0 if κ and τ are large enough. Intuitively, if input prices are highly correlated, there is little benefit from multivariate (R5) relative to univariate (R3).
Proof of Corollary 1.1. By Lemma 6, $\Delta\mathbb{F}x^{R3}_{1,t+1} = c_1\Delta\mathbb{F}y_{t+1}$. Thus, by (47),

$$\begin{aligned}\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R3}_{1,t+1} &= \Delta x^*_{1,t+1} - c_1\Delta\mathbb{F}y_{t+1}\\ &= \left(\frac{1}{a+b} - c_1\right)\Delta\mathbb{F}y_{t+1} + \frac{b}{a+b}(\Delta p_{2,t+1} - \Delta p_{1,t+1})\\ &= [1 - c_1(a+b)]\left[\frac{1}{a+b}\Delta\mathbb{F}y_{t+1} + \frac{b}{a+b}(\Delta p_{2,t+1} - \Delta p_{1,t+1})\right]\\ &+ bc_1(\Delta p_{2,t+1} - \Delta p_{1,t+1}).\end{aligned}$$

Thus, by (42),

$$\mathbb{E}\left[\Delta x^*_{1,t+1} - c_1\Delta\mathbb{F}y_{t+1}|\Delta x^*_{1,t+1}\right] = (1 - c_1(a+b))\Delta x^*_{1,t+1} + bc_1\mathbb{E}\left[\Delta p_{2,t+1} - \Delta p_{1,t+1}|\Delta x^*_{1,t+1}\right].$$

By (61),

$$\lim_{\tau\to 1,\tau\in[0,1)} \mathbb{E}\left[\Delta p_{2,t+1} - \Delta p_{1,t+1}|\Delta x^*_{1,t+1} = \Delta x\right] = 0.$$

Thus,

$$\lim_{\tau\to 1,\tau\in[0,1)} \mathbb{E}\left[\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R3}_{1,t+1}|\Delta x^*_{1,t+1} = \Delta x\right] = (1 - c_1(a+b))\Delta x.$$

Next, by Lemma 6, $\mathrm{Corr}(\Delta p^F_{1,t} - \Delta p^F_{2,t}, \Delta\bar{y}^F_t) > 0$ implies that $c_1 < 1/(a+b)$ and thus $1 - c_1(a+b) > 0$. Lastly, Lemma 6 shows that $\lim_{\kappa\to 1,\kappa\in[0,1)} c_1(\kappa) = 1/(a+b)$. It is then straightforward to show that:

$$\lim_{\substack{\tau\to 1,\tau\in[0,1)\\ \kappa\to 1,\kappa\in[0,1)}} \mathbb{E}\left[\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R3}_{1,t+1}|\Delta x^*_{1,t+1} = \Delta x\right] = 0.$$

This establishes the result.
QED

Corollary 1.2 (Rationalizing R2). *Under (R2), it holds that,*

$$\lim_{\tau\to 1,\tau\in[0,1)} \mathbb{E}[\Delta x^*_{1,t+1} - \Delta\mathbb{F}x^{R2}_{1,t+1}|\Delta x^*_{1,t+1} = \Delta x] = (1 - a - b)\Delta x. \tag{94}$$

Corollary 1.2 shows (R2) yields incoherent forecasts, which differ from those implied by (R3). The bias of (R2) tends to 0 as long as the co-movement in input prices is sufficiently large and the technology approaches constant returns to scale, $a + b \to 1$. Intuitively, under decreasing returns, it is necessary to forecast a larger increase in input quantities than the sales growth target, but (R2) dictates to forecast the same growth in input as in sales. As a result, (R2) systematically underpredicts input growth.

Proof of Corollary 1.2. By equation (47),

$$\begin{aligned}\Delta x^*_{1,t+1} - c\Delta \mathbb{F} y_{t+1} &= \left(\frac{1}{a+b} - c\right)\Delta \mathbb{F} y_{t+1} + \frac{b}{a+b}(\Delta p_{2,t+1} - \Delta p_{1,t+1}) \\ &= [1 - c(a+b)]\left[\frac{1}{a+b}\Delta \mathbb{F} y_{t+1} + \frac{b}{a+b}(\Delta p_{2,t+1} - \Delta p_{1,t+1})\right] \\ &+ bc(\Delta p_{2,t+1} - \Delta p_{1,t+1}).\end{aligned}$$

Thus, by (42),

$$\mathbb{E}\left[\Delta x^*_{1,t+1} - c\Delta \mathbb{F} y_{t+1} | \Delta x^*_{1,t+1}\right] = (1 - c(a+b))\Delta x^*_{1,t+1} + bc\mathbb{E}\left[\Delta p_{2,t+1} - \Delta p_{1,t+1} | \Delta x^*_{1,t+1}\right].$$

By (61),

$$\lim_{\tau \to 1, \tau \in [0,1)} \mathbb{E}\left[\Delta p_{2,t+1} - \Delta p_{1,t+1} | \Delta x^*_{1,t+1} = \Delta x\right] = 0.$$

Thus,

$$\lim_{\tau \to 1, \tau \in [0,1)} \mathbb{E}\left[\Delta x^*_{1,t+1} - c\Delta \mathbb{F} y_{t+1} | \Delta x^*_{1,t+1} = \Delta x\right] = (1 - c(a+b))\Delta x.$$

Choosing $c = 1$ establishes the result since $\Delta \mathbb{F} x^{R2}_{1,t+1} = \Delta \mathbb{F} y_{t+1}$.
QED

Corollary 1.3 (Rationalizing R1). *Under (R1), it holds that,*

$$\begin{aligned}\lim_{\tau \to 1, \tau \in [0,1)} &\mathbb{E}[\Delta x^*_{1,t+1} - \Delta \mathbb{F}^{R1}_{1,t+1} | \Delta x^*_{1,t+1} = \Delta x] \\ &= \Delta x\left[1 + \frac{1}{k}\frac{\sigma_1^2(1-\gamma_1^k)(1-\gamma_1)\lambda + \sigma^2(1-\iota^k)(1-\iota)}{2\sigma_1^2(1-\gamma_1)\lambda + 2\sigma^2(1-\iota)}\right],\end{aligned} \tag{95}$$

where λ is a non-negative constant.

Corollary 1.3 shows the sign of (R1)'s bias depends on the sign of the cost-minimizing target, Δx. If past input prices follow an upward trend (so that input quantities follow a downward trend), then by mean reversion one would expect a reduction in input prices between t and $t+1$, and thus a growth in the cost-minimizing quantities, $\Delta \mathbb{F} x_{1,t+1} > 0$. In this case, however, (R1) would forecast a decrease in the quantity of input 1, thereby underpredicting its growth rate. This bias decreases as the forecaster uses more data, $k \to +\infty$, but it never goes to 0 as long as past input growth is non-zero. We prove the bias of (R1) is larger than the bias of each of the other rules.

Proof of Corollary 1.3. First, notice that by (48),

$$\Delta x^*_{1,t+1} - \Delta \mathbb{F} x^{R1}_{1,t+1} = \frac{1}{a+b}\left[\Delta \mathbb{F} y_{t+1} - \frac{1}{k}\sum_{j=1}^{k} \Delta \mathbb{F} y_{t+1-j}\right] + \frac{b}{a+b}\left[\Delta p_{2,t+1} - \Delta p_{1,t+1} - \frac{1}{k}\sum_{j=1}^{k}(\Delta p_{2,t+1-j} - \Delta p_{1,t+1-j})\right].$$

Next, note that by (65) and (66),

$$\begin{aligned}
&\lim_{\tau\to 1, \tau\in[0,1)} \mathbb{E}\left[\frac{1}{a+b}\left[\Delta \mathbb{F} y_{t+1} - \frac{1}{k}\sum_{j=1}^{k} \Delta \mathbb{F} y_{t+1-j}\right]\middle| \Delta x^*_{1,t+1} = \Delta x\right] \\
&= \Delta x\left[1 + \frac{1}{k}\frac{(\tilde\theta_1+\tilde\theta_2)^2\sigma_1^2(1-\gamma_1)^2\sum_{j=1}^{k}\gamma_1^{j-1} + \sigma^2(1-\iota)^2\sum_{j=1}^{k}\iota^{j-1}}{2\sigma_1^2(1-\gamma_1)(\tilde\theta_1+\tilde\theta_2)^2 + 2(1-\iota)\sigma^2}\right] \\
&= \Delta x\left[1 + \frac{1}{k}\frac{(\tilde\theta_1+\tilde\theta_2)^2\sigma_1^2(1-\gamma_1)(1-\gamma_1^k) + \sigma^2(1-\iota)(1-\iota^k)}{2\sigma_1^2(1-\gamma_1)(\tilde\theta_1+\tilde\theta_2)^2 + 2(1-\iota)\sigma^2}\right].
\end{aligned}$$

By (62), (63), and (61),

$$\lim_{\tau\to 1, \tau\in[0,1)} \mathbb{E}[\Delta x^*_{1,t+1} - \Delta \mathbb{F} x^{R1}_{1,t+1} | \Delta x^*_{1,t+1} = \Delta x] \tag{96}$$

$$= \Delta x\left[1 + \frac{1}{k}\frac{(\tilde\theta_1+\tilde\theta_2)^2\sigma_1^2(1-\gamma_1)(1-\gamma_1^k) + \sigma^2(1-\iota)(1-\iota^k)}{2\sigma_1^2(1-\gamma_1)(\tilde\theta_1+\tilde\theta_2)^2 + 2(1-\iota)\sigma^2}\right]. \tag{97}$$

This establishes the result with $\lambda = (\tilde\theta_1 + \tilde\theta_2)^2$. Next, suppose $\gamma_i \geq 0$ for $i = 1, 2$. Define the mapping $f(x) = (1-\gamma_i^x)/x$ for $x \geq 1$. Since $\gamma_i \in [0, 1)$, it follows that,

$$\frac{\partial f(x)}{\partial x} = \frac{-\gamma_1^{x-1}(x^2 - \gamma_1) - 1}{x^2} < 0 \quad \forall x \geq 1.$$

This implies that the RHS of (96) is strictly decreasing in k.
QED

Proof of Implication 1. We first prove that $c_0 = c_1 = 0$ under rule (R5). First, note that since $x^*_{1,s} = \bar x_{1,s}$ for all $s \leq t$ by assumption,

$$\Delta \bar x_{1,t} = \log \bar x_{1,t} - \log \bar x_{1,t-1} = \log x^*_{1,t} - \log x^*_{1,t-1} = \Delta x^*_{1,t}.$$

From (46),

$$\begin{aligned}
\mathbb{E}\left[\Delta x^*_{1,t+1} - \mathbb{E}_t\left[\Delta x^*_{1,t+1}\right]\middle|\Delta \bar x_{1,t}\right] &= \frac{b}{a+b}(\mathbb{E}\left[\epsilon_{1,t+1}|\Delta \bar x_{1,t}\right] - \mathbb{E}\left[\epsilon_{2,t+1}|\Delta \bar x_{1,t}\right]) \\
&= \frac{b}{a+b}(\mathbb{E}\left[\epsilon_{1,t+1}|\Delta x^*_{1,t}\right] - \mathbb{E}\left[\epsilon_{2,t+1}|\Delta x^*_{1,t}\right]) \\
&= \frac{b}{a+b}\frac{\operatorname{Cov}\left(\epsilon_{2,t+1} - \epsilon_{1,t+1}, \Delta x^*_{1,t}\right)}{\operatorname{Var}\left(\Delta x^*_{1,t}\right)}\Delta \bar x_{1,t}.
\end{aligned}$$

Let c_0, c_1 be the coefficients from a linear projection of $\Delta x^*_{1,t+1} - \Delta \mathbb{F} x^{R5}_{1,t+1}$ on a constant and the previous period implemented growth in input 1 $\Delta \bar{x}_{1,t}$. By Lemma 4, $c_0 = 0$ and

$$\begin{aligned}
c_1 &= \frac{b}{a+b} \frac{\operatorname{Cov}\left(\epsilon_{2,t+1} - \epsilon_{1,t+1}, \Delta x^*_{1,t}\right)}{\operatorname{Var}\left(\Delta x^*_{1,t}\right)} \\
&= \frac{b}{a+b} \frac{\frac{1}{a+b} \operatorname{Cov}\left(\epsilon_{2,t+1} - \epsilon_{1,t+1}, \Delta \mathbb{F} y_t\right) + \frac{b}{a+b} \operatorname{Cov}\left(\epsilon_{2,t+1} - \epsilon_{1,t+1}, \Delta p_{2,t} - \Delta p_{1,t}\right)}{\operatorname{Var}\left(\Delta x^*_{1,t}\right)} \\
&= 0,
\end{aligned}$$

where the last equality follows because:

$$\operatorname{Cov}\left(\epsilon_{2,t+1} - \epsilon_{1,t+1}, \Delta \mathbb{F} y_t\right) = \operatorname{Cov}\left(\epsilon_{2,t+1} - \epsilon_{1,t+1}, \theta_1(\tau)\Delta p_{1,t} + \theta_2(\tau)\Delta p_{2,t} + e_t - e_{t-1}\right) = 0.$$

Thus, $c_0 = c_1 = 0$. Moreover, $\mathbb{E}\left[\Delta x^*_{1,t+1} - \mathbb{E}_t\left[\Delta x^*_{1,t+1}\right] \middle| \Delta \bar{x}_{1,t}\right] = 0$. This proves the result for rule (R5).

Now we derive the linear projection coefficients under rule (R2). First, note that since $x^*_{1,s} = \bar{x}_{1,s}$ for all $s \leq t$ by assumption,

$$\Delta \bar{x}_{1,t} = \log \bar{x}_{1,t} - \log \bar{x}_{1,t-1} = \log x_{1,t} - \log x^*_{1,t-1} = \Delta x^*_{1,t}.$$

From (47),

$$\begin{aligned}
&\mathbb{E}\left[\Delta x^*_{1,t+1} - \Delta \mathbb{F} x^{R2}_{1,t+1} \middle| \Delta \bar{x}_{1,t}\right] \\
&\quad = \left(\frac{1}{a+b} - 1\right) \mathbb{E}\left[\Delta \mathbb{F} y_{t+1} \middle| \Delta x^*_{1,t}\right] + \frac{b}{a+b} \mathbb{E}\left[\Delta p_{2,t+1} - \Delta p_{1,t+1} \middle| \Delta x^*_{1,t}\right] \\
&\quad = \frac{\left[\left(\frac{1}{a+b} - 1\right) \operatorname{Cov}\left(\Delta \mathbb{F} y_{t+1}, \Delta x^*_{1,t}\right) + \frac{b}{a+b} \operatorname{Cov}\left(\Delta p_{2,t+1} - \Delta p_{1,t+1}, \Delta x^*_{1,t}\right)\right]}{\operatorname{Var}\left(\Delta x^*_{1,t}\right)} \Delta \bar{x}_{1,t}.
\end{aligned}$$

Let c_0, c_1 be the coefficients from a linear projection of $\Delta x^*_{1,t+1} - \Delta \mathbb{F} x^{R5}_{1,t+1}$ on a constant and the previous period implemented growth in input 1 $\Delta \bar{x}_{1,t}$. By Lemma 4, $c_0 = 0$ and:

$$c_1 = \frac{\left[\left(\frac{1}{a+b} - 1\right) \operatorname{Cov}\left(\Delta \mathbb{F} y_{t+1}, \Delta x^*_{1,t}\right) + \frac{b}{a+b} \operatorname{Cov}\left(\Delta p_{2,t+1} - \Delta p_{1,t+1}, \Delta x^*_{1,t}\right)\right]}{\operatorname{Var}\left(\Delta x^*_{1,t}\right)}.$$

Note that by (64) and (60),

$$\begin{aligned}
&\lim_{\tau \to 1, \tau \in [0,1)} \mathbb{E}\left[\Delta x^*_{1,t+1} - \Delta \mathbb{F} x^{R2}_{1,t+1} \middle| \Delta \bar{x}_{1,t} = \Delta x\right] \\
&\qquad = -\frac{(1-a-b)[(\tilde{\theta}_1 + \tilde{\theta}_2)^2 \sigma_1^2 (1-\gamma_1)^2 + \sigma^2(1-\iota)^2]}{2\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1 + \tilde{\theta}_2)^2 + 2\sigma^2(1-\iota)} \Delta x.
\end{aligned}$$

Thus, since $\mathbb{E}\left[\Delta x^*_{1,t+1} - \Delta \mathbb{F} x^{R2}_{1,t+1} \middle| \Delta \bar{x}_{1,t} = \Delta x\right] = c_1 \Delta x$ almost surely,

$$c_1 = \frac{\mathbb{E}\left[\Delta x^*_{1,t+1} - \Delta \mathbb{F} x^{R2}_{1,t+1} \middle| \Delta \bar{x}_{1,t} = \Delta x\right]}{\Delta x} \overset{\substack{\tau\to 1\\ \tau\in[0,1)}}{\rightarrow} -\frac{(1-a-b)[(\tilde{\theta}_1+\tilde{\theta}_2)^2\sigma_1^2(1-\gamma_1)^2+\sigma^2(1-\iota)^2]}{2\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1+\tilde{\theta}_2)^2+2\sigma^2(1-\iota)},$$

which holds since $\mathbb{P}\left[\Delta \bar{x}_{1,t} \neq 0\right] = \mathbb{P}\left[\Delta x^*_{1,t} \neq 0\right] = 1$ and the intersection of two almost-sure events is non-empty. This establishes the result for rule (R2).

Now we prove the result for rule (R1). First, note that since $x^*_{1,s} = \bar{x}_{1,s}$ for all $s \leq t$ by assumption,

$$\Delta \bar{x}_{1,t} = \log \bar{x}_{1,t} - \log \bar{x}_{1,t-1} = \log x^*_{1,t} - \log x^*_{1,t-1} = \Delta x^*_{1,t}.$$

From (48),

$$\mathbb{E}\left[\Delta x^*_{1,t+1} - \Delta \mathbb{F} x^{R1}_{1,t+1} \middle| \Delta \bar{x}_{1,t}\right] = \frac{1}{a+b}\left[\mathbb{E}\left[\Delta \mathbb{F} y_{t+1} \middle| \Delta x^*_{1,t}\right] - \frac{1}{k}\sum_{j=1}^{k} \mathbb{E}\left[\Delta \mathbb{F}_{y,t+1-j} \middle| \Delta x^*_{1,t}\right]\right] \tag{98}$$

$$+ \frac{b}{a+b}\left[\mathbb{E}\left[\Delta p_{2,t+1} - \Delta p_{1,t+1} \middle| \Delta x^*_{1,t}\right] - \frac{1}{k}\sum_{j=1}^{k} \mathbb{E}\left[\Delta p_{2,t+1} - \Delta p_{1,t+1} \middle| \Delta x^*_{1,t}\right]\right]. \tag{99}$$

From (60), (62), (63), (64), (65), and (66),

$$\lim_{\tau\to 1,\tau\in[0,1)} \mathbb{E}\left[\Delta x^*_{1,t+1} - \Delta \mathbb{F} x^{R1}_{1,t+1} \middle| \Delta \bar{x}_{1,t} = \Delta x\right]$$
$$= -\Delta x \left[\frac{\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1+\tilde{\theta}_2)^2\left[(1-\gamma_1)+\dfrac{1+\gamma_1^{k-1}}{k}\right]+\sigma^2(1-\iota)\left[2-\iota+\iota^{k-1}\right]}{2\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1+\tilde{\theta}_2)^2+2\sigma^2(1-\iota)}\right].$$

Let (c_0, c_1) be the coefficients from a linear projection of $\Delta x^*_{1,t+1} - \Delta \mathbb{F} x^{R1}_{1,t+1}$ on a constant and the previous period implemented growth in input 1, $\Delta \bar{x}_{1,t}$. Examining (98), we have $c_0 = 0$ and $\mathbb{E}\left[\Delta x^*_{1,t+1} - \Delta \mathbb{F} x^{R1}_{1,t+1} \middle| \Delta \bar{x}_{1,t} = \Delta x\right] = c_1 \Delta x$ almost surely. Thus, since $\mathbb{P}\left[\Delta \bar{x}_{1,t} \neq 0\right] = \mathbb{P}\left[\Delta x^*_{1,t} \neq 0\right] = 1$, and the intersection of two almost-sure events is non-empty, it follows that there exists $\Delta x \neq 0$ such that:

$$c_1 = \frac{\mathbb{E}\left[\Delta x^*_{1,t+1} - \Delta \mathbb{F} x^{R1}_{1,t+1} \middle| \Delta \bar{x}_{1,t} = \Delta x\right]}{\Delta x}$$
$$\overset{\substack{\tau\to 1\\ \tau\in[0,1)}}{\rightarrow} -\frac{\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1+\tilde{\theta}_2)^2\left[(1-\gamma_1)+\dfrac{1+\gamma_1^{k-1}}{k}\right]+\sigma^2(1-\iota)\left[2-\iota+\iota^{k-1}\right]}{2\sigma_1^2(1-\gamma_1)(\tilde{\theta}_1+\tilde{\theta}_2)^2+2\sigma^2(1-\iota)}.$$

This establishes the result for rule (R1).
QED

A.5 Underinvestment with Costly Ex Post Adjustment Costs: Structural Derivation of Implication 2

This subsection provides a structural foundation for the reduced-form underinvestment mechanism stated in subsection 3.2 of the main text. There, realized input growth is written as forecasted growth plus a partial correction toward the cost-minimizing target:

$$\Delta x_{1,t+1} = \Delta \mathbb{F}^{Rk}_{x1,t+1} + \lambda\Big(\Delta x^*_{1,t+1} - \Delta \mathbb{F}^{Rk}_{x1,t+1}\Big), \qquad \lambda \in (0,1]. \tag{100}$$

Here we derive this reduced form from a structural problem in which, after issuing a forecast based on rule (Rk), the firm can revise its input plan ex post, but only at convex adjustment cost.

Let $\Delta \mathbb{F}^{Rk}_{x1,t+1}$ and $\Delta \mathbb{F}^{Rk}_{x2,t+1}$ denote the forecasted growth rates of the two inputs under rule (Rk), and let $\Delta x^*_{1,t+1}$ and $\Delta x^*_{2,t+1}$ denote the corresponding cost-minimizing growth rates. After forecasts are issued, the firm chooses ex post adjustments $\mathbb{A}_1$ and $\mathbb{A}_2$, so that realized input growth is

$$\Delta x_{1,t+1} = \Delta \mathbb{F}^{Rk}_{x1,t+1} + \mathbb{A}_1, \qquad \Delta x_{2,t+1} = \Delta \mathbb{F}^{Rk}_{x2,t+1} + \mathbb{A}_2. \tag{101}$$

Let $(\mathbb{A}^*_1, \mathbb{A}^*_2)$ denote the optimal adjustment.

Define the investment gap for input 1 under rule (Rk) as

$$G^{Rk}_{t+1} \equiv \Delta x^*_{1,t+1} - \Delta x_{1,t+1} = \Delta x^*_{1,t+1} - \big(\Delta \mathbb{F}^{Rk}_{x1,t+1} + \mathbb{A}^*_1\big). \tag{102}$$

Hence

$$G^{Rk}_{t+1} = \Big(\Delta x^*_{1,t+1} - \Delta \mathbb{F}^{Rk}_{x1,t+1}\Big) - \mathbb{A}^*_1. \tag{103}$$

Whenever rule (Rk) is downward biased, so that $\Delta x^*_{1,t+1} - \Delta \mathbb{F}^{Rk}_{x1,t+1} > 0$, define

$$\lambda_1 \equiv \frac{\mathbb{A}^*_1}{\Delta x^*_{1,t+1} - \Delta \mathbb{F}^{Rk}_{x1,t+1}}. \tag{104}$$

Under convex adjustment costs, the optimal correction is partial, so $0 < \lambda_1 < 1$ except in the knife-edge case of costless full adjustment. Substituting (104) into (103) yields

$$G^{Rk}_{t+1} = (1-\lambda_1)\Big(\Delta x^*_{1,t+1} - \Delta \mathbb{F}^{Rk}_{x1,t+1}\Big), \tag{105}$$

which is exactly the reduced-form expression used in subsection 3.2 of the main text, with $\lambda \equiv \lambda_1$.

Therefore, the main-text formulation can be interpreted as a compact representation of the structural ex post adjustment problem studied below: when a forecasting rule underpredicts the cost-minimizing growth rate of the relevant input, convex adjustment costs prevent full ex post correction, and realized input growth remains below the benchmark.

Consider a CFO at time t, facing a sales growth target, $\Delta \mathbb{F} y_{t+1}$, and making input expenditures growth forecasts, $\Delta \mathbb{F} x_{i,t+1}, i = 1,2$. After these forecasts are implemented, absent any further adjustment, the firm uses $\bar{x}_{1,t} \cdot e^{\Delta \mathbb{F} x_{i,t+1}}$ units of input i at time $t+1$. However, between times t and $t+1$, the CFO observes the realization of input prices and thus the cost minimizing growth, $\Delta x^*_{i,t+1}$, for both inputs $i = 1,2$.

As a result, the CFO can change the quantities of each input $i \in (1,2)$ by choosing a percent adjustment, $\mathbb{A}_i$, at a cost, $\mathbb{C}_i$. The final realized quantity of input i used by the firm at time $t+1$ is then $\bar{x}_{i,t} \cdot e^{(\Delta \mathbb{F} x_{i,t+1} + \mathbb{A}_i)}$. Total adjustment costs are,

$$\mathbb{C}(\mathbb{A}_1, \mathbb{A}_2) = C\left(\left\{ m_i \cdot \bar{x}_{i,t} \left[e^{\Delta \mathbb{F} x_{i,t+1} + \mathbb{A}_i} - e^{\Delta \mathbb{F} x_{i,t+1}} \right] \right\}_{i=1,2} \right),$$

where $\{m_i > 0\}_{i=1,2}$ are adjustment cost parameters for inputs $i = 1, 2$, and $C(z_1, z_2)$ is a convex cost function such that $C(0,0) = 0$ (zero cost in the absence of ex post adjustment), $C(z_1, z_2) > 0 \quad \forall (z_1, z_2) \neq (0,0)$ (strictly positive when adjustment is not zero), $C_{ii} = \frac{\partial^2 C}{\partial z_i^2} > 0$, and $C_{ij} = \frac{\partial^2 C(0,0)}{\partial z_i \partial z_j} = 0$, for any $i \neq j$. Because we assume $\mathbb{C}(\mathbb{A}_1, \mathbb{A}_2)$ is convex in both arguments, marginal adjustment costs are increasing in the magnitude of the ex post correction.

Therefore, given initial forecasts, $\Delta \mathbb{F} x_{1,t+1}$ and $\Delta \mathbb{F} x_{2,t+1}$, and the sales growth target, $\Delta \mathbb{F} y_{t+1}$, the CFO's problem amounts to choosing levels of ex-post adjustment that minimize total cost while achieving the sales growth target. That is,

$$\begin{aligned} \min_{\mathbb{A}_1, \mathbb{A}_2 \in \mathbb{R}} \quad & p_{1,t+1} \bar{x}_{1,t} e^{\Delta \mathbb{F} x_{1,t+1} + \mathbb{A}_1} + p_{2,t+1} \bar{x}_{2,t} e^{\Delta \mathbb{F} x_{2,t+1} + \mathbb{A}_2} + \mathbb{C}(\mathbb{A}_1, \mathbb{A}_2) \\ \text{s.t.} \quad & \left[\bar{x}_{1,t} e^{\Delta \mathbb{F} x_{1,t+1} + \mathbb{A}_1} \right]^a \cdot \left[\bar{x}_{2,t} e^{\Delta \mathbb{F} x_{2,t+1} + \mathbb{A}_2} \right]^b \geq \bar{y}_t e^{\Delta \mathbb{F} y_{t+1}}. \end{aligned} \tag{106}$$

We assume throughout that the following holds:

$$\frac{m_1^2 C_{11}}{(p_{1,t+1})^2} \cdot a[\Delta x^*_{1,t+1} - \Delta \mathbb{F} x_{1,t+1}] > \frac{m_2^2 C_{22}}{(p_{2,t+1})^2} \cdot b[\Delta x^*_{2,t+1} - \Delta \mathbb{F} x_{2,t+1}]. \tag{107}$$

Suppose the forecasts of both inputs are downward biased, $\Delta x^*_{i,t+1} - \Delta \mathbb{F} x_{i,t+1} > 0$, $i = 1, 2$. Then the realized growth rate of at least one input $i \in \{1, 2\}$ is strictly above the initial forecast and strictly below the cost-minimizing target,

$$\Delta \mathbb{F} x_{i,t+1} < \Delta \mathbb{F} x_{i,t+1} + \mathbb{A}^*_i < \Delta x^*_{i,t+1},$$

where $\mathbb{A}^*_1$ is the optimal ex-post adjustment for input 1,

$$\mathbb{A}^*_1 = R_1(m)[\Delta x^*_{1,t+1} - \Delta \mathbb{F} x_{1,t+1}] + R_2(m)[\Delta x^*_{2,t+1} - \Delta \mathbb{F} x_{2,t+1}], \tag{108}$$

and $R_1(m) > 0$ and $R_2(m) > 0$ are formally defined below and do not depend on initial forecasts, $\Delta \mathbb{F} x_{i,t+1}$, $i = 1, 2$.

Proof of Implication 2. We proceed in steps. First, we derive the optimal ex-post adjustment $\mathbb{A}^*_1$. After that, we will establish the under-investment result.

Optimal Ex-Post Adjustment. Convexity and twice continuous differentiability of $(x_1, x_2) \mapsto C(x_1, x_2)$ imply the Karush-Kuhn-Tucker (KKT) conditions are both necessary and sufficient to identify the solution of the ex-post cost minimization problem. Since $(\Delta x^*_{1,t+1}, \Delta x^*_{2,t+1})$ are the growth rates of inputs that achieve the sales revenue growth target, $\Delta \mathbb{F} y_{t+1}$, at minimum cost, the following KKT conditions hold:

$$\begin{aligned} \frac{a \bar{y}_t e^{\Delta \mathbb{F} y_{t+1}}}{\bar{x}_{1,t} e^{\Delta x^*_{1,t+1}} p_{1,t+1}} &= \frac{b \bar{y}_t e^{\Delta \mathbb{F} y_{t+1}}}{\bar{x}_{2,t} e^{\Delta x^*_{2,t+1}} p_{2,t+1}} \quad (\bar{x}_{1,t} e^{\Delta x^*_{1,t+1}})^a (\bar{x}_{2,t} e^{\Delta x^*_{2,t+1}})^b \\ &= \bar{y}_t e^{\Delta \mathbb{F} y_{t+1}}. \end{aligned} \tag{109}$$

A solution of ex-post cost minimization, $(\mathbb{A}^*_1, \mathbb{A}^*_2)$, satisfies the following KKT conditions:

$$
\begin{aligned}
&\frac{\bar{x}_{1,t}\exp(\Delta\mathbb{F}x_{1,t+1}+\mathbb{A}_1^*)}{a\bar{y}_t e^{\Delta\mathbb{F}y_{t+1}}} \\
&\cdot\left[p_{1,t+1}+m_1\frac{\partial C(m_1\cdot\bar{x}_{1,t}\exp(\Delta\mathbb{F}x_{1,t+1})[\exp(\mathbb{A}_1^*)-1],m_2\cdot\bar{x}_{2,t}\exp(\Delta\mathbb{F}x_{2,t+1})[\exp(\mathbb{A}_2^*)-1])}{\partial x_1}\right] \\
&=\frac{\bar{x}_{2,t}\exp(\Delta\mathbb{F}x_{2,t+1}+\mathbb{A}_2^*)}{b\bar{y}_t e^{\Delta\mathbb{F}y_{t+1}}} \\
&\cdot\left[p_{2,t+1}+m_2\frac{\partial C(m_1\cdot\bar{x}_{1,t}\exp(\Delta\mathbb{F}x_{1,t+1})[\exp(\mathbb{A}_1^*)-1],m_2\cdot\bar{x}_{2,t}\exp(\Delta\mathbb{F}x_{2,t+1})[\exp(\mathbb{A}_2^*)-1])}{\partial x_2}\right] \\
&(\bar{x}_{1,t}\exp(\Delta\mathbb{F}x_{1,t+1}+\mathbb{A}_1^*))^a(\bar{x}_{2,t}\exp(\Delta\mathbb{F}x_{2,t+1}+\mathbb{A}_2^*))^b=\bar{y}_t e^{\Delta\mathbb{F}y_{t+1}}.
\end{aligned}
$$

Define the mappings f_1, f_2, and g by:

$$
\begin{aligned}
&f_1(\delta_1,\delta_2,\xi_1,\xi_2)= \\
&\quad\frac{\bar{x}_{1,t}\exp(\delta_1+\xi_1)}{a}\left[p_{1,t+1}+m_1\frac{\partial C(m_1\cdot\bar{x}_{1,t}\exp(\delta_1)[\exp(\xi_1)-1],m_2\cdot\bar{x}_{2,t}\exp(\delta_2)[\exp(\xi_2)-1])}{\partial x_1}\right] \\
&f_2(\delta_1,\delta_2,\xi_1,\xi_2)= \\
&\quad\frac{\bar{x}_{2,t}\exp(\delta_2+\xi_2)}{b}\left[p_{2,t+1}+m_2\frac{\partial C(m_1\cdot\bar{x}_{1,t}\exp(\delta_1)[\exp(\xi_1)-1],m_2\cdot\bar{x}_{2,t}\exp(\delta_2)[\exp(\xi_2)-1])}{\partial x_2}\right] \\
&g(\delta_1,\delta_2,\xi_1,\xi_2)=(\bar{x}_{1,t}\exp(\delta_1+\xi_1))^a(\bar{x}_{2,t}\exp(\delta_2+\xi_2))^b.
\end{aligned}
$$

Since $C(0,0)=0$ and $C(x_1,x_2)>0$ for all $(x_1,x_2)\neq 0$, it holds that $\partial C(0,0)/\partial x_1 = \partial C(0,0)/\partial x_2 = 0$. This implies the following equations hold:

$$f_1(\Delta x^*_{1,t+1},\Delta x^*_{2,t+1},0,0)=f_2(\Delta x^*_{1,t+1},\Delta x^*_{2,t+1},0,0) \tag{110}$$
$$g(\Delta x^*_{1,t+1},\Delta x^*_{2,t+1},0,0)=\bar{y}_t e^{\Delta\mathbb{F}y_{t+1}} \tag{111}$$
$$f_1(\Delta\mathbb{F}x_{1,t+1},\Delta\mathbb{F}x_{2,t+1},\mathbb{A}_1^*,\mathbb{A}_2^*)=f_2(\Delta\mathbb{F}x_{1,t+1},\Delta\mathbb{F}x_{2,t+1},\mathbb{A}_1^*,\mathbb{A}_2^*) \tag{112}$$
$$g(\Delta\mathbb{F}x_{1,t+1},\Delta\mathbb{F}x_{2,t+1},\mathbb{A}_1^*,\mathbb{A}_2^*)=\bar{y}_t e^{\Delta\mathbb{F}y_{t+1}}. \tag{113}$$

Denote $\tilde{\Delta}=(\Delta\mathbb{F}x_{1,t+1},\Delta\mathbb{F}x_{2,t+1},\mathbb{A}_1^*,\mathbb{A}_2^*)$ and $\Delta=(\Delta x^*_{1,t+1},\Delta x^*_{2,t+1},0,0)$. A first order Taylor expansion of $\log f_i$ around Δ yields,

$$
\begin{aligned}
&\log f_i(\tilde{\Delta})=\log f_i(\Delta) \\
&\qquad+\frac{1}{f_i(\Delta)}\left[\frac{\partial f_i(\Delta)}{\partial\delta_1}[\Delta\mathbb{F}x_{1,t+1}-\Delta x^*_{1,t+1}]+\frac{\partial f_i(\Delta)}{\partial\delta_2}[\Delta\mathbb{F}x_{2,t+1}-\Delta x^*_{2,t+1}]+\frac{\partial f_i(\Delta)}{\partial\xi_1}\mathbb{A}_1^*+\frac{\partial f_i(\Delta)}{\partial\xi_2}\mathbb{A}_2^*\right].
\end{aligned}
$$

A first order Taylor expansion of $\log g$ around Δ yields,

$$
\begin{aligned}
&\log g(\tilde{\Delta})=\log g(\Delta) \\
&\qquad+\frac{1}{g(\Delta)}\left[\frac{\partial g(\Delta)}{\partial\delta_1}[\Delta\mathbb{F}x_{1,t+1}-\Delta x^*_{1,t+1}]+\frac{\partial g(\Delta)}{\partial\delta_2}[\Delta\mathbb{F}x_{2,t+1}-\Delta x^*_{2,t+1}]+\frac{\partial g(\Delta)}{\partial\xi_1}\mathbb{A}_1^*+\frac{\partial g(\Delta)}{\partial\xi_2}\mathbb{A}_2^*\right],
\end{aligned}
$$

where,

$$\frac{\partial f_1(\Delta)}{\partial \delta_1} = a^{-1} p_{1,t+1} \bar{x}_{1,t} \exp(\Delta x^*_{1,t+1}) \quad \frac{\partial f_1(\Delta)}{\partial \delta_2} = 0 \quad \frac{\partial f_1(\Delta)}{\partial \xi_2} = 0$$
$$\frac{\partial f_1(\Delta)}{\partial \xi_1} = a^{-1} p_{1,t+1} \bar{x}_{1,t} \exp(\Delta x^*_{1,t+1}) + \frac{m_1^2}{a} (\bar{x}_{1,t} \exp(\Delta x^*_{1,t+1}))^2 C_{11}$$
$$\frac{\partial f_2(\Delta)}{\partial \delta_1} = 0 \quad \frac{\partial f_2(\Delta)}{\partial \delta_2} = b^{-1} p_{2,t+1} \bar{x}_{2,t} \exp(\Delta x^*_{2,t+1}) \quad \frac{\partial f_1(\Delta)}{\partial \xi_1} = 0$$
$$\frac{\partial f_2(\Delta)}{\partial \xi_2} = b^{-1} p_{2,t+1} \bar{x}_{2,t} \exp(\Delta x^*_{2,t+1}) + \frac{m_2^2}{b} (\bar{x}_{2,t} \exp(\Delta x^*_{2,t+1}))^2 C_{22}$$
$$\frac{\partial g(\Delta)}{\partial \delta_1} = a \bar{y}_t e^{\Delta \mathbb{F} y_{t+1}} \quad \frac{\partial g(\Delta)}{\partial \delta_2} = b \bar{y}_t e^{\Delta \mathbb{F} y_{t+1}} \quad \frac{\partial g(\Delta)}{\partial \xi_1} = a \bar{y}_t e^{\Delta \mathbb{F} y_{t+1}} \quad \frac{\partial g(\Delta)}{\partial \xi_2} = b \bar{y}_t e^{\Delta \mathbb{F} y_{t+1}}.$$

Thus, using (110), the first order Taylor expansions, and the assumption that $C_{12} = 0$, we obtain,

$$\left[\frac{\partial f_1(\Delta)}{\partial \delta_1}[\Delta \mathbb{F} x_{1,t+1} - \Delta x^*_{1,t+1}] + \frac{\partial f_1(\Delta)}{\partial \delta_2}[\Delta \mathbb{F} x_{2,t+1} - \Delta x^*_{2,t+1}] + \frac{\partial f_1(\Delta)}{\partial \xi_1}\mathbb{A}^*_1 + \frac{\partial f_1(\Delta)}{\partial \xi_2}\mathbb{A}^*_2\right]$$
$$= \left[\frac{\partial f_2(\Delta)}{\partial \delta_1}[\Delta \mathbb{F} x_{1,t+1} - \Delta x^*_{1,t+1}] + \frac{\partial f_2(\Delta)}{\partial \delta_2}[\Delta \mathbb{F} x_{2,t+1} - \Delta x^*_{2,t+1}] + \frac{\partial f_2(\Delta)}{\partial \xi_1}\mathbb{A}^*_1 + \frac{\partial f_2(\Delta)}{\partial \xi_2}\mathbb{A}^*_2\right]$$

$$\frac{\partial g(\Delta)}{\partial \delta_1}[\Delta \mathbb{F} x_{1,t+1} - \Delta x^*_{1,t+1}] + \frac{\partial g(\Delta)}{\partial \delta_2}[\Delta \mathbb{F} x_{2,t+1} - \Delta x^*_{2,t+1}] + \frac{\partial g(\Delta)}{\partial \xi_1}\mathbb{A}^*_1 + \frac{\partial g(\Delta)}{\partial \xi_2}\mathbb{A}^*_2 = 0. \quad (114)$$

Solving for $(\mathbb{A}^*_1, \mathbb{A}^*_2)$, we obtain:

$$\mathbb{A}^*_1 = R_1(m)[\Delta x^*_{1,t+1} - \Delta \mathbb{F} x_{1,t+1}] + R_2(m)[\Delta x^*_{2,t+1} - \Delta \mathbb{F} x_{2,t+1}]$$

$$\mathbb{A}^*_2 = \frac{a}{b}(1 - R_1(m))[\Delta x^*_{1,t+1} - \Delta \mathbb{F} x_{1,t+1}] + (1 - \frac{a}{b})R_2(m)[\Delta x^*_{2,t+1} - \Delta \mathbb{F} x_{2,t+1}], \quad (115)$$

where:

$$R_1(m) = \frac{ak_1 + ak_2 + \frac{a^2}{b}k_4}{ak_1 + ak_2 + bk_3 + \frac{a^2}{b}k_4} \quad R_2(m) = \frac{ak_4}{ak_1 + ak_2 + bk_3 + \frac{a^2}{b}k_4}, \quad (116)$$

with:

$$k_1 = (\bar{x}_{1,t} \exp(\Delta x^*_{1,t+1}))^a (\bar{x}_{2,t} \exp(\Delta x^*_{2,t+1}))^{b-1} p_{1,t+1}$$
$$k_2 = (\bar{x}_{1,t} \exp(\Delta x^*_{1,t+1}))^{a-1} (\bar{x}_{2,t} \exp(\Delta x^*_{2,t+1}))^{b} p_{2,t+1}$$
$$k_3 = m_1^2 (\bar{x}_{1,t} \exp(\Delta x^*_{1,t+1}))^{a+1} (\bar{x}_{2,t} \exp(\Delta x^*_{2,t+1}))^{b-1} C_{11}$$
$$k_4 = m_2^2 (\bar{x}_{1,t} \exp(\Delta x^*_{1,t+1}))^{a-1} (\bar{x}_{2,t} \exp(\Delta x^*_{2,t+1}))^{b+1} C_{22}.$$

Next, under $m_1, m_2 > 0$ and $C_{11}, C_{22} > 0$, it holds that $R_1(m) \in (0, 1)$ and $R_2(m) > 0$ almost surely. This establishes the optimal ex post adjustment, $(\mathbb{A}^*_1, \mathbb{A}^*_2)$.

Under-Investment under Algebraic Condition. Suppose that $m_1, m_2, C_{11}, C_{22} > 0$. We first prove that $\Delta\mathbb{F}x_{1,t+1} < \Delta\mathbb{F}x_{1,t+1} + \mathbb{A}_1^*$. We proved previously that,

$$\mathbb{A}_1^* = R_1(m)[\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1}] + R_2(m)[\Delta x_{2,t+1}^* - \Delta\mathbb{F}x_{2,t+1}],$$

where $R_1(m), R_2(m) > 0$. Thus, if forecasts for both inputs are downwards biased, $\Delta x_{i,t+1}^* - \Delta\mathbb{F}x_{i,t+1} > 0$, then optimal ex-post adjustment for input 1 is strictly positive, $\mathbb{A}_1^* > 0$, and we have $\Delta\mathbb{F}x_{1,t+1} < \Delta\mathbb{F}x_{1,t+1} + \mathbb{A}_1^*$.

Next, we show that $\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1} - \mathbb{A}_1^* > 0$. From (116),

$$\begin{aligned}\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1} - \mathbb{A}_1^* &= (1 - R_1(m))[\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1}] - R_2(m)[\Delta x_{2,t+1}^* - \Delta\mathbb{F}x_{2,t+1}] \qquad (117)\\ &= \lambda_1 \left[\frac{m_1^2 C_{11}}{(p_{1,t+1})^2} \cdot a[\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1}] - \frac{m_2^2 C_{22}}{(p_{2,t+1})^2} \cdot b[\Delta x_{2,t+1}^* - \Delta\mathbb{F}x_{2,t+1}]\right],\end{aligned}$$

with

$$\lambda_1 = \frac{\frac{b}{a}(\bar{x}_{1,t}\exp(\Delta x_{1,t+1}^*))^{a+1}(\bar{x}_{2,t}\exp(\Delta x_{2,t+1}^*))^{b-1}(p_{1,t+1})^2}{ak_1 + ak_2 + bk_3 + \frac{a^2}{b}k_4}.$$

Under the algebraic condition:

$$\frac{m_1^2 C_{11}}{(p_{1,t+1})^2} \cdot a[\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1}] > \frac{m_2^2 C_{22}}{(p_{2,t+1})^2} \cdot b[\Delta x_{2,t+1}^* - \Delta\mathbb{F}x_{2,t+1}],$$

it follows from (117) that $\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1} - \mathbb{A}_1^* > 0$, so the realized growth rate of input $i = 1$ is strictly above the initial forecast and strictly below the cost-minimizing target. Otherwise, under the opposite algebraic condition,

$$\frac{m_1^2 C_{11}}{(p_{1,t+1})^2} \cdot a[\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1}] < \frac{m_2^2 C_{22}}{(p_{2,t+1})^2} \cdot b[\Delta x_{2,t+1}^* - \Delta\mathbb{F}x_{2,t+1}],$$

it is the realized growth rate of input $i = 2$ to be strictly above the initial forecast and strictly below the cost-minimizing target.

Next, we prove that, under the algebraic condition:

$$\frac{m_1^2 C_{11}}{(p_{1,t+1})^2} \cdot a[\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1}] > \frac{m_2^2 C_{22}}{(p_{2,t+1})^2} \cdot b[\Delta x_{2,t+1}^* - \Delta\mathbb{F}x_{2,t+1}],$$

the extent of under-investment, $\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1} - \mathbb{A}_1^* > 0$ increases with the extent of downward bias. From (117),

$$\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1} - \mathbb{A}_1^* = (1 - R_1(m))[\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1}] - R_2(m)[\Delta x_{2,t+1}^* - \Delta\mathbb{F}x_{2,t+1}].$$

Since $R_1(m) \in (0,1)$ and does not depend on $\Delta\mathbb{F}x_{1,t+1}$, it follows that $\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1} - \mathbb{A}_1^*$ increases with greater downward bias in the forecast for input 1, $\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1}$.

Now, we prove that, if (107) holds,

$$\log TC(\Delta\mathbb{F}x_{1,t+1}, \Delta\mathbb{F}x_{2,t+1}, \mathbb{A}_1^*, \mathbb{A}_2^*) - \log TC(\Delta x_{1,t+1}^*, \Delta x_{2,t+1}^*, 0, 0) \tag{118}$$

$$= \underbrace{\frac{a}{2b}[\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1} - \mathbb{A}_1^*]^2}_{\text{Forecasting Bias Effect}} + \underbrace{\varkappa_1 \cdot (\mathbb{A}_1^*)^2 + \varkappa_2 \cdot (\mathbb{A}_2^*)^2}_{\text{Adjustment Effect}}, \tag{119}$$

where $\varkappa_1$ and $\varkappa_2$ are constants, and $TC(\Delta\mathbb{F}x_{1,t+1}, \Delta\mathbb{F}x_{2,t+1}, \mathbb{A}_1^*, \mathbb{A}_2^*)$ are total costs, inclusive of input costs and adjustment costs.

To do so, consider the total cost function,

$$TC(\Delta\mathbb{F}x_{1,t+1}, \Delta\mathbb{F}x_{2,t+1}, \mathbb{A}_1, \mathbb{A}_2) =$$
$$p_{1,t+1}\bar{x}_{1,t}\exp(\Delta\mathbb{F}x_{1,t+1} + \mathbb{A}_1) + p_{2,t+1}\bar{x}_{2,t}\exp(\Delta\mathbb{F}x_{2,t+1} + \mathbb{A}_2) + \mathbb{C}(\mathbb{A}_1, \mathbb{A}_2).$$

Denote $\tilde{\Delta} = (\Delta\mathbb{F}x_{1,t+1}, \Delta\mathbb{F}x_{2,t+1}, \mathbb{A}_1^*, \mathbb{A}_2^*)$ and $\Delta = (\Delta x_{1,t+1}^*, \Delta x_{2,t+1}^*, 0, 0)$. A second order Taylor Expansion of $(\delta_1, \delta_2, \xi_1, \xi_2) \mapsto \log TC(\delta_1, \delta_2, \xi_1, \xi_2)$ around Δ yields,

$$\log TC(\tilde{\Delta}) = \log TC(\Delta) + \frac{1}{TC(\Delta)}\frac{\partial TC(\Delta)}{\partial x^{\mathrm{T}}}(\tilde{\Delta} - \Delta)$$
$$+ \frac{1}{2}(\tilde{\Delta} - \Delta)^{\mathrm{T}}\left[-\frac{1}{TC(\Delta)^2}\frac{\partial TC(\Delta)}{\partial x}\frac{\partial TC(\Delta)}{\partial x^{\mathrm{T}}} + \frac{1}{TC(\Delta)}\frac{\partial^2 TC(\Delta)}{\partial x \partial x^{\mathrm{T}}}\right](\tilde{\Delta} - \Delta).$$

Since $C(0,0) = 0$ and $C(x_1, x_2) > 0$ for all $(x_1, x_2) \neq 0$, it holds that $\partial C(0,0)/\partial x_1 = \partial C(0,0)/\partial x_2 = 0$. This combined with the assumption that $C_{12} = 0$ implies,

$$\frac{\partial TC(\Delta)}{\partial x} = \begin{bmatrix} r_1 \\ r_2 \\ r_1 \\ r_2 \end{bmatrix} \quad \frac{\partial TC(\Delta)}{\partial x}\frac{\partial TC(\Delta)}{\partial x^{\mathrm{T}}} = \begin{bmatrix} r_1^2 & r_1 r_2 & r_1^2 & r_1 r_2 \\ r_1 r_2 & r_2^2 & r_1 r_2 & r_2^2 \\ r_1^2 & r_1 r_2 & r_1^2 & r_1 r_2 \\ r_1 r_2 & r_2^2 & r_1 r_2 & r_2^2 \end{bmatrix} \quad TC(\Delta) = r_1 + r_2$$

$$\frac{\partial^2 TC(\Delta)}{\partial x \partial x^{\mathrm{T}}} = \begin{bmatrix} r_1 & 0 & r_1 & 0 \\ 0 & r_2 & 0 & r_2 \\ r_1 & 0 & r_1 + v_1 & 0 \\ 0 & r_2 & 0 & r_2 + v_2 \end{bmatrix} \quad \begin{bmatrix} r_1 \\ r_2 \\ v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} p_{1,t+1}\bar{x}_{1,t}\exp(\Delta x_{1,t+1}^*) \\ p_{2,t+1}\bar{x}_{2,t}\exp(\Delta x_{2,t+1}^*) \\ (m_1\bar{x}_{1,t}\exp(\Delta x_{1,t+1}^*))^2 C_{11} \\ (m_2\bar{x}_{2,t}\exp(\Delta x_{2,t+1}^*))^2 C_{22}. \end{bmatrix}$$

Next, using (109),

$$r_1 = \frac{aTC(\Delta)}{a+b} \quad r_2 = \frac{bTC(\Delta)}{a+b} \quad r_1 = \frac{a}{b}r_2. \tag{120}$$

This, combined with (114) and (115), yields,

$$r_1\mathbb{A}_1^* + r_2\mathbb{A}_2^* = \frac{a}{b}r_2[\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1}] + r_2[\Delta x_{2,t+1}^* - \Delta\mathbb{F}x_{2,t+1}]$$
$$= r_1[\Delta x_{1,t+1}^* - \Delta\mathbb{F}x_{1,t+1}] + r_2[\Delta x_{2,t+1}^* - \Delta\mathbb{F}x_{2,t+1}]$$

Thus,

$$\frac{1}{TC(\Delta)}\frac{\partial TC(\Delta)}{\partial x^{\mathrm{T}}}(\tilde{\Delta}-\Delta)=\frac{-r_1[\Delta x^*_{1,t+1}-\Delta\mathbb{F}x_{1,t+1}]-r_2[\Delta x^*_{2,t+1}-\Delta\mathbb{F}x_{2,t+1}]+r_1\mathbb{A}^*_1+r_2\mathbb{A}^*_2}{TC(\Delta)}=0 \tag{121}$$

Next, using (120), we obtain,

$$-\frac{1}{TC(\Delta)^2}\frac{\partial TC(\Delta)}{\partial x}\frac{\partial TC(\Delta)}{\partial x^{\mathrm{T}}}=-\frac{1}{(a+b)^2}\begin{bmatrix}a^2 & ab & a^2 & ab\\ ab & b^2 & ab & b^2\\ a^2 & ab & a^2 & ab\\ ab & b^2 & ab & b^2\end{bmatrix}$$

$$\frac{1}{TC(\Delta)}\frac{\partial^2 TC(\Delta)}{\partial x\partial x^{\mathrm{T}}}=\frac{1}{a+b}\begin{bmatrix}a & 0 & a & 0\\ 0 & b & 0 & b\\ a & 0 & a & 0\\ 0 & b & 0 & b\end{bmatrix}+\begin{bmatrix}0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & v_1/TC(\Delta) & 0\\ 0 & 0 & 0 & v_2/TC(\Delta).\end{bmatrix}$$

Thus,

$$\begin{aligned}&\frac{1}{TC(\Delta)}\frac{\partial^2 TC(\Delta)}{\partial x\partial x^{\mathrm{T}}}-\frac{1}{TC(\Delta)^2}\frac{\partial TC(\Delta)}{\partial x}\frac{\partial TC(\Delta)}{\partial x^{\mathrm{T}}}\\ &\qquad=\frac{1}{(a+b)^2}\begin{bmatrix}ab & -ab & ab & -ab\\ -ab & ab & -ab & ab\\ ab & -ab & ab & -ab\\ -ab & ab & -ab & ab\end{bmatrix}+\begin{bmatrix}0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & v_1/TC(\Delta) & 0\\ 0 & 0 & 0 & v_2/TC(\Delta).\end{bmatrix}\end{aligned}$$

Let $\zeta=(1,-1,1,-1)^T$. Using (115) and (114),

$$(\tilde{\Delta}-\Delta)^{\mathrm{T}}\zeta=-\left(\frac{a+b}{b}\right)[\Delta x^*_{1,t+1}-\Delta\mathbb{F}x_{1,t+1}-\mathbb{A}^*_1].$$

Thus,

$$\begin{aligned}\frac{1}{(a+b)^2}(\tilde{\Delta}-\Delta)^{\mathrm{T}}\begin{bmatrix}ab & -ab & ab & -ab\\ -ab & ab & -ab & ab\\ ab & -ab & ab & -ab\\ -ab & ab & -ab & ab\end{bmatrix}(\tilde{\Delta}-\Delta)&=\frac{ab}{(a+b)^2}(\tilde{\Delta}-\Delta)^{\mathrm{T}}\zeta\zeta^{\mathrm{T}}(\tilde{\Delta}-\Delta)\\ &=\frac{ab}{(a+b)^2}\frac{(a+b)^2}{b^2}[\Delta x^*_{1,t+1}-\Delta\mathbb{F}x_{1,t+1}-\mathbb{A}^*_1]^2\\ &=\frac{a}{b}[\Delta x^*_{1,t+1}-\Delta\mathbb{F}x_{1,t+1}-\mathbb{A}^*_1]^2.\end{aligned}$$

Moreover,

$$(\tilde{\Delta}-\Delta)^{\mathrm{T}}\begin{bmatrix}0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & v_1/TC(\Delta) & 0\\ 0 & 0 & 0 & v_2/TC(\Delta)\end{bmatrix}(\tilde{\Delta}-\Delta)=\frac{v_1}{TC(\Delta)}(\mathbb{A}^*_1)^2+\frac{v_2}{TC(\Delta)}(\mathbb{A}^*_2)^2.$$

Thus, using (121),

$$\log TC(\tilde{\Delta}) = \log TC(\Delta) + \frac{a}{2b}[\Delta x^*_{1,t+1} - \Delta \mathbb{F} x_{1,t+1} - \mathbb{A}^*_1]^2 + \frac{v_1}{2 \cdot TC(\Delta)}(\mathbb{A}^*_1)^2 + \frac{v_2}{2 \cdot TC(\Delta)}(\mathbb{A}^*_2)^2,$$

which establishes the result with $\varkappa_i = \dfrac{v_i}{2 \cdot TC(\Delta)} > 0$ (a.s.) since $C_{ii} \neq 0$ by assumption.
QED

A.6 CES Case

This subsection extends the benchmark Cobb-Douglas analysis to a CES technology. Relative to Lemma 2 in the main text, the CES case weakens exact forecast-error equalities into local monotonicity, sign, and inequality restrictions. The purpose of this subsection is to show that the conceptual distinction between coherence and accuracy does not hinge on exact log-linearity.

In subsection A.6.1 we derive coherence restrictions on forecasts under CES production function, and in subsection A.6.2 we derive coherence restrictions on forecast errors under CES production function.

A.6.1 Coherence in Forecasts—CES Case

We derive coherence restrictions on forecasts under CES production function. This subsection derives an inequality condition, which we will then implement empirically in section B.2 of this Web Appendix below.

Proposition A1 (Forecast Coherence Under CES Production). *Forecast coherence requires that forecasts of output and inputs, $\mathbb{F}_t[y_{t+1}]$, $\mathbb{F}_t[x_{1,t+1}]$, and $\mathbb{F}_t[x_{2,t+1}]$, satisfy an inequality. For $a + b \leq \xi \leq 1$, the CES function is concave and forecast coherence requires,*

$$\mathbb{F}_t[y_{t+1}] \leq f(\mathbb{F}_t[x_{1,t+1}], \mathbb{F}_t[x_{2,t+1}]) = \left(\frac{a}{a+b}\mathbb{F}_t[x_{1,t+1}]^\xi + \frac{b}{a+b}\mathbb{F}_t[x_{2,t+1}]^\xi\right)^{\frac{a+b}{\xi}}. \tag{122}$$

For $\xi \geq 1$ and $a + b \geq 1$, the CES function is convex and the inequality swaps sign.

$$\mathbb{E}_t[y_{t+1}] \geq f(\mathbb{E}_t[x_{1,t+1}], \mathbb{E}_t[x_{2,t+1}]) = \left(\frac{a}{a+b}\mathbb{E}_t[x_{1,t+1}]^\xi + \frac{b}{a+b}\mathbb{E}_t[x_{2,t+1}]^\xi\right)^{\frac{a+b}{\xi}}. \tag{123}$$

Proof of Proposition A1. Recall that for a concave function, f, it holds that $\mathbb{E}[f(x)] \leq f(\mathbb{E}[x])$. Assume $\xi \leq 1$ and start by assuming that $a + b = 1$. The CES function f is homogeneous of degree one because, for a scalar λ,

$$f(\lambda \mathbf{x}) = \left[a(\lambda x_1)^\xi + b(\lambda x_2)^\xi\right]^{\frac{1}{\xi}} = \lambda\left(a x_1^\xi + b x_2^\xi\right)^{\frac{1}{\xi}}.$$

Furthermore, note that f is also quasiconcave because it is a monotone transformation of a concave function. In fact,

$$f = g^{\frac{1}{\xi}},$$

and to see that g is concave, compute its Hessian, H_g,

$$H_g = \begin{bmatrix} \frac{\partial^2 g}{\partial x_1^2} & \frac{\partial^2 g}{\partial x_2 \partial x_1} \\ \frac{\partial^2 g}{\partial x_1 \partial x_2} & \frac{\partial^2 g}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} a\xi\left(\xi-1\right)x_1^{\xi-2} & 0 \\ 0 & b\xi\left(\xi-1\right)x_2^{\xi-2} \end{bmatrix}.$$

Since H_g is negative semi-definite, we can conclude that g is concave.

Now, let $a+b \leq 1$. As a result, $f = \left(g^{\frac{1}{\xi}}\right)^{a+b}$, where g is concave, as shown above. Then, f is a concave increasing function of a concave function, from which we can conclude that f is concave, which proves the proposition.

QED

Proposition A1 provides a first restriction, in the form of an inequality condition, ex ante coherent forecasts of output and inputs should satisfy. This inequality can be implemented empirically. We provide an illustration in section B of this Web Appendix. Relative to some of the restrictions and test statistics involving forecast errors we derive in the paper, the inequality condition of Proposition A1 has two appealing features. First, it applies to a fairly general class of production functions. Second, it requires data on forecasts only. However, a credible empirical implementation of this inequality would require firm-specific estimates of the technology parameters our data does not allow.

A.6.2 Coherence in Forecast Errors—CES Case

We now derive coherence restrictions on forecast errors under a CES production function. The purpose of this subsection is to show that the empirical results on forecast errors reported in the main text remain interpretable in light of theory even after relaxing the Cobb-Douglas assumption.

Setup. Let output be generated by a CES technology,

$$y_{t+1} = f(x_{1,t+1}, x_{2,t+1}), \qquad \text{with } f \text{ increasing in each input.} \tag{124}$$

A CFO issues forecasts $(\mathbb{F}_t y_{t+1}, \mathbb{F}_t x_{1,t+1}, \mathbb{F}_t x_{2,t+1})$ at time t. We define forecast errors as,

$$\mathbb{FE}_t[z_{t+1}] \equiv z_{t+1} - \mathbb{F}_t z_{t+1}, \qquad z \in \{y, x_1, x_2\}. \tag{125}$$

Coherence under CES (levels). By Proposition A1 above, when f is concave, forecast coherence implies the Jensen-type inequality,

$$\mathbb{F}_t y_{t+1} \leq f(\mathbb{F}_t x_{1,t+1}, \mathbb{F}_t x_{2,t+1}), \tag{126}$$

(with the inequality reversing when f is convex). This is the CES analogue of the Cobb–Douglas log-linearity condition and is developed above in subsection A.6.1 in this Web Appendix.

From coherence to a forecast-error restriction. Write realized inputs as $x_{i,t+1} = \mathbb{F}_t x_{i,t+1} + \mathbb{FE}_t[x_{i,t+1}]$ and expand f around the forecast point, $\mathbb{F}_t x_{t+1} \equiv (\mathbb{F}_t x_{1,t+1}, \mathbb{F}_t x_{2,t+1})$:

$$\begin{aligned} y_{t+1} &= f(x_{1,t+1}, x_{2,t+1}) \\ &= f(\mathbb{F}_t x_{t+1}) + \nabla f(\mathbb{F}_t x_{t+1})^\top \mathbb{FE}_t[x_{t+1}] + R_{t+1}, \end{aligned}$$

where $\mathbb{FE}_t[x_{t+1}] \equiv (\mathbb{FE}_t[x_{1,t+1}], \mathbb{FE}_t[x_{2,t+1}])$ and R_{t+1} is a second-order remainder.

Subtracting $\mathbb{F}_t y_{t+1}$ and using coherence ($\mathbb{F}_t y_{t+1} \leq f(\mathbb{F}_t x_{t+1})$ under concavity) yields the *CES forecast-error inequality*,

$$\mathbb{FE}_t[y_{t+1}] \;\; \geq \;\; \underbrace{\nabla f(\mathbb{F}_t x_{t+1})^\top \mathbb{FE}_t[x_{t+1}]}_{\text{local "technology-weighted" input } \mathbb{FE}} \;\; + R_{t+1} \qquad \text{(concave CES).} \tag{127}$$

If f is convex, the inequality reverses.

Interpretation and link to the empirical tests. Equation (127) is the counterpart of Lemma 2 in the main text. Under Cobb–Douglas log-linearity, Lemma 2, which is the analogue of (127), holds *with equality and without a remainder term,* R_{t+1}. Under CES, coherence still links output forecast errors to input forecast errors, but: (i) the relationship is *local* (weights are marginal products evaluated at the forecast point), (ii) it is typically an *inequality*, and (iii) there is a second-order remainder term, R_{t+1}.

Two practical implications follow.

(1) Qualitative validity. Because CES technologies are increasing in each input, the components of $\nabla f(\mathbb{F}_t x_{t+1})$ are nonnegative. Therefore, coherence imposes monotonicity/sign restrictions: ceteris paribus, unusually positive (negative) input forecast errors push the output forecast error in the same direction locally. This is the sense in which, under CES, coherence delivers *qualitative* forecast-error restrictions rather than exact equalities.

(2) Why the Cobb–Douglas-based regression tests remain interpretable. If forecast errors are not too large (so that the second-order remainder, R_{t+1}, is small on average), then (127) implies that the empirical residual from the Cobb–Douglas equality-based restriction is approximately one-sided (up to a small approximation error). Equivalently, the Cobb–Douglas equality used in the main text can be viewed as a convenient *first-order approximation* to the CES restriction (127). This is why the forecast-error tests remain informative about coherence even if the true technology is CES (but not Cobb–Douglas): under CES, coherence still requires that output forecast errors co-move with (technology-weighted) input forecast errors, though the restriction is weaker (inequality + remainder) than under Cobb–Douglas.

Implementation note. In principle, one could implement (127) directly by replacing constant elasticities with CES marginal products at $\mathbb{F}_t x_{t+1}$. In practice, this would require additional measurement of the forecast-point technology weights (firm- or industry-specific); the main text focuses on the Cobb–Douglas benchmark, while the CES case clarifies that the conceptual distinction between coherence and accuracy does not hinge on exact log-linearity. Therefore, the Cobb–Douglas equality-based tests in the main text can be interpreted as a transparent first-order benchmark, while the CES case preserves the same qualitative restrictions on the joint behavior of forecast errors.

A.7 A Model of Narrow Bracketing in Corporate Forecasts

Consider a CFO's self making a forecast about the first input, $\mathbb{F} \log x_1$, by

$$\min_{\mathbb{F} \log x_1} \mathbb{E}\left[(\log x_1 - \mathbb{F} \log x_1)^2 \, | \Omega \right].$$

The forecaster observes two noisy signals, one about the output, η_y, and one about the second input, η_2.[21] Specifically, $\eta_y = \log y + \epsilon_y$ and $\eta_2 = \log x_2 + \epsilon_2$, where $\epsilon_y \sim \mathcal{N}\left(\mu_y, s_y^2\right)$, $\epsilon_2 \sim \mathcal{N}\left(\mu_2, s_2^2\right)$, and $y = x_1^a \cdot x_2^b$. Under these assumptions, we obtain Proposition 2 in the main text in the case of uncorrelated prices. Below, we state and prove Proposition 2 for the more general case with correlated prices.

Proof of Proposition 2 (Optimal forecasts with noisy signals). The Proposition is stated in the text for the case of $\rho_{1,2} = 0$. Here we prove the Proposition for the general case with correlated prices, i.e., for a generic value of $\rho_{1,2} \in [0.1]$,

$$\mathbb{E}\left[\log x_1 | \eta_y, \eta_2\right] = \mu_1 + \beta_y\left(\eta_y - \mu_y\right) + \beta_2\left(\eta_2 - \mu_2\right),$$

where coefficients equal

$$\beta_y = \frac{\operatorname{Cov}\left(\eta_y, \log x_1\right) - \left(\frac{\operatorname{Cov}(\eta_y,\eta_2)\operatorname{Cov}(\log x_1,\eta_2)}{\operatorname{Var}(\eta_2)}\right)}{\operatorname{Var}\left(\eta_y\right) - \frac{\operatorname{Cov}(\eta_y,\eta_2)^2}{\operatorname{Var}(\eta_2)}}, \quad \beta_2 = \frac{\operatorname{Cov}\left(\eta_2, \log x_1\right) - \left(\frac{\operatorname{Cov}(\eta_y,\eta_2)\operatorname{Cov}(\log x_1,\eta_y)}{\operatorname{Var}(\eta_y)}\right)}{\operatorname{Var}\left(\eta_2\right) - \frac{\operatorname{Cov}(\eta_y,\eta_2)^2}{\operatorname{Var}(\eta_y)}}.$$

The following hold:
$\operatorname{Cov}\left(\eta_y, \log x_1\right) = cov\left(a\log x_1 + b\log x_2 + \epsilon_y, \log x_1\right) = a\sigma_1^2 + b\rho_{1,2}$,
$\operatorname{Cov}\left(\eta_2, \log x_1\right) = \rho_{1,2}$,
$\operatorname{Cov}\left(\eta_y, \eta_2\right) = cov\left(a\log x_1 + b\log x_2 + \epsilon_y, \log x_2 + \epsilon_2\right) = b\sigma_2^2 + a\rho_{1,2}$,
$\operatorname{Var}\left(\eta_y\right) = a^2\sigma_1^2 + b^2\sigma_2^2 + s_y^2 + 2ab\rho_{1,2}$,
$\operatorname{Var}\left(\eta_2\right) = \sigma_2^2 + s_2^2$.
Substituting the above in the expressions for β_y and β_2 yields,

$$\begin{aligned}
\beta_y &= \frac{a\sigma_1^2 + b\rho_{1,2} - \frac{\rho_{1,2}\left(a\rho_{1,2}+b\sigma_2^2\right)}{\sigma_2^2+s_2^2}}{a^2\sigma_1^2 + b^2\sigma_2^2 + s_y^2 + 2ab\rho_{1,2} - \frac{\left(a\rho_{1,2}+b\sigma_2^2\right)^2}{\sigma_2^2+s_2^2}} \\
\beta_2 &= \frac{\rho_{1,2} - \frac{\left(a\rho_{1,2}+b\sigma_2^2\right)\left(a\sigma_1^2+b\rho_{1,2}\right)}{a^2\sigma_1^2+b^2\sigma_2^2+s_y^2+2ab\rho_{1,2}}}{\sigma_2^2 + s_2^2 - \frac{\left(a\rho_{1,2}+b\sigma_2^2\right)^2}{a^2\sigma_1^2+b^2\sigma_2^2+s_y^2+2ab\rho_{1,2}}}.
\end{aligned}$$

For $\rho_{1,2} = 0$, we obtain the result stated in the text,

$$\begin{aligned}
\beta_y &= \frac{a\sigma_1^2}{a^2\sigma_1^2 + b^2\sigma_2^2 + s_y^2 - \frac{b^2\sigma_2^4}{\sigma_2^2+s_2^2}}, \\
\beta_2 &= \frac{-\frac{\left(b\sigma_2^2\right)\left(a\sigma_1^2\right)}{a^2\sigma_1^2+b^2\sigma_2^2+s_y^2}}{\sigma_2^2 + s_2^2 - \frac{b^2\sigma_2^4}{a^2\sigma_1^2+b^2\sigma_2^2+s_y^2}} = -\frac{ab\sigma_1^2\sigma_2^2}{a^2\sigma_1^2 + b^2\sigma_2^2 + s_y^2} \times \frac{a^2\sigma_1^2 + b^2\sigma_2^2 + s_y^2}{\left(\sigma_2^2 + s_2^2\right)\left(a^2\sigma_1^2 + b^2\sigma_2^2 + s_y^2\right) - b^2\sigma_2^4} \\
&= \frac{ab\sigma_1^2\sigma_2^2}{b^2\sigma_2^4 - \left(\sigma_2^2 + s_2^2\right)\left(a^2\sigma_1^2 + b^2\sigma_2^2 + s_y^2\right)}.
\end{aligned}$$

QED

Proof of Corollary 2.1 (Narrow Bracketing).

[21] This is wlog, as inputs can be relabelled $(i, \neg i)$. Also, by stationarity we drop time subscripts.

$$\begin{aligned}
\lim_{s_y, s_2 \to +\infty} \beta_y &= \lim_{s_y, s_2 \to +\infty} \frac{a\sigma_1^2}{a^2\sigma_1^2 + b^2\sigma_2^2 + s_y^2 - \frac{b^2\sigma_2^4}{\sigma_2^2+s_2^2}} = 0, \\
\lim_{s_y, s_2 \to +\infty} \beta_2 &= \lim_{s_y, s_2 \to +\infty} \frac{ab\sigma_1^2\sigma_2^2}{b^2\sigma_2^4 - \left(\sigma_2^2 + s_2^2\right)\left(a^2\sigma_1^2 + b^2\sigma_2^2 + s_y^2\right)} = 0.
\end{aligned}$$

QED

Proof of Corollary 2.2 (Univariate Projections).

$$\begin{aligned}
\lim_{s_2 \to +\infty} \beta_y &= \lim_{s_2 \to +\infty} \frac{a\sigma_1^2}{a^2\sigma_1^2 + b^2\sigma_2^2 + s_y^2 - \frac{b^2\sigma_2^4}{\sigma_2^2+s_2^2}} = \frac{a\sigma_1^2}{a^2\sigma_1^2 + b^2\sigma_2^2 + s_y^2}, \\
\lim_{s_2 \to +\infty} \beta_2 &= \lim_{s_2 \to +\infty} \frac{ab\sigma_1^2\sigma_2^2}{b^2\sigma_2^4 - \left(\sigma_2^2 + s_2^2\right)\left(a^2\sigma_1^2 + b^2\sigma_2^2 + s_y^2\right)} = 0.
\end{aligned}$$

QED

Because nothing necessarily implies that $\beta_y = 1$ and $\mu_1 - \beta_y\mu_y = 0$, (R2) is generally inferior to (R3) even in a second-best world.

Finally, we consider for completeness the case in which both signals are infinitely precise.

Corollary 2.3 (Precise Signals). *When* $s_y^2, s_2^2 \to 0$, *the optimal forecast is* $\mathbb{E}\left[\log x_1 | \eta_y, \eta_2\right] = \frac{1}{a}\left(\eta_y - b\eta_2\right)$.

Proof of Corollary 2.3.

$$\begin{aligned}
\lim_{s_y, s_2 \to 0} \beta_y &= \lim_{s_y, s_2 \to 0} \frac{a\sigma_1^2}{a^2\sigma_1^2 + b^2\sigma_2^2 + s_y^2 - \frac{b^2\sigma_2^4}{\sigma_2^2+s_2^2}} = \frac{a\sigma_1^2}{a^2\sigma_1^2 + b^2\sigma_2^2 + - \frac{b^2\sigma_2^4}{\sigma_2^2}} = \frac{1}{a}, \\
\lim_{s_y, s_2 \to 0} \beta_2 &= \lim_{s_y, s_2 \to 0} \frac{ab\sigma_1^2\sigma_2^2}{b^2\sigma_2^4 - \left(\sigma_2^2 + s_2^2\right)\left(a^2\sigma_1^2 + b^2\sigma_2^2 + s_y^2\right)} = \frac{ab\sigma_1^2\sigma_2^2}{b^2\sigma_2^4 - \sigma_2^2\left(a^2\sigma_1^2 + b^2\sigma_2^2\right)} = -\frac{b}{a}.
\end{aligned}$$

QED

Note the results in this subsection rationalize versions of rules (R5), (R1), and (R3) under different signal structures; they are not intended to reproduce every empirical implementation detail of the textbook rules.

B. Empirical Implementation of Theoretical Results

B.1 Dependence of Proposition 1 on Input Price Comovement

The ranking in Proposition 1 depends on the degree of comovement in input price innovations. In the limit case where input prices move very closely together, the multivariate rule (R5) weakly dominates more restrictive rules in expected squared error. When input price innovations are largely independent, pairwise orderings among simpler rules may vary.

To assess empirical relevance, we compute the correlation between annual growth in U.S. labor input prices (Employment Cost Index, private industry wages and salaries) and capital goods prices (PPI for Final Demand Capital Equipment) over 2001-2019, corresponding to our firm-level sample period. The correlation is positive ($\rho = 0.24$), indicating non-trivial input price comovement. For completeness, we note that considering the longer period 2001-2025, including the more recent high-inflation years outside our sample window, the same correlation rises to 0.7.

B.2 Implementation of Inequality Condition

Here we implement the inequality condition of Proposition A1, developed under a general CES function. We view this inequality as imposing on the data as little restriction as possible. We compute a and b using the universe of industries from the Bureau of Economic Analysis and find that $a + b \leq 1$. Furthermore, the elasticity of substitution between capital and labor in the US economy, denoted with χ, is typically found to be between 0.5 and 1 (e.g., see Berndt (1976) and Oberfield and Raval (2021)), where $\chi = 1$ defines the Cobb-Douglas production function. Therefore, the CES function is weakly concave and thus inequality (122) is the relevant one. We account for heterogeneity by allowing a and b to vary by industry and by presenting our results for three different values of the elasticity of substitution between capital and labor, $\chi = 0.5, 0.7, 0.9$. We implement our inequality restriction both in levels and in growth rates.[22]

Table A11 of the Web Appendix reports our results. Panel A shows that most CFOs' forecasts violate the inequality restriction of Proposition 1. In levels, almost all CFOs give joint forecasts of capital, labor, and output that jointly violate the inequality. However, as just discussed, results in levels should be seen with caution, as they refer to a much smaller sample given the limitations of Compustat data on wages. In growth rates, about 73% of CFOs' forecasts violate the inequality. These results are quite stable across different values of the elasticity of substitution between capital and labor. If anything, moving toward $\chi = 1$ (Cobb-Douglas) appears to give a slightly better shot at CFOs to give coherent forecasts, perhaps because much business teaching uses examples based on the Cobb-Douglas.

Panel B reports summary statistics of the difference between the left-hand side and the right-hand side of the inequality of Proposition 1. Most CFOs forecast a growth of output that is larger than the output growth implied by feeding into a CES production function the CFOs' forecasts of capital and labor input growth.

[22]The CES coherence restriction is naturally expressed in levels. Around the forecast point, it induces local sign and monotonicity restrictions on forecast errors and their first-order approximations in growth rates, which is the sense in which it remains empirically informative here.

C. Individual-Level Tests

Under the assumptions of Section 3, we derive,
Proposition A2 (Individual-Level Test). *If $\xi \to 0$ and $\rho_{1,2} = 0$, under the null hypothesis of coherent forecasts, the following holds:*

$$\text{C1-stat} \equiv \frac{\frac{\mathbb{F}_t \log y_{t+1} - a\mathbb{F}_t \log x_{1,t+1}}{b} - \log \frac{b}{a+b} Z}{\gamma_2 \sigma_2} \sim \mathcal{N}(0,1) \tag{128}$$

and

$$\text{C2-stat} \equiv \frac{\mathbb{F}\mathbb{E}_t \log y_{t+1} - a\mathbb{F}\mathbb{E}_t \log x_{1,t+1}}{\sigma_2 b} \sim \mathcal{N}(0,1), \tag{129}$$

where $\mathbb{F}\mathbb{E}_t \log y_{t+1} = \log y_{t+1} - \mathbb{F}_t \log y_{t+1}$ *and* $\mathbb{F}\mathbb{E}_t \log x_{1,t+1} = \log x_{1,t+1} - \mathbb{F}_t \log x_{1,t+1}$.
Proof of Proposition A2. To derive the test statistic, note that we have $\mathbb{E}_t \log y_{t+1} = a\mathbb{E}_t \log x_{1,t+1} + b\mathbb{E}_t \log x_{2,t+1}$ and $\log x_{2,t+1} = \log \frac{b}{a+b} Z - \pi_{2,t+1}$, implying that the CFO forecast for input 2 is

$$\mathbb{E}_t \log x_{2,t+1} = \log \frac{b}{a+b} Z - \gamma_2 \pi_{2,t}. \tag{130}$$

Note that (130) depends on the technological parameters, a and b, and on Z, which we assume are known to the CFO at the time of the forecast and are stable over time. From (130) we obtain that

$$\frac{\mathbb{E}_t \log x_{2,t+1} - \log \frac{b}{a+b} Z}{\gamma_2 \sigma_2} \sim \mathcal{N}(0,1).$$

We can then derive our test statistic C1-stat based on the joint forecasts of the first input and output by recalling that $\log y = a \log x_1 + b \log x_2$ as follows:

$$\text{C1-stat} \equiv \frac{\frac{\mathbb{E}_t \log y_{t+1} - a\mathbb{E}_t \log x_{1,t+1}}{b} - \log \frac{b}{a+b} Z}{\gamma_2 \sigma_2} \sim \mathcal{N}(0,1),$$

where the distribution here is obtained under the null hypothesis of coherent forecasts.

To derive our C2-stat, we start by defining the forecast error of a generic variable x forecasted at t and realized at $t+1$ as the difference between the realization and the forecast, $\text{FE}_t x_{t+1} = x_{t+1} - \mathbb{E}_t x_{t+1}$. We then have that

$$\begin{aligned} \text{FE}_t \log x_{2,t+1} &= \log x_{2,t+1} - \mathbb{E}_t \log x_{2,t+1} \\ &= \log \frac{b}{a+b} Z - \pi_{2,t+1} - \mathbb{E}_t \left[\log \frac{b}{a+b} Z - \pi_{2,t+1} \right] \\ &= -\text{FE}_t \pi_{2,t+1} = -\epsilon_{2,t+1}. \end{aligned}$$

As a result, the forecast error of the log of the second input is the negative of the innovation of the second log-price process. It follows that

$$\frac{\text{FE}_t \log x_{2,t+1}}{\sigma_2} \sim \mathcal{N}(0,1).$$

Noting that $\text{FE}_t \log y_{t+1} = a\text{FE}_t \log x_{1,t+1} + b\text{FE}_t \log x_{2,t+1}$, we obtain our C2-stat,

$$\text{C2-stat} \equiv \frac{\text{FE}_t \log y_{t+1} - a\text{FE}_t \log x_{1,t+1}}{\sigma_2 b} \sim \mathcal{N}(0,1).$$

QED

Coherence Test Statistic: Multiple Inputs Case. Here we generalize our C2-stat to a multivariate case with N inputs. For this subsection, the production setting is

$$y = \prod_{i=1}^{N} x_i^{a_i}$$

$$\boldsymbol{p}'\boldsymbol{x} = Z,$$

where $\boldsymbol{p}$ and $\boldsymbol{x}$ are the column vectors of factor prices and quantities, respectively. As in the $N = 2$ case we have a linear relationship between the logs of inputs and output,

$$\log y = \sum_{i=1}^{N} a_i \log x_i,$$

where the same equation holds for the forecast errors. Analogously to the bivariate case, we have that $\text{FE}_t \log x_{1,t+1} = -\epsilon_{1,t+1}$ so that

$$\frac{\text{FE}_t \log x_{1,t+1}}{\sigma_1} \sim \mathcal{N}(0, 1).$$

Then, using the linear relationship between the logs of inputs and output we obtain our generalized C2-stat,

$$\frac{\text{FE}_t \log y_{t+1} - \sum_{i=2}^{N} a_i \text{FE}_t \log x_{i,t+1}}{\sigma_1 a_1} \sim \mathcal{N}(0, 1).$$

Coherence Test Statistic: Discussion. This statistic has an intuitive interpretations. Under the null of coherence, the FEs of output and one input cannot be "too far" from each other in the sense of (129).[23] Proposition A2 is written for the 2-input case, as it is the one we implement empirically below. Relative to (4), whose implementation requires observing FEs on all n inputs, implementation of Proposition A2 only requires observing FEs for $n - 1$ inputs. This is an implication of the $\xi \to 0$ assumption.

Under the assumptions of Proposition A2, forecast accuracy for a generic variable x, corresponding to the output or any of the input, can be tested using $\frac{\mathbb{FE}_t \log x_{t+1}}{\sigma_x} \sim \mathcal{N}(0, 1)$. Implementation of this test along with the test in (129) enables us to empirically distinguish forecast accuracy from forecast coherence.

Figure 3 illustrates the theoretical connection between the concepts of forecast accuracy and forecast coherence for a firm whose input has a loading of 1/3. There are four conceptual cases, corresponding to the four areas of the figure. In the first area, the forecasts are both accurate and coherent, as both $\frac{\text{FE}_t \log x_{1,t+1}}{\sigma_1}$ and $\frac{\text{FE}_t \log y_{t+1}}{\sigma_y}$ are close to zero and also close to each other. In the second area, the forecasts are inaccurate but coherent, as $\frac{\text{FE}_t \log x_{1,t+1}}{\sigma_1}$, $\frac{\text{FE}_t \log y_{t+1}}{\sigma_y}$, or both are statistically different from zero but quite close to each other. In the third area, the forecasts are accurate but incoherent, as both $\frac{\text{FE}_t \log x_{1,t+1}}{\sigma_1}$ and $\frac{\text{FE}_t \log y_{t+1}}{\sigma_y}$ are close to zero but sufficiently

[23]Relative to Subsection 2.2.3, Proposition A2 clarifies that output-input FEs of opposite sign do not necessarily imply incoherence, as long as they are sufficiently close to each other. On the other hand, FEs of the same sign can reflect incoherence, if they are sufficiently far from each other.

apart from each other. In the fourth area, the forecasts are both inaccurate and incoherent, as $\frac{\mathrm{FE}_t \log x_{1,t+1}}{\sigma_1}$, $\frac{\mathrm{FE}_t \log y_{t+1}}{\sigma_y}$, or both are statistically different from zero and also far apart from each other.

We test for coherence using the C2-stat, as it does not require FEs on the second input (wages) or other firm-level characteristics hard to observe. For each CFO, we observe four items (two forecasts and two realizations) and we estimate three parameters (a, b, and σ_2) using aggregate industry data from the Bureau of Economic Analysis and Bureau of Labor Statistics. As a result, the C2-stat is distributed as a Student t with $N - K = 4 - 3 = 1$ degree of freedom.

Table A12 presents the results. In Panel A, we reject the null of coherence at the 95% confidence level for 55.7% of CFOs in our sample.[24] This result parallels our previous findings that 48% of CFOs make forecasts closest to incoherent RoTs, (R1) and (R2), and that 52% of CFOs violate the intuitive restrictions on FEs discussed in Subsection 2.2.3

Figures in Panel A further confirm that CFO forecasts are fairly accurate on output and fairly inaccurate on capital. Specifically, we reject the null of forecast accuracy at the 95% confidence level for 27.2% of CFOs with respect to sales revenue and for 47.9% of CFOs with respect to CapEx. When considering output and capital input forecasts together, we reject the null of accuracy for 57% of CFOs in our sample.

These results rule out an alternative interpretation of our earlier findings, that is, that CFOs wield control over sales and capital to a different extent. While CFOs have ultimate authority over their firm's capital expenditures, their firm's sales depend also on external factors, including customer behavior and product market competition. Thus, one might expect CFOs to be more accurate about CapEx than sales. This is the opposite of what we find. CFOs are much less accurate in forecasting CapEx, likely reflecting the fact that their CapEx forecast is incoherently linked to their planned sales.

Panel B of Table A12 assesses coherence and accuracy together. It shows that 31.1% of CFOs in our sample are coherent and accurate, 13.2% are coherent but inaccurate, 12.0% are accurate but incoherent, and 43.7% are incoherent and inaccurate (all at the 95% confidence level).[25] The summary statistics of the cross-sectional distribution of the calculated C2-stat and of the FEs on output and capital input are shown in Panel C.

One concern with these results is the extent to which the computed C2-stat, as well as the accuracy statistics, are sensitive to the uncertainty coming from estimating the parameters a, b, and σ_2. To address this concern, we perform a bootstrap procedure, as follows. For each CFO, we generate 1,000 bootstrap repetitions of the C2-stat.[26] Using these 1,000 replications, we compute the fraction of cases out of 1,000 for which we reject the null of coherence at the 95% and 99% confidence levels. Hence, for each CFO and confidence level, the computed statistic is a number between 0 and 1, where 0 means that the null of coherence was never rejected across all 1,000 repetitions and 1 means that the null of coherence was rejected for all 1,000 repetitions.

[24] At the 99% confidence level, we reject coherence for 7.7% of CFOs. This difference reflects the distribution of the Student t with 1 degree of freedom.

[25] The figures at the 99% confidence level are 89.4%, 2.9%, 3.4%, and 4.3%, respectively.

[26] Specifically, for each of the 15 BEA industries and pair of consecutive years between 1987 and 2018, we resample observations with replacement 1,000 times (aka bootstrap replications). At each replication, we obtain an estimate of σ_1, based on the residual sum of squares (RSS) of the regression of total capital compensation on its lag, and an estimate of σ_2, based on the RSS of the regression of total labor compensation on its lag. We use cluster bootstrap with 6 clusters, corresponding to the following year windows: 1988-1992, 1993-1997, 1998-2002, 2003-2007, 2008-2012, and 2013-2018. We additionally generate bootstrap estimates of a and b, as the capital and labor shares of total factor compensation. Endowed with bootstrap estimates of a, b, σ_1^2, and σ_2^2, we derive corresponding estimates for σ_y^2, for the FEs of the log of output (y) and of the log of capital input (x_1), and thus for the C2-stat.

In Figure A2, we plot the value of this statistic (on the y-axis) against its empirical cumulative distribution function across CFOs (on the x-axis). The top graph refers to the calculation of the statistic at the 95% confidence level. The null of coherence is rejected in all bootstrap repetitions for about 40% of CFOs, in none for about 15% of CFOs, and in strictly more than none and less than all repetitions for the remaining 45% of CFOs. Consistent with the results in Table A12, the proportion of CFOs for whom the null of coherence is rejected more than half of the times is approximately 55%.[27]

[27] The bottom graph refers to the calculation of the statistic at the 99% confidence level.

D. Further Figures and Tables

Figure A1: **Survey Questions of Firm Forecasts**

4. Relative to the previous 12 months, what will be your company's PERCENTAGE CHANGE during the next 12 months? (e.g., +3%, -2%, etc.) [Leave blank if not applicable]	
% Prices of your products	% Technology spending
% Overtime	% Earnings
% Advertising/Marketing spending	% Revenues
% Number of employees	% Inventory
% Productivity (output per hour worked)	% M&A activity
% Wages/Salaries	% Capital spending
% Health care costs	% Dividends

Figure A2: **Bootstrap of Coherence Test Statistic C2-stat**

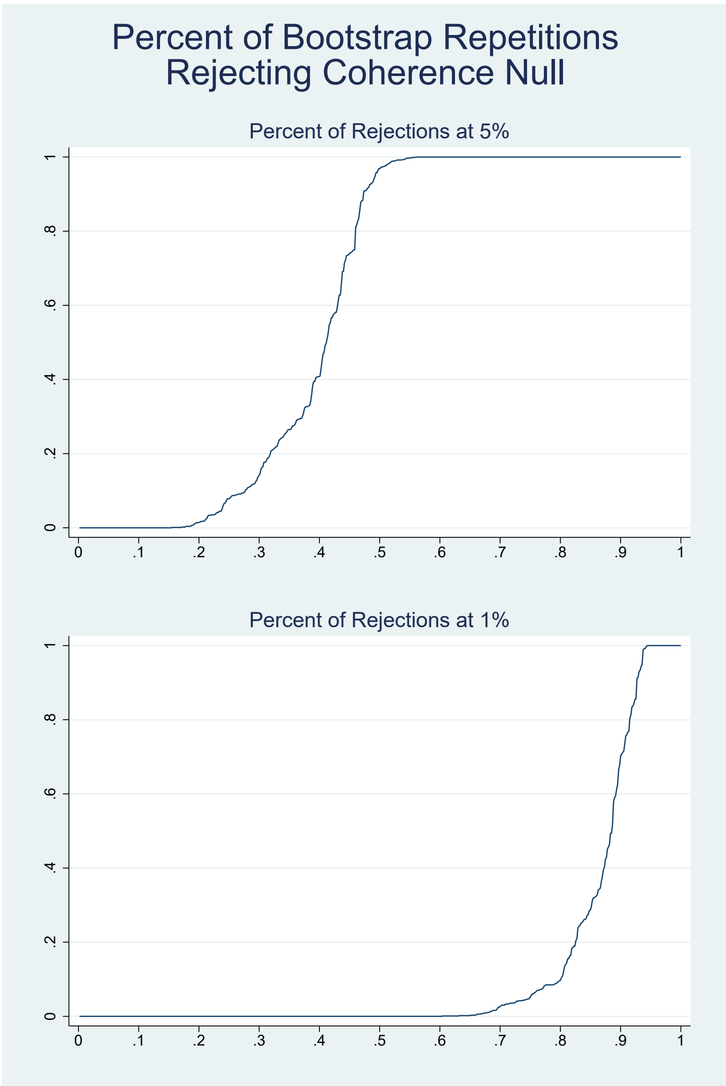


Table A1: **Summary Statistics**

The table reports summary statistics on selected firm characteristics. Panel A reports statistics in the matched Duke-Compustat sample. Panel B reports statistics in the Compustat universe over 1999-2019.

Panel A – Matched Duke-Compustat sample

	Mean	Std. Dev.	P05	Median	P95	N Obs.
Market-to-book	1.845	1.629	0.875	1.402	4.157	15,929
ROA	0.025	0.235	-0.154	0.037	0.167	16,591
Sales	10,617.34	28,482.18	53.66	2,043.96	49,545.00	17,799
Log(sales)	7.591	2.093	4.022	7.629	10.813	17,757
Assets	37,698.73	187,568.6	74.12	2,894.43	113,960.0	17,799
Log(assets)	7.993	2.221	4.306	7.971	11.644	17,799
Book Leverage	0.411	1.755	0.000	0.371	0.910	17,733
Capital Expenditure	0.045	0.058	0.001	0.031	0.134	16,200
R & D	0.060	0.202	0.000	0.025	0.211	9,043
Cash Flow	0.302	13.075	-1.127	0.413	3.005	17,010
Cash	4.895	62.328	0.014	0.563	15.188	17,269
Advertising	0.025	0.043	0.000	0.009	0.098	6,729
Dividends	0.117	1.420	0.000	0.060	0.349	17,391
Dividends (0/1)	0.607	0.488	0.000	1.000	1.000	17,391

Panel B – Compustat data

	Mean	Std. Dev.	P05	Median	P95	N Obs.
Market-to-book	1.843	2.526	0.708	1.297	4.489	105,769
ROA	0.006	0.279	-0.273	0.021	0.187	123,155
Sales	3,521.58	15,220.67	17.992	315.011	14,687.00	127,307
Log(sales)	5.911	2.062	2.890	5.753	9.595	127,307
Assets	12,626.69	98,056.55	24.147	597.555	30,241.99	140,894
Log(assets)	6.529	2.170	3.184	6.393	10.317	140,894
Book Leverage	0.408	22.120	0.000	0.362	0.996	139,264
Capital Expenditure	0.063	0.159	0.000	0.032	0.212	108,909
R & D	0.075	0.135	0.000	0.027	0.290	52,955
Cash Flow	0.050	0.261	-0.222	0.061	0.256	118,905
Cash	0.192	0.396	0.002	0.081	0.685	111,475
Advertising	0.034	0.103	0.000	0.009	0.131	39,813
Dividends	0.144	7.286	0.000	0.000	0.557	122,340
Dividends (0/1)	0.431	0.495	0.000	0.000	1.000	122,340

Table A2: **CFO Growth Forecasts and Realizations of Selected Balance Items**

The table reports summary statistics on selected forecasts and realizations. Panel A reports statistics on twelve-month ahead growth forecasts in the Duke sample. Panel B reports statistics on the corresponding realizations in the matched Duke-Compustat sample over 1999-2019.

Panel A – CFO Growth Forecasts (percent)

	Mean	Std. Dev.	Q10	Median	Q90	N Obs.
Expected Growth in Revenues and in Earnings						
Revenues	9.30	27.13	-5.00	5.00	20.00	14,490
Earnings	11.00	42.37	-10.00	5.00	30.00	25,472
Expected Growth in Capital-Related Expenditures						
Capital Expenditures	8.11	43.90	-15.00	3.00	25.00	25,305
R & D	4.51	21.65	0.00	0.00	15.00	8,325
Technology Spending	6.68	28.02	-5.00	3.00	20.00	22,404
Expected Growth in Labor-Related Costs						
Wages	3.90	12.41	0.00	3.00	7.00	27,472
Employees	3.95	30.16	-5.00	1.00	10.00	25,471
Outsourced Employees	3.74	21.19	0.00	0.00	10.00	10,990
Health Spending	8.59	11.65	1.00	8.00	15.00	25,064
Expected Growth in Productivity, Product Prices, and Advertising						
Productivity	3.91	9.38	0.00	3.00	10.00	18,197
Product Prices	2.08	8.22	-3.00	2.00	7.00	24,499
Advertising	4.75	21.83	-5.00	2.00	15.00	20,989
Expected Growth in Cash Holdings and Corporate Payout						
Cash	5.02	38.56	-20.00	0.00	20.00	16,876
Dividends	4.54	30.52	0.00	0.00	15.00	5,227
Share Repurchases	1.55	24.40	0.00	0.00	5.00	5,487

Panel B – Realizations, Matched Compustat-Duke Sample (percent)

	Mean	Std. Dev.	Q10	Median	Q90	N Obs.
Actual Growth in Revenues and in Earnings						
Revenues	6.80	21.32	-13.56	5.23	27.25	14,549
Earnings	-10.36	307.02	-161.71	2.36	124.59	14,580
Actual Growth in Capital-Related Expenditures						
Capital Expenditures	15.87	67.07	-42.59	3.96	75.70	13,770
R & D	7.09	29.57	-19.85	4.27	33.33	6,456
Technology Spending	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Actual Growth in Labor-Related Costs						
Wages	7.02	14.65	-7.23	5.35	22.10	2,836
Employees	2.98	16.95	-11.88	1.19	17.71	14,359
Outsourced Employees	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Health Spending	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Actual Growth in Productivity, Product Prices, and Advertising						
Productivity	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Product Prices	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Advertising	8.03	42.14	-26.76	2.79	38.46	5,735
Actual Growth in Cash Holdings and Corporate Payout						
Cash	35.42	132.66	-46.26	5.76	113.78	14,520
Dividends	12.68	52.88	-12.22	6.15	38.44	8,762
Share Repurchases	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

Table A3: **Growth Realizations of Selected Balance Items**

The table reports summary statistics on selected balance sheet items in the Compustat universe over 1999-2019.

Realizations in Compustat (percent)

	Mean	Std. Dev.	Q10	Median	Q90	N Obs.
Growth in Revenues and in Earnings						
Revenues	13.38	35.96	-16.64	6.89	46.24	105,866
Earnings	-16.21	432.06	-207.66	-3.92	174.42	105,841
Growth in Capital-Related Expenditures						
Capital Expenditures	35.71	132.99	-56.41	5.26	129.28	100,633
R & D	15.91	53.09	-25.24	6.75	57.64	40,715
Technology Spending	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Growth in Labor-Related Costs						
Wages	10.57	24.22	-9.57	6.94	31.96	29,491
Employees	6.18	25.30	-14.29	2.07	29.17	107,435
Outsourced Employees	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Health Spending	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Growth in Productivity, Product Prices, and Advertising						
Productivity	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Product Prices	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Advertising	19.12	80.07	-37.35	4.39	71.33	34,251
Growth in Cash Holdings and Corporate Payout						
Cash	76.55	308.36	-57.50	5.23	184.62	103,833
Dividends	18.74	97.74	-56.43	5.42	60.56	54,841
Share Repurchases	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

Table A4: **Minimum Distance of Sales Forecasts from Rules of Thumb: Robustness to Alternative Definition of Rule 5**

The top part of the table reports selected statistics on the orthogonal distance between the CFO forecast of sales growth and the forecast of sales growth implied by the five rules of thumb. Here we define R5 using advertising growth instead of earnings growth. The first column reports distance statistics for the CFOs whose forecast is closest to the forecast implied by R1. The second column reports distance statistics for the CFOs whose forecast is closest to the forecast implied by R2. The third column reports distance statistics for the CFOs whose forecast is closest to the forecast implied by R3. The fourth column reports distance statistics for the CFOs whose forecast is closest to the forecast implied by R4. The fifth column reports statistics for the CFOs whose forecast is closest to the forecast implied by R5. The last column reports statistics for the full sample, whereby each CFO's distance is the minimum distance between that CFO's forecast and the forecasts implied by the five rules of thumb. The bottom part of the table reports the absolute and relative frequencies of CFOs whose forecast of sales growth is closest to the corresponding rule of thumb by column.

	R1	R2	R3	R4	R5	All
Mean	0.040	0.025	0.030	0.041	0.023	0.029
Std. Dev.	0.031	0.044	0.031	0.056	0.016	0.039
Frac. Zeros	0.000	0.365	0.000	0.000	0.000	0.146
P10	0.005	0.000	0.006	0.002	0.004	0.000
P25	0.017	0.000	0.009	0.003	0.009	0.005
P50	0.031	0.007	0.022	0.023	0.019	0.016
P75	0.064	0.032	0.040	0.065	0.031	0.035
P90	0.094	0.071	0.074	0.122	0.048	0.071
P95	0.094	0.106	0.094	0.222	0.049	0.106
N of Observations	9	52	30	18	21	130
Fraction	0.069	0.400	0.231	0.138	0.162	1.000

Table A5: **Minimum Distance of Earnings Forecasts from Rules of Thumb**

The top part of the table reports selected statistics on the orthogonal distance between the CFO forecast of earnings growth and the forecast of earnings growth implied by the five rules of thumb. The first column reports distance statistics for the CFOs whose forecast is closest to the forecast implied by R1. The second column reports distance statistics for the CFOs whose forecast is closest to the forecast implied by R2. The third column reports distance statistics for the CFOs whose forecast is closest to the forecast implied by R3. The fourth column reports distance statistics for the CFOs whose forecast is closest to the forecast implied by R4. The fifth column reports statistics for the CFOs whose forecast is closest to the forecast implied by R5. The last column reports statistics for the full sample, whereby each CFO's distance is the minimum distance between that CFO's forecast and the forecasts implied by the five rules of thumb. The bottom part of the table reports the absolute and relative frequencies of CFOs whose forecast of earnings growth is closest to the corresponding rule of thumb by column.

	R1	R2	R3	R4	R5	All
Mean	0.045	0.021	0.028	0.032	0.031	0.026
Std. Dev.	0.054	0.028	0.046	0.035	0.030	0.033
Frac. Zeros	0.000	0.356	0.000	0.000	0.000	0.197
P10	0.002	0.000	0.002	0.005	0.004	0.000
P25	0.003	0.000	0.008	0.007	0.009	0.004
P50	0.027	0.014	0.010	0.017	0.015	0.014
P75	0.064	0.035	0.017	0.050	0.052	0.035
P90	0.101	0.057	0.099	0.071	0.073	0.068
P95	0.177	0.085	0.182	0.111	0.090	0.101
N of Observations	24	219	35	48	70	396
Fraction	0.061	0.553	0.088	0.121	0.177	1.000

Table A6: **Minimum Distance of CapEx Forecasts from Rules of Thumb**

The top part of the table reports selected statistics on the orthogonal distance between the CFO forecast of capex growth and the forecast of capex growth implied by the five rules of thumb. The first column reports distance statistics for the CFOs whose forecast is closest to the forecast implied by R1. The second column reports distance statistics for the CFOs whose forecast is closest to the forecast implied by R2. The third column reports distance statistics for the CFOs whose forecast is closest to the forecast implied by R3. The fourth column reports distance statistics for the CFOs whose forecast is closest to the forecast implied by R4. The fifth column reports statistics for the CFOs whose forecast is closest to the forecast implied by R5. The last column reports statistics for the full sample, whereby each CFO's distance is the minimum distance between that CFO's forecast and the forecasts implied by the five rules of thumb. The bottom part of the table reports the absolute and relative frequencies of CFOs whose forecast of capex growth is closest to the corresponding rule of thumb by column.

	R1	R2	R3	R4	R5	All
Mean	0.079	0.054	0.161	0.079	0.104	0.064
Std. Dev.	0.150	0.072	0.189	0.121	0.083	0.099
Frac. Zeros	0.000	0.146	0.000	0.000	0.000	0.106
P10	0.009	0.000	0.006	0.005	0.012	0.000
P25	0.022	0.014	0.008	0.029	0.039	0.014
P50	0.046	0.035	0.089	0.039	0.090	0.035
P75	0.076	0.071	0.321	0.062	0.157	0.071
P90	0.144	0.127	0.488	0.357	0.173	0.141
P95	0.264	0.163	0.508	0.390	0.293	0.220
N of Observations	58	288	14	25	11	396
Fraction	0.147	0.727	0.035	0.063	0.028	1.000

Table A7: **Additional Cross-Sectional Regressions in Compustat Data for Rules of Thumb**

The table reports results from OLS estimations using the Compustat universe over 1999-2019. The dependent variable is Capex Growth, the annual percent growth rate of capital expenditures. The regressors are Lag Capex Growth, the one-year lagged annual percent growth rate of capital expenditures; Past Capex Growth, the average of the past two lagged annual percent growth rates of capital expenditures; Sales Growth, the annual percent growth rate of total revenues; Wages Growth, the annual percent growth rate of total wages; Earnings Growth, the annual percent growth rate of net income; and Advertising Growth, the annual percent growth rate of advertising expenditures. All variables are winsorized at the 1st and 99th percentiles. Standard errors are clustered at the firm-year level and t-statistics are reported in parentheses. ***, **, and * indicate two-tailed significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Lag CapEx Growth	-0.060***									
	(0.010)									
Past CapEx Growth		-0.089***								-0.089***
		(0.016)								(0.040)
Sales Growth			1.055***				1.042***	0.996***	0.652***	0.702***
			(0.036)				(0.036)	(0.045)	(0.060)	(0.178)
Wages Growth				1.036***					0.589***	0.380***
				(0.082)					(0.081)	(0.159)
Earnings Growth					0.053***		0.018***			-0.034
					(0.005)		(0.004)			(0.024)
Advertising Growth						0.317***		0.178***		0.263***
						(0.021)		(0.016)		(0.076)
Constant	0.327***	0.316***	0.217***	0.215***	0.355***	0.274***	0.217***	0.178***	0.192***	0.166***
	(0.042)	(0.041)	(0.025)	(0.026)	(0.039)	(0.034)	(0.025)	(0.027)	(0.023)	(0.034)
R^2	0.004	0.004	0.081	0.047	0.005	0.042	0.082	0.096	0.063	0.074
N observations	85,838	74,413	100,441	15,955	100,040	33,202	100,040	33,202	15,955	3,013

Table A8: **Incoherence and Rules of Thumb: Alternative Definition of R5**
The top part of the table reports results from OLS estimations. The dependent variable is our measure of incoherence, the minimum orthogonal distance between the CFO forecast of sales growth and the forecast of sales growth implied by Rule 5. Here we define Rule 5 using advertising growth instead of earnings growth. The regressors are indicator variables. Rule 1 ('narrow bracketing') equals one if the CFO's sales growth forecast is closest to the forecast implied by R1. Rule 2 ('sales anchoring') equals one if the CFO's sales growth forecast is closest to the forecast implied by R2. Rule 3 ('economies of scale') equals one if the CFO's sales growth forecast is closest to the forecast implied by R3. Rule 4 ('industry based') equals one if the CFO's sales growth forecast is closest to the forecast implied by R4. Rule 5 is the excluded category. t-statistics are reported in parentheses. ***, **, and * indicate two-tailed significance at the 1%, 5%, and 10% level, respectively. The bottom part of the table reports selected summary statistics of the dependent variable.

	(1)	(2)	(3)	(4)	(5)
Rule 1 ('narrow bracketing')	0.099***				0.134***
	(0.022)				(0.025)
Rule 2 ('sales anchoring')		0.018			0.053***
		(0.012)			(0.016)
Rule 3 ('economies of scale')			-0.024*		0.023
			(0.014)		(0.018)
Rule 4 ('industry based')				0.003	0.044**
				(0.018)	(0.020)
Constant	0.058***	0.058***	0.071***	0.065***	0.023*
	(0.006)	(0.008)	(0.007)	(0.007)	(0.014)
R^2	0.126	0.009	0.014	-0.008	0.175
N observations	130	130	130	130	130
Summary Statistics of the dependent variable					
Mean		0.065	P10		0.012
Std. Dev.		0.069	Median		0.045
			P90		0.153

Table A9: **Incoherence and Rules of Thumb: Distance from Optimal Forecast, Robustness to (R6)**

The dependent variable is our measure of incoherence, the minimum orthogonal distance between the CFO forecast of sales growth and the forecast of sales growth implied by Rule 5. The regressors are indicator variables. Rule 1 ('narrow bracketing') equals one if the CFO's sales growth forecast is closest to the forecast implied by R1, and within a $+/-$ 0.05 interval around it. Rule 2 ('sales anchoring') equals one if the CFO's sales growth forecast is closest to the forecast implied by R2, and within a $+/-$ 0.05 interval around it. Rule 3 ('economies of scale') equals one if the CFO's sales growth forecast is closest to the forecast implied by R3, and within a $+/-$ 0.05 interval around it. Rule 4 ('industry based') equals one if the CFO's sales growth forecast is closest to the forecast implied by R4, and within a $+/-$ 0.05 interval around it. Rule 6 ('residual') equals one if the CFO's sales growth forecast has not been classified as either Rule 1, Rule 2, Rule 3, or Rule 4. Rule 5 is the excluded category. t-statistics are reported in parentheses. ***, **, and * indicate two-tailed significance at the 1%, 5%, and 10% level, respectively. The bottom part of the table reports selected summary statistics of the dependent variable.

	(1)	(2)	(3)	(4)	(5)	(6)
Rule 1 ('narrow bracketing')	0.037**					0.088***
	(0.018)					(0.017)
Rule 2 ('sales anchoring')		0.010				0.060***
		(0.008)				(0.011)
Rule 3 ('economies of scale')			-0.058***			0.001
			(0.013)			(0.014)
Rule 4 ('industry based')				-0.043***		0.019
				(0.009)		(0.012)
Rule 6 ('residual')					0.097***	0.132***
					(0.009)	(0.012)
Constant	0.071***	0.069***	0.079***	0.082***	0.054***	0.020
	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.009)
N observations	396	396	396	396	396	396

Table A10: **Incoherence, Rules of Thumb, and Corporate Performance: Robustness Using Alternative Definition of Rule 5**

The table reports results from OLS estimations. The dependent variable is return on assets, operating income divided by beginning-of-period total assets. The main regressor is our measure of incoherence, the minimum orthogonal distance between the CFO forecast of sales growth and the forecast of sales growth implied by Rule 5. Here we define Rule 5 using advertising growth instead of earnings growth. Miscalibration is the negative imputed individual volatility, computed as the mean of the imputed individual volatility within each survey date subtracted from the imputed individual volatility; the result is then multiplied by -1 and divided the standard deviation of the imputed individual volatilities within the survey date. Optimism is a standardized expected return measure, computed by subtracting the mean within each survey date from the expected return estimate and dividing by the standard deviation within the survey date. Both miscalibration and optimism apply the same procedure to Short-Term (one-year) and Long-Term (10-year) forecasts. Firm size is the log of total assets. Market-to-book is the book value of assets plus the market value of common equity minus the book value of common equity and deferred taxes divided by total assets. Dividends is an indicator equal to one if common dividend is strictly larger than zero. One-digit SIC industry fixed effects and survey fixed effects are included but their coefficients are not reported. Standard errors are clustered by firm and bootstrapped following Cameron, Gelbach, and Miller (2009). t-statistics are reported in parentheses. ***, **, and * indicate two-tailed significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Incoherence	-0.279*	-0.335**	-0.320**	-0.317**			
	(0.156)	(0.151)	(0.159)	(0.152)			
Rule 1 (‘narrow bracketing’)					-0.062**	-0.071*	-0.071*
					(0.031)	(0.042)	(0.042)
Rule 2 (‘sales anchoring’)					0.009	0.012	0.014
					(0.032)	(0.035)	(0.035)
Rule 3 (‘economies of scale’)					0.018	0.023	0.020
					(0.037)	(0.037)	(0.038)
Rule 4 (‘industry based’)					0.020	0.020	0.026
					(0.033)	(0.039)	(0.036)
Miscalibration ST		0.009				0.004	
		(0.005)				(0.007)	
Optimism ST		0.011				0.009	
		(0.011)				(0.011)	
Miscalibration LT			0.014				0.012
			(0.010)				(0.011)
Optimism LT			0.008				0.007
			(0.011)				(0.011)
Constant	0.051***	0.051***	0.048***	0.014	0.024	0.015	0.015
	(0.011)	(0.011)	(0.012)	(0.013)	(0.031)	(0.034)	(0.034)
Industry FE	No	No	No	Yes	No	No	No
Survey FE	No	No	No	Yes	No	No	No
R^2	0.055	0.075	0.073	0.140	0.017	0.020	0.026
N of Observations	136	122	121	136	136	122	121

Table A11: **Violations of Coherence Inequality Restrictions**

The table presents results of inequality tests of the null hypothesis of coherence at the individual CFO level based on ex ante forecasts. Panel A reports the percentage of CFOs for which coherence is rejected. The results are reported for the coherence inequality in levels and in growth rates, separately, and for three different values of χ, which denotes the elasticity of substitution between capital and labor. Panel B reports summary statistics of the difference between the left-hand-side and the right-hand-side of the coherence inequality.

Panel A – Inequality Test of Coherence

	$\chi = 0.5$	$\chi = 0.7$	$\chi = 0.9$
Inequality in Levels			
% Incoherent	100.00	100.00	99.07
% Coherent	0.00	0.00	0.93
% Total	100.00	100.00	100.00
N. Obs.	107	107	107
Inequality in Growth Rates			
% Incoherent	73.31	73.14	72.96
% Coherent	26.69	26.86	27.04
% Total	100.00	100.00	100.00
N. Obs.	577	577	577

Panel B – Summary stats of difference, LHS − RHS

	$\chi = 0.5$	$\chi = 0.7$	$\chi = 0.9$
Inequality in Levels			
Mean	15,419.59	15,201.16	15,033.78
Std. Dev.	28,574.83	28,233.56	27,990.89
Q10	213.3352	207.60	194.3495
Median	3,252.30	3228.92	3,205.96
Q90	39,737.28	39,505.75	39,360.82
N. Observations	107	107	107
Inequality in Growth Rates			
Mean	0.047	0.044	0.042
Std. Dev.	0.122	0.125	0.128
Q10	-0.061	-0.067	-0.067
Median	0.034	0.033	0.033
Q90	0.164	0.163	0.162
N. Observations	577	577	577

Table A12: **The Coherence and Accuracy Sides of Rationality**

The table presents results of tests of the null hypotheses of coherence and of accuracy at the individual CFO level based on ex post forecast errors. We consider four hypotheses, the coherence hypothesis based on the C2-stat, the accuracy hypothesis based on the forecast error of capex growth, the accuracy hypothesis based on the forecast error of sales growth, and both accuracy hypotheses at the same time. Forecast errors are computed as realization minus forecast. Under the null hypothesis the above statistics equal zero. Critical values are those of the t-student distribution with one degree of freedom, namely $+/-12.706$ at the 5% and $+/-63.657$ at the 1%. Panel A reports the percentage of CFOs for which each hypothesis is separately rejected. Panel B reports the percentage of CFOs for which the null hypotheses of coherence and accuracy are jointly not rejected, for which coherence is not rejected but accuracy is rejected; for which coherence is rejected but accuracy is not rejected; and for which both coherence and accuracy are rejected. In Panel B, accuracy means both forecasts of capex growth and sales growth are accurate; and inaccuracy means at least one forecast is inaccurate. Panel C reports summary statistics of the C2-stat, the forecast error of sales growth, and the forecast error of capex growth.

Panel A – Separate Assessment of Coherence and Accuracy (Percent of Rejections of Null)

Significance level α	Coherence Sales-CapEx	Accuracy Sales	Accuracy CapEx	Accuracy Both
	(1)	(2)	(3)	(4)
5%	55.7%	27.2%	47.9%	57.0%
1%	7.7%	1.8%	6.4%	7.1%

Panel B – Joint Assessment of Coherence and Accuracy

Significance level α	Coherent + Accurate	Coherent + Inaccurate	Incoherent + Accurate	Incoherent + Inaccurate
	(1)	(2)	(3)	(4)
5%	31.1%	13.2%	12.0%	43.7%
1%	89.4%	2.9%	3.4%	4.3%

Panel C – Test Statistics: Summary Statistics

	Mean	Std.Dev.	P05	Median	P95	N Obs.
C2-statistic	-0.193	4.846	-8.335	-0.135	7.871	560
FE Sales	-0.538	19.07	-23.53	0.554	22.24	563
FE CapEx	-0.988	31.28	-54.18	1.186	41.20	560

Electronic Access to the Working Paper Series

The full set of ECGI working papers can be accessed through the Institute's Web-site (www.ecgi.global/content/working-papers) or SSRN:

Finance Paper Series	http://www.ssrn.com/link/ECGI-Fin.html
Law Paper Series	http://www.ssrn.com/link/ECGI-Law.html