

R&D Expenditure and Revenue-Based Productivity: Measurement Sensitivity in a Pilot of Shenzhen-Listed Firms

Jiao Di*

Shenyang Chongwen High School, Shenyang, China

**Corresponding author: Jiao Di.*

Abstract

The resource-based view treats R&D expenditure as an input into firm-specific knowledge. Its empirical link with productivity, however, depends on how the two constructs are measured. This pilot uses 60 full annual reports disclosed through CNINFO for 12 Shenzhen manufacturing and technology issuers during 2019–2023. After the pre-specified completeness and history rule, the analytic panel contains 48 firm-years from 10 firms. The author estimates a revenue-based productivity residual and compare pooled, fixed-effect, lagged, alternative-outcome, denominator, tail, resampling, temporal, and first-difference specifications. Revenue-share R&D is negative in the firm and year fixed-effect model, whereas R&D scaled by fixed assets is imprecise. A 1,999-resample full-chain firm bootstrap also includes zero. The evidence does not identify a causal R&D effect or describe all Chinese listed firms. It instead shows why an auditable reporting pipeline and denominator checks are needed before resource-based arguments are turned into productivity claims.

Keywords

R&D expenditure, revenue-based productivity, measurement sensitivity, resource-based view, Shenzhen-listed firms

1. Introduction

Firm R&D spending can add to knowledge resources that competitors may not readily reproduce. RBV is concerned with heterogeneous, hard-to-trade knowledge assets and with the organizational conditions that turn them into performance [1, 2]. Reported R&D is a visible indicator of knowledge-building effort, but its economic meaning is not straightforward. Spending may come before commercialization, accompany restructuring costs, or be recorded differently across reports. In an observational panel, current expenditure and a measured productivity gain therefore need not appear at the same time.

The Chinese context intensifies this issue without getting rid of it. New research on listed companies indicates that productivity has to do with ownership, knowledge-based capital, and policy exposure, but R&D is just one of the inputs in a more comprehensive package of resources [3-5]. A representative database and a structural productivity design would have been required to calculate such mechanisms. The paper cannot

replace the small disclosed-report panel to perform this. It rather makes use of the pilot to inquire if an intuitively obvious R&D-productivity relationship can withstand a clear shift in the denominator of R&D intensity and a series of predetermined sensitivity tests. The smaller question is achievable using publicly available documents and is meaningful to scholars who need to choose which of the reported ratios are really indicative.

This paper asks a deliberately limited question: what does a transparent public-report pilot show about the sensitivity of the R&D-productivity association to operationalization? Instead of claiming a market-wide causal effect, the author builds a reproducible CNINFO-based panel and sets out a small specification ladder before interpreting an estimate. The paper links source URLs, identifiers, document hashes, extraction anchors, and exclusions for each report. It then compares revenue-share R&D with R&D-to-fixed-assets, which does not share the revenue denominator with the outcome. Negative, null, and imprecise estimates are reported together. Their disagreement is treated as measurement evidence rather than as a reason to select a favorable result.

The rest of the paper is organized as follows. Section 2 reviews the relevant literature. Section 3 describes the annual-report panel and the empirical specifications. Section 4 presents the pre-specified estimates and sensitivity checks. Section 5 interprets the results through RBV while considering measurement alternatives. Section 6 concludes the paper.

2. Related Work

The difference between a resource input and a realized productivity has existed in the fields of strategy and economics for quite some time. The account of firm growth given by Penrose focuses on the productive service generated by the resources in use instead of merely a resource label [6]. New China-specific literature also views R&D as a component of a broader innovation process: it can influence productivity as a form of product and process innovation and its relationship can be different under various conditions of ownership and organizations [7-9]. These mechanisms allow contemporaneous accounting ratios to be useful empirical signals, but do not directly measure an intangible capability. Small panels have the ability to capture R&D prior to commercialization or during changes in the product mix.

Annual report of Chinese listed companies allows observing the R&D, revenue, fixed assets and employment at the issuer-year level, but also reveals a measurement issue. When productivity is estimated in terms of revenue and R&D is divided by revenue, the regressor and outcome have the same denominator. A significant coefficient could therefore be a result of construction as opposed to a transformation of resources. The literature on productivity has long cautioned that production-function estimates are prone to the careful consideration of inputs and unobserved heterogeneity [10, 11]. Since the published reports have not offered a consistent set of intermediate-input or investment information, the current work does not aim at identifying structure TFP, but tests an auditable proxy and its sensitivity to the denominator. The new research on Chinese listed companies also notes different TFP trends under various institutional and local settings [5, 12].

3. Data and Method

3.1 Auditable Annual-report Panel

The author compiled 60 full annual reports for 12 Shenzhen-listed manufacturing and technology firms over fiscal years 2019–2023. The reports came from CNINFO's historical-announcement interface, China's designated information-disclosure platform [13]. For each document, the extraction pipeline retained the announcement identifier, disclosure URL, local file name, and SHA-256 hash. Report-table anchors provided revenue, reported R&D expenditure (the R&D-expense field), fixed assets, and employee counts. A firm-year entered the analytic sample only if all four anchors were present; no field was manually completed or imputed. Fifty-one firm-years met this rule, and the pre-specified requirement of at least three observations per firm left 48 firm-years from 10 firms.

This screening rule separates document traceability from econometric adequacy. The four anchors are needed to calculate the pre-specified ratios and the revenue-based proxy, while the minimum-history rule is needed for within-firm comparisons. The rule does not make the retained observations random. Excluded

reports were kept in the source manifest, so the sample flow can be checked rather than inferred from a final regression table. Values are fiscal-year accounting measures. The analysis does not rebuild a perpetual-inventory R&D stock or harmonize unreported project-level information.

The design is a pilot, not a representative census of Chinese listed firms. It was chosen because every observation can be traced to a public primary document. The analysis therefore gives priority to provenance and clear sensitivity checks over breadth. This is consistent with RBV's concern with context-specific resources, without claiming that the observed firms represent all sectors, exchanges, or ownership forms.

3.2 Variables and Specifications

Let Y_{it} denote revenue, K_{it} fixed assets, L_{it} employees, and RD_{it} reported R&D expenditure for firm i in year t . The revenue-based productivity proxy is the residual $\hat{\omega}_{it}$ from

$$\log Y_{it} = \beta_K \log K_{it} + \beta_L \log L_{it} + \tau_t + u_{it}. \quad (1)$$

The author treats this residual as a revenue-based productivity proxy, not as a structural measure of physical total factor productivity. The revenue-scaled intensity is $RD I_{it}^{Rev} = RD_{it}/Y_{it} \times 100$, and the denominator-sensitive alternative is $RD I_{it}^{FA} = RD_{it}/K_{it} \times 100$. The association model is

$$\hat{\omega}_{it} = \alpha + \gamma RD I_{it} + \mu_i + \tau_t + \epsilon_{it}, \quad (2)$$

where $RD I_{it}$ denotes either stated intensity, μ_i and τ_t are firm and year effects, and ϵ_{it} is the remaining error. Firm-clustered covariance estimates are reported for fixed-effect models and HC3 estimates for pooled models.

All monetary anchors were normalized to RMB 10,000 before ratios and logarithms were formed. The author used natural logarithms and retained only observations with non-missing R&D, revenue, fixed assets, and employment, with strictly positive revenue, fixed assets, and employment; R&D was allowed to equal zero. No missing value was imputed. The residual model is OLS with year indicators and HC3 covariance; the association models use OLS with issuer and year indicators, with issuer-clustered covariance where fixed effects are included. Thus, the reported adjusted R^2 values describe fitted regressions rather than a separate production-function estimator.

The replication routine reads the frozen observation file, chooses report rows that are marked as ready, applies the complete-anchor and minimum-history principles, and computes ratios, logs, residuals, and the one-year lag after sorting issuer-years. It creates the pre-specified estimate table, analytic panel, and sensitivity results of the same retained sample. A reproducibility companion record associates every pre-specified executable specification with its paper label, its presence of an R&D term, its sample and covariance option, so that a narrative label does not take the place of a unique regression.

The residual structure is intentionally low-key. This eliminates the fitted correlation between revenue and two observed factors of production and common year effects; this does not recapture physical output, markups, materials, capital utilization or a structural productivity process. Traditional methods like Olley-Pakes and Levinsohn-Petrin deal with unobserved productivity with more detailed information about inputs and investments [14, 15]. The selected reports do not have those data regularly. The author hence refers to the term using the phrase revenue-based productivity proxy in all cases and considers how an inference would change if R&D were measured on a scale of fixed assets instead of revenue. The other denominator is not a cure to endogeneity; however, it removes the direct shared-revenue part in the central comparison.

The final specification ladder includes pooled estimation with log-input and year fixed effects, fixed effects on both firm-years, 1-year R&D lag, log revenue per employee as an alternate outcome, and the denominator of fixed assets. No-R&D fixed-effect model was calculated solely as an outcome comparison and it does not contain coefficient of R&D. Further, the author had pre-specified five symmetric winsorization levels, leave-one-firm-out estimates, six overlapping temporal windows, and first-difference diagnostics. To compensate the few numbers of firms, the author also conducted 1,999 firm-cluster resamples (seed 20260823) and reestimated both the residual equation and the association model on each sample. Such a full chain distribution is a sensitivity diagnostic rather than an alternative to proper causal inference or a definitive small cluster inference.

The ladder is supposed to enable putting the interpretation into falsifiability in the pilot. A constant positive finding should be directionally consistent after eliminating the common denominator, altering the time slice, and excluding the individual companies. On the other hand, the negative revenue-share coefficient should not be considered as a RBV contradiction because it vanishes with this measure change. The protocol prevents the statement to go beyond the observed relationship in that it records the reason why more causal language would involve larger panels, lag orders, and exogenous variation.

4. Results

The 48 retained firm-years vary substantially in revenue, fixed assets, employment, and reported R&D expenditure (Table 1). Mean revenue-scaled R&D expenditure was 7.883%, while the fixed-assets-scaled ratio averaged 44.916%. Figure 1 shows the raw relation between revenue-share R&D and the productivity residual. It is descriptive rather than an effect estimate because firm-specific factors and the common revenue component have not been removed.

Table 1: Summary statistics for the analytic panel. Monetary values are RMB 10,000; ratio values are percentages

Variable	N	Mean	SD	Min.	Median	Max.
Revenue	48	12,887,111	13,882,278	641,164	7,580,516	60,231,535
Fixed assets	48	5,223,750	7,155,910	127,875	1,851,950	23,090,382
Employees	48	91,219	133,873	9,104	45,578	703,504
R&D expenditure	48	708,564	676,420	45,866	555,659	3,957,495
R&D / revenue (%)	48	7.883	3.852	3.464	6.553	17.716
R&D / fixed assets (%)	48	44.916	38.836	3.390	32.160	123.240
Productivity residual	48	0.000	0.349	-0.683	0.028	0.670

Figure 1: Descriptive pattern between R&D intensity and the revenue-based productivity residual

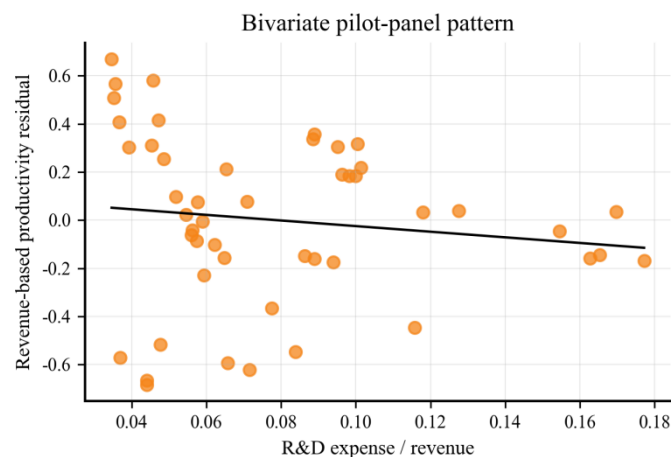


Table 2 reports the pre-specified estimates. The pooled R&D/revenue coefficient was -0.0278 per percentage point (95% CI $[-0.0663, 0.0107]$). With firm and year effects, the coefficient was -0.1215 (95% CI $[-0.1710, -0.0719]$). The one-year lag estimate, based on 38 observations, was -0.0771 and its interval included zero. Replacing the residual with log revenue per employee yielded -0.1207 , but this outcome remains tied to revenue. These estimates are not read as evidence that R&D reduces productivity because the principal intensity and the principal outcome share revenue.

Table 2: Pre-specified association estimates. Ratio coefficients are per percentage point

Specification	Estimate	95% CI	N	Firms	Effects
Pooled R&D / revenue	-0.0278	$[-0.0663, 0.0107]$	48	10	Year
FE, R&D / revenue	-0.1215	$[-0.1710, -0.0719]$	48	10	Firm + year
Lagged R&D / revenue, FE	-0.0771	$[-0.1780, 0.0237]$	38	10	Firm + year
Labor-productivity outcome, FE	-0.1207	$[-0.1736, -0.0678]$	48	10	Firm + year
FE, R&D / fixed assets	0.0020	$[-0.0030, 0.0070]$	48	10	Firm + year
First difference, R&D / fixed assets	0.0034	$[0.0014, 0.0054]$	38	10	Year

Pooled OLS includes log fixed assets, log employees, and year effects with HC3 covariance. Fixed-effect and first-difference rows use firm-clustered covariance (10 clusters). The no-R&D fixed-effect benchmark is used only for outcome fit and is not listed because it has no R&D coefficient. Conventional adjusted R^2 is not reported as a within-fit statistic, and fit values are not comparable across the alternative outcome or first-difference model.

The denominator change is associated with a different pattern. For R&D divided by fixed assets, the firm/year fixed-effect estimate was 0.0020 per percentage point with a 95% interval of $[-0.0030, 0.0070]$. The point estimate was positive but imprecise. Table 2 shows this contrast across the pre-specified ladder. The comparison indicates that the association is sensitive to the R&D denominator and cannot support a single directional interpretation of the pilot panel.

Sensitivity checks did not convert the capital-scaled estimate into a precise effect. The 2019–2021 and 2021–2023 windows produced estimates of 0.0038 (95% CI $[-0.0050, 0.0127]$) and 0.0022 (95% CI $[-0.0056, 0.0099]$), respectively. The 1,999-resample full-chain firm bootstrap percentile interval was $[-0.0077, 0.0073]$, with a median estimate of 0.0018. Thus, resampling firms did not turn the small positive point estimate into a consistently positive result. Figure 2 summarizes the six temporal windows, and Figure 3 displays the bootstrap coefficient distribution.

Figure 2: Temporal-window sensitivity for R&D divided by fixed assets. All confidence intervals include zero

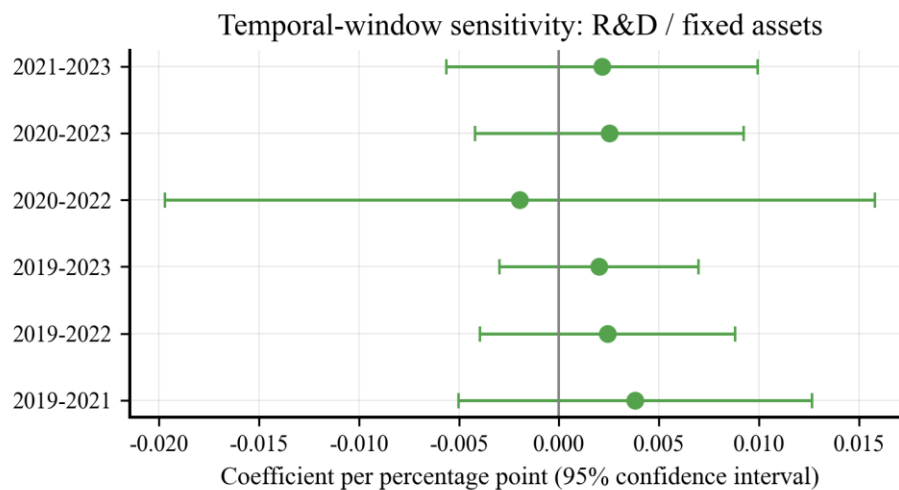
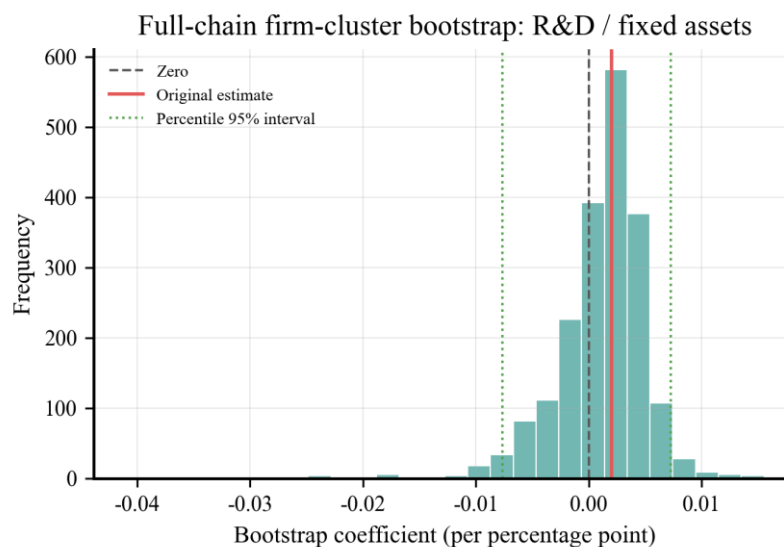


Figure 3: Full-chain firm-cluster bootstrap distribution for the R&D-to-fixed-assets fixed-effect estimate, based on $B=1,999$ bootstrap resamples. Each draw re-estimates the revenue-residual and association equations. The red line is the original estimate, the dashed line is zero, and dotted lines mark the percentile 95% interval



Tail trimming and leave-one-firm-out checks serve a different purpose from the confidence intervals. The factors ask whether a result is driven by a small number of unusually high-intensity observations or by a single issuer that exerts unusually strong influence in a short panel. Across the pre-specified symmetric winsorization grid, the revenue-share fixed-effect coefficient remained negative, from -0.0936 to -0.1345 , with all intervals below zero. This stability does not make the coefficient a resource effect because the shared revenue component remains in every version. In contrast, each leave-one-firm-out interval for the capital-scaled model crossed zero; point estimates ranged from -0.0012 to 0.0029 .

First differences provide a mixed diagnostic. Pooling all annual changes with year effects produced 0.0034 (95% CI $[0.0014, 0.0054]$) for the capital-scaled measure, whereas the 2021–2022 change was close to zero and imprecise. The first-difference estimate is reported in Table 2 rather than treated as a hidden favorable result. Its smaller effective sample, short period, and overlapping diagnostic windows do not distinguish a delayed innovation return from period-specific composition or reporting variation. It is therefore a record of within-firm changes, not confirmation of a causal pathway. In combination, the diagnostics favor a measurement-sensitive account and do not support a stable ranking of firms by innovation effectiveness within this pilot panel.

5. Discussion

The pilot does not advocate a straightforward assumption that greater reported R&D intensity will increase measured productivity. Its more acceptable contribution is diagnostic. According to RBV, knowledge resources have potential to generate advantage if integrated into supplementary organizational and business assets [2, 16]. The annual report records an input to this process in the form of R&D expenditure and does not record the outcome capacity, when commercialized or how appropriable it was. An association that is weak or unstable contemporaneously may also be consistent with RBV and with more recent findings that returns to innovation vary by ownership and organizational contexts [4, 9].

The RBV interpretation also clarifies what R&D can and cannot represent here. R&D is a visible input, but a resource-based explanation depends on the package around it: people, routines, complementary assets, and the capacity to turn technical knowledge into sales. The annual reports provide standardized spending, selected balance-sheet items, and employee counts, but reports do not measure those complements directly. The paper therefore does not test a mediation mechanism or claim that RBV is validated. RBV is used as an interpretive framework for why expenditure alone may be insufficient and why time and firm differences matter.

The negative revenue-share coefficients have a particularly important rival explanation. Revenue is used both in the intensity denominator and in the construction of the productivity proxy. A decline in revenue can mechanically raise RD/Y while reducing a revenue-based residual, producing a negative relation even if the underlying stock of knowledge is valuable. The R&D-to-fixed-assets estimate avoids that exact coupling, but it is imprecise in this small panel. The appropriate conclusion is neither that R&D is harmful nor that a positive effect has been recovered. Instead, the outcome shows that the interpretation is measurement-sensitive.

Several limitations set the boundary of the paper. The source frame is restricted to selected Shenzhen manufacturing and technology issuers, and the short 2019–2023 window offers limited power for firm-clustered inference. Revenue is not physical output, fixed assets omit intangible capital, and public-report extraction cannot recover all controls used in larger commercial databases. The design also lacks an exogenous R&D shock, so it cannot distinguish R&D causing productivity from productivity influencing R&D budgets. These limitations are reported because such limitations define the next empirical step: a broader verified panel, alternative output measures, and a credible source of R&D variation.

With a larger issuer panel, the analysis could add industry-by-year controls, longer lags, and explicit treatment of entry, exit, and missing disclosure. When production data separate quantities from prices and include intermediate inputs, a TFP interpretation would be more defensible. A causal design could also use a policy change, a funding-eligibility threshold, or another external source of R&D variation with credible exclusion restrictions. The current pilot does not provide those conditions. It instead shows the evidence gap between a reproducible association study and a policy recommendation.

For conference readers, the practical implication is modest. Public disclosure can support reproducible evidence when its provenance chain is explicit. It should not be used to manufacture a precise general

conclusion. Researchers who rely on R&D intensity should report denominator alternatives and clarify whether productivity is revenue-based, quantity-based, or estimated with richer production-function data.

6. Conclusion

This study built a traceable annual-report pilot to examine the association between reported R&D expenditure and a revenue-based productivity proxy among selected Shenzhen-listed firms. The pre-specified models gave negative revenue-share estimates but an inconclusive R&D-to-fixed-assets estimate. Temporal, full-chain bootstrap, first-difference, and leave-one-firm-out diagnostics did not establish a stable positive effect. The conclusion is therefore narrow: in a public-report pilot, the R&D–productivity relation depends materially on operationalization. Future research should use broader verified panels, non-revenue output measures, and designs that address the joint determination of innovation expenditure and productivity.

References

- [1] Wernerfelt, B., “A resource-based view of the firm,” *Strategic Management Journal*, 5, (2), 171–180, 1984.
- [2] Barney, J., “Firm resources and sustained competitive advantage,” *Journal of Management*, 17, (1), 99–120, 1991.
- [3] Wang, J., “Managing R&D in chinese listed firms: The interactive effects of ownership concentration and corporate governance,” *Thunderbird International Business Review*, 63, (5), 607–621, 2021, <https://doi.org/10.1002/tie.22207>.
- [4] He, Q., Guijarro-Garcia, M., and Costa-Climent, J., “Impact of knowledge-based capital on firm productivity: The contingent effect of ownership,” *Journal of Business Research*, 140, 85–94, 2022, <https://doi.org/10.1016/j.jbusres.2021.11.088>.
- [5] Ren, X., Zhang, X., Yan, C., and Gozgor, G., “Climate policy uncertainty and firm-level total factor productivity: Evidence from china,” *Energy Economics*, 113, 106209, 2022, <https://doi.org/10.1016/j.eneco.2022.106209>.
- [6] Penrose, E. T., *The Theory of the Growth of the Firm*, Basil Blackwell, Oxford, 1959.
- [7] Zhu, F., Li, Q., Yang, S., and Balezentis, T., “How ICT and R&D affect productivity? Firm-level evidence for china,” *Economic Research-Ekonomska Istraživanja*, 34, (1), 3468–3486, 2021, <https://doi.org/10.1080/1331677X.2021.1875861>.
- [8] Leung, T. Y., and Sharma, P., “Differences in the impact of R&D intensity and R&D internationalization on firm performance: Mediating role of innovation performance,” *Journal of Business Research*, 131, 81–91, 2021, <https://doi.org/10.1016/j.jbusres.2021.03.060>.
- [9] Anwar, S., and Sun, S., “R&D, product and process innovation, and firm performance: A case study of chinese firms,” *Arthaniti: Journal of Economic Theory and Practice*, 2024, <https://doi.org/10.1177/09767479241247615>.
- [10] Griliches, Z., *R&D and productivity: The econometric evidence*, University of Chicago Press, 1998.
- [11] Hall, B. H., and Mairesse, J., “Exploring the relationship between R&D and productivity in french manufacturing firms,” *Journal of Econometrics*, 65, (1), 263–293, 1995.
- [12] Xiong, J., and Chen, L., “Dialect diversity and total factor productivity: Evidence from chinese listed companies,” *Frontiers in Psychology*, 13, 1017397, 2022, <https://doi.org/10.3389/fpsyg.2022.1017397>.
- [13] CNINFO, *Historical announcements and listed company disclosure database*, <https://www.cninfo.com.cn/new/commonUrl?url=disclosure/list/notice>, accessed August 2026.
- [14] Olley, G. S., and Pakes, A., “The dynamics of productivity in the telecommunications equipment industry,” *Econometrica*, 64, (6), 1263–1297, 1996.

- [15] Levinsohn, J., and Petrin, A., “Estimating production functions using inputs to control for unobservables,” *Review of Economic Studies*, 70, (2), 317–341, 2003.
- [16] Teece, D. J., “Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy,” *Research Policy*, 15, (6), 285–305, 1986, [https://doi.org/10.1016/0048-7333\(86\)90027-2](https://doi.org/10.1016/0048-7333(86)90027-2).

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Open Access

This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

