Published by Yayasan Amerta Insan Unggul

DOI: 10.65246

p-ISSN 3063-9840 | e-ISSN 3063-8194

GoodWill Journal of Economics, Management, and Accounting

journal homepage: https://jurnal.amertainstitute.com/index.php/GoodWill/index



Transfer Pricing Dispute and Its Implications for Judicial Review in Indonesia

Syafri Maulana Akbar[1], Resi Ariyasa Qadri[2]
Polytechnic of State Finance STAN
Email: resi.ariyasa@pknstan.ac.id

ARTICLE INFO

Keywords:
Transfer Pricing Dispute;
Judicial Review;
Indonesia

Publication History:
Received: 15 May 2026
Screening: 25 May 2026
Review: 21 June 2026
Revision requested: 28 July 2026
Revised: 15 August 2026
Accepted: 29 October 2026

ABSTRACT

This study examines a transfer pricing dispute in Indonesia and its implications for evaluating Judicial Review, focusing on Tax Court Decision Number PUT-014532.15/2020/PP/M.XIIIB of 2022. Using a qualitative interpretive case-study approach, data were collected through document analysis and a semi-structured interview with a Tax Court Judge and analyzed using Thematic Analysis. The findings show that the dispute did not primarily concern the selection of the Transactional Net Margin Method (TNMM), but rather its application through functional characterization, period selection, comparable-company selection, comparability adjustments, and evidentiary support. The Tax Court accepted the Directorate General of Taxes' use of 2017 single-year data, four comparable companies, and a median operating margin of 7.49 percent, resulting in the full upholding of an IDR 76.21 billion adjustment. Conversely, the Taxpayer's explanation of underutilization and commercial losses was not sufficiently operationalized into a measurable capacity adjustment supported by adequate causal evidence. These findings indicate that the implications for Judicial Review depend on distinguishing potential legal issues from disputes that predominantly involve reassessment of facts and evidence. Issues concerning the traceability of judicial reasoning, the attribution of adjustments, and the basis of particular methodological decisions may warrant further legal evaluation only insofar as they can be linked to the statutory grounds for Judicial Review. The study demonstrates that transfer pricing disputes are shaped by the interaction of methodological choices, evidentiary quality, and judicial reasoning, which collectively determine the scope of further legal remedies.

1. Introduction

The rapid growth of related-party transactions among multinational manufacturing entities often creates a high degree of complexity in the application of the Arm's Length Principle (Permatasari & Husnasari, 2022; Tambunan, 2020; Udoayang et al., 2010). Information asymmetry and differences in methodological perspectives between Taxpayers and the Directorate General of Taxes frequently constitute the primary causes of transfer pricing disputes that proceed to the Tax Court (Dwindahany & Efendi, 2024; Tambunan & Abbas, 2023; Zulfiqar et al., 2023). This phenomenon of methodological disagreement is clearly reflected in Tax Court Decision Number PUT-014532.15/2020/PP/M.XIIIB of 2022, involving PT X as the Appellant. In this case, the Taxpayer argued that the operating loss incurred by the Power Generation Factory segment was solely attributable to commercial business factors, particularly a decline in sales volume below the break-even point, as demonstrated using a multiple-year data approach. In contrast, the Directorate General of Taxes rejected this argument and firmly applied a single-year approach using 2017 data. The tax authority also made a technical adjustment by eliminating eight of the twelve comparable companies considered to have extreme values in order to determine the median arm's-length profit. The complete rejection of the appeal and the upholding of an upward arm's-length profit adjustment amounting to IDR 76.21 billion raised significant legal concerns regarding the consistency of evidentiary standards and legal certainty in tax practices within the manufacturing

DOI: 10.65246

sector.

To examine the root causes of these issues, this study employs Information Asymmetry Theory to analyze the unequal control of information between the Taxpayer and the Directorate General of Taxes, as well as Agency Theory to examine differences in interests, strategic behavior, and the need for monitoring within the taxation relationship (Akerlof, 1970; Eisenhardt, 1989; Jensen & Meckling, 1976). The relevance of these theoretical frameworks is consistent with the findings of Sebele-Mpofu et al. (2025) and Sukarno (2022), who emphasize that information asymmetry in transfer pricing documentation is a major factor triggering fiscal adjustments by tax authorities. Furthermore, research by Ganesh Lal Meena and Neetu Zora (2025) and Muzychuk and Fomina (2021) shows that disagreements concerning the application of the TNMM and Functions, Assets, and Risks (FAR) analysis frequently lead to the escalation of disputes to the Tax Court. This is reinforced by the study of Matavulj (2021), which demonstrates that differences in the use of single-year and multiple-year data have a highly significant effect on the determination of the interquartile range of arm's-length profits. Meanwhile, Martensen (2020) highlights that the elimination of comparable companies with extreme values directly affects the reliability of evidence presented during court proceedings. Complementing these findings, Arrisman and Fauzi (2025) conclude that the strength of the Taxpayer's factual evidence substantially influences the direction of the judicial panel's legal considerations in deciding disputes. Collectively, these previous studies reinforce the urgency of this research in examining the factors contributing to the Taxpayer's defeat and their implications for judicial review strategies.

Previous studies have generally focused on examining transfer pricing disputes from the perspective of compliance levels or solely through normative quantitative analyses of the selection of tax calculation methods. These studies consistently confirm that information asymmetry and differences in the choice of margin-testing methodologies are the primary factors triggering fiscal adjustments by tax authorities. Therefore, the existing literature has not extensively integrated the legal, technical-methodological, and factual evidentiary dimensions within a single case study of a specific court decision. The main distinction of this study lies in its in-depth empirical analysis of Tax Court Decision Number PUT-014532.15/2020/PP/M.XIIIB of 2022 to precisely identify the factors contributing to the Taxpayer's defeat in the manufacturing sector. This study specifically examines the judicial panel's legal considerations regarding the validity of using 2017 single-year data and the elimination of comparable companies with extreme values in determining the interquartile range of arm's-length profits. In addition, this study provides a strategic overview of the implications of the Tax Court's legal considerations for the potential and formulation of a Judicial Review before the Supreme Court. Accordingly, this study seeks to address the academic gap concerning the evidentiary dynamics of transfer pricing disputes in the manufacturing sector through a more focused analysis.

The primary research novelty offered by this study lies in its empirical novelty within the field of transfer pricing dispute studies in Indonesia. This empirical novelty is demonstrated through an in-depth and comprehensive examination of Tax Court Decision Number PUT-014532.15/2020/PP/M.XIIIB of 2022, which adjudicated a transfer pricing dispute involving a manufacturing entity. More specifically, this novelty is reflected in the detailed examination of the judicial panel's legal considerations regarding the validity of using 2017 single-year data and the elimination of comparable companies with extreme values. In addition, this study extends its empirical contribution by linking the facts presented before the Tax Court to the identification of potential grounds and the formulation of a Judicial Review strategy before the Supreme Court. This empirical analysis is further strengthened by Information Asymmetry Theory and Agency Theory to examine the dynamics of conflicting interests and information imbalances between the Taxpayer and the tax authority. This approach presents a form of novelty that has not been widely explored in previous studies, namely the simultaneous and integrated synthesis of legal, technical-methodological, and factual evidentiary dimensions. By emphasizing this empirical novelty, the findings are expected to enrich the academic taxation literature while also providing strategic practical guidance for Taxpayers and tax authorities in managing tax dispute risks.

This study employs a qualitative method and is designed based on an interpretive paradigm to

DOI: 10.65246

obtain an in-depth understanding of the substantive meaning, argumentative dynamics, and ratio decidendi underlying Tax Court Decision Number PUT-014532.15/2020/PP/M.XIIIB of 2022 (Aisya & Nuryanah, 2024; Suharsono et al., 2024; Turwanto et al., 2022). In line with this paradigm, this study aims to (1) analyze the factors supporting the Directorate General of Taxes' victory, (2) examine the application of the Arm's Length Principle, and (3) assess its implications for Judicial Review. To address these objectives, secondary data consisting of court decision documents, relevant tax regulations, and supporting documents, as well as primary data obtained through interviews, were systhemetically analyzed using thematic analysis to identify the principal patterns and themes underlying the legal debate (Muhammadi et al., 2016; Parmalia & Rosid, 2023; Prayudi & Rosid, 2025). In practical terms, this study provides a tangible contribution to manufacturing-sector Taxpayers and tax consultants in preparing responsive transfer pricing documentation and mitigating the risk of fiscal adjustments before the Tax Court. On the other hand, the findings are also useful for tax authorities as evaluative material for improving the objectivity of comparability testing and the consistency of law enforcement. Accordingly, this study is expected to serve as an additional reference for stakeholders in formulating more precise evidentiary arguments in transfer pricing disputes.

2. Literature Review

2.1. Conceptual and Theoretical Foundations

This study is grounded in Agency Theory and Information Asymmetry Theory, which provide complementary perspectives for examining transfer pricing disputes between taxpayers and tax authorities. Agency Theory explains potential conflicts of interest arising from differences in objectives between parties, particularly when multinational enterprises seek to optimize their global tax positions through related-party transactions (Eisenhardt, 1989; Jensen & Meckling, 1976). Information Asymmetry Theory further explains how differences in access to internal financial, operational, and functional information may complicate the assessment of transfer pricing compliance (Keen & Slemrod, 2017; Kleven et al., 2011; Murphy et al., 2019). In such circumstances, disagreements may arise regarding profit allocation, the functions performed, assets employed, risks assumed, and the reliability of supporting documentation (Arabyan, 2023; Mahmudi, 2022; Melega et al., 2022). These theories provide the analytical foundation for understanding the competing arguments presented by taxpayers and tax authorities and the judicial reasoning applied in resolving transfer pricing disputes.

The Arm's Length Principle (ALP) constitutes the principal normative framework for evaluating whether related-party transactions reflect conditions that would have prevailed between independent parties. Under the OECD Transfer Pricing Guidelines and Regulation of the Director General of Taxes Number PER-32/PJ/2011, the application of the ALP requires an appropriate comparability analysis. One method frequently applied in practice is the Transactional Net Margin Method (TNMM), which evaluates the net profit earned relative to an appropriate base, such as costs, sales, or assets. The Operating Margin (OM) may serve as the profit level indicator, depending on the functional and economic characteristics of the tested party. The reliability of the TNMM depends on the appropriateness of the selected method, the comparability of the companies examined, and the consistency of the financial data used (Khris & Whiteside, 2020; Matavulj, 2021; Sulik-Górecka, 2018).

In addition to these theoretical and methodological foundations, the study incorporates the legal framework governing evidence and judicial review. Under Law Number 14 of 2002 concerning the Tax Court, the assessment of evidence is central to establishing the material facts underlying a tax dispute. Taxpayers must substantiate their arguments through relevant financial records, accounting documents, and other supporting evidence. Where a party subsequently challenges the Tax Court's decision, judicial review before the Supreme Court provides an extraordinary legal remedy subject to the applicable statutory grounds. Accordingly, the interaction between information asymmetry, transfer pricing methodology, evidentiary reliability, and judicial review forms the principal analytical basis of this study.

2.2. Review of Empirical Studies

Previous studies have examined transfer pricing from the perspectives of corporate taxation, international profit allocation, comparability analysis,

DOI: 10.65246

and the economic consequences of multinational enterprises' tax planning. Research on international profit shifting demonstrates that differences in tax rates and the allocation of economic activities may influence the distribution of reported profits across jurisdictions (Bartelsman & Beetsma, 2003; Davies et al., 2018). More recent discussions emphasize the importance of examining the economic substance of related-party transactions and the consistency between reported financial outcomes and the functions, assets, and risks of the entities involved (Beuselinck & Pierk, 2024; Brychta et al., 2024; Mwape et al., 2025).

Another strand of literature focuses on the methodological challenges associated with applying the TNMM. The selection of comparable companies, the use of single-year or multiple-year financial data, and the treatment of extreme observations can materially affect the arm's length range and the resulting transfer pricing adjustment. Studies discussing comparability and financial data selection highlight the importance of consistent criteria and the need to account for differences in operating conditions (Barton et al., 2021; De Robertis & Kondi, 2022; Dessatniuk, 2015; Ökten et al., 2025; Steens et al., 2022). In particular, the use of single-year data may provide contemporaneous evidence of market conditions, whereas multiple-year data may help explain business cycles or unusual operating circumstances. Similarly, excluding outliers may improve the representativeness of the comparable sample, but the decision requires a defensible methodological basis because it can alter the interquartile range and the assessment of arm's length profitability (Chyzhykova, 2019; Muyaa, 2014).

Despite these methodological discussions, transfer pricing disputes cannot be resolved solely through numerical comparisons. Claims that losses result from declining production capacity, adverse market conditions, or other commercial circumstances must be evaluated against the available accounting records and supporting evidence (Berdonce et al., 2021). The evidentiary process is therefore essential in determining whether a taxpayer's methodological and commercial explanations are sufficiently substantiated. Nevertheless, the reviewed literature, including discussions of judicial decision-making and legal review (Pelekis, 2021; Patel et al., 2023), indicates the importance of examining how technical transfer pricing arguments are translated into legal reasoning. This issue is particularly relevant when disputes progress from the Tax Court to judicial review, where the applicable legal grounds differ from a general reassessment of factual disagreements.

2.3. Identification of the Research Gap

The reviewed literature establishes the importance of the ALP, the TNMM, comparability analysis, and the quality of evidence in determining transfer pricing compliance. However, three interrelated gaps motivate the present study. First, much of the existing discussion emphasizes transfer pricing practices, profit shifting, or methodological considerations without integrating these dimensions into a comprehensive analysis of judicial reasoning. Second, disagreements concerning the selection of comparable companies, the use of single-year financial data, and the treatment of outliers require further examination in relation to the evidentiary standards applied in resolving actual tax disputes. Third, the legal implications of these methodological disagreements at the judicial review stage remain insufficiently explained when considered alongside the statutory grounds for challenging a Tax Court decision.

Accordingly, this study addresses the gap by examining how transfer pricing methodologies and comparability arguments are presented, evaluated, and substantiated in judicial proceedings. It focuses on the relationship between the taxpayer's arguments, the tax authority's assessment, the evidentiary considerations of the Tax Court, and the legal grounds relevant to judicial review before the Supreme Court. The study's contribution lies in integrating Agency Theory, Information Asymmetry Theory, the ALP, and evidentiary analysis within a legal framework for understanding transfer pricing dispute resolution. The analysis also distinguishes methodological disagreements concerning the merits of a transfer pricing assessment from legal errors or other statutory grounds that may justify judicial review.

2.4. Development of the Conceptual Framework

The conceptual framework positions Agency Theory and Information Asymmetry Theory as the overarching analytical perspectives for examining transfer pricing disputes. Agency Theory explains the

DOI: 10.65246

potential divergence between taxpayers' commercial and tax-planning interests and the tax authority's responsibility to ensure compliance with applicable tax regulations. Information Asymmetry Theory explains why the parties may reach different conclusions when evaluating related-party transactions because they do not possess equivalent access to financial and operational information.

Within this framework, the ALP provides the normative standard for assessing the arm's length nature of related-party transactions, while the TNMM represents the transfer pricing method examined in the dispute. Its application is evaluated through the selection of comparable companies, the choice of financial data periods, the treatment of outliers, and the determination of an appropriate profitability indicator. These methodological issues generate competing arguments that must be assessed against the available documentary and financial evidence.

The evidentiary assessment subsequently informs the Tax Court's evaluation of the parties' arguments and the factual basis of its decision. If the decision is challenged, the judicial review stage examines whether the challenge satisfies the applicable statutory grounds, rather than automatically reopening every disagreement concerning the assessment of facts. The framework therefore connects theoretical explanations of the dispute, the technical application of transfer pricing methods, the evaluation of evidence, and the legal review of judicial decisions.

2.5. Research Propositions

As this study adopts a legal-analytical approach to examining transfer pricing disputes rather than testing statistical relationships among variables, the following research propositions guide the analysis.

Proposition 1: Differences in information access and economic interests between taxpayers and tax authorities contribute to disagreements over the arm's length nature of related-party transactions.

Proposition 2: The selection of comparable companies, the use of financial data periods, and the treatment of outliers influence the methodological arguments used to assess compliance with the ALP through the TNMM.

Proposition 3: The strength and credibility of documentary and financial evidence play a central role in determining how the Tax Court evaluates competing transfer pricing arguments and establishes the material facts of the dispute.

Proposition 4: The legal significance of transfer pricing methodological disagreements at the judicial review stage depends on whether the challenge can be linked to the statutory grounds for judicial review, rather than merely reflecting dissatisfaction with the Tax Court's factual assessment.

These propositions provide a structured basis for examining the relationship between theoretical explanations, methodological disagreements, evidentiary assessment, and judicial reasoning in transfer pricing disputes.

3. Research Methods

3.1. Research Design

This study employs a qualitative research design grounded in an interpretive paradigm to examine the dynamics of transfer pricing disputes in Indonesia (Qadri, 2022). The interpretive approach recognizes that legal realities and economic practices are understood through the interpretations of legal subjects dealing with complex tax transactions (Qadri & Fatmawati, 2021). In transfer pricing disputes, the acceptance or rejection of a particular methodology depends not only on mathematical calculations but also on the interpretation of comparability evidence, functional analysis, and business risks (Dwindahany & Efendi, 2024; Tambunan, 2020). This design enables an in-depth reconstruction of the arguments presented by the Taxpayer and the Directorate General of Taxes (DGT), the legal considerations of the Tax Court, and the implications of those considerations for potential Judicial Review. The study seeks analytical depth and contextual understanding rather than statistical generalization.

3.2. Research Context and Setting

The research focuses on an Indonesian transfer pricing dispute involving PT X, as documented in Tax Court Decision Number PUT-014532.15/2020/PP/M.XIIIB of 2022. The case provides a context for examining the application of the Arm's Length Principle (ALP), the Transactional Net Margin Method (TNMM), the Operating Margin (OM) as a profit level indicator, and the determination of an arm's-length profit adjustment.

DOI: 10.65246

Particular attention is given to disagreements concerning the tested party, the functional, assets, and risks (FAR) analysis, the use of single-year and multiple-year financial data, the selection of comparable companies, the treatment of extreme values, and the use of the median in calculating the adjustment. The research also examines the implications of the Tax Court's reasoning for assessing the possibility of Judicial Review before the Supreme Court.

3.3. Population and Sample / Research Participants

As a qualitative case study, this research does not employ a statistical population or probability-based sampling. The principal unit of analysis is the transfer pricing dispute involving PT X, represented by the relevant Tax Court decision and associated legal documents. The primary research participant is one key informant, identified as Tax Court Judge HK.01, who was selected for the relevance of their judicial knowledge and experience to the research questions. The informant was interviewed using a semi-structured approach to obtain additional insights into judicial reasoning, the establishment of material facts, the evaluation of comparability analyses, and the assessment of technical tax evidence. The selection of a key informant supports the study's interpretive objective, although the findings are not intended to represent the views of all Tax Court judges.

3.4. Data Sources and Data Collection

This study integrates primary and secondary data. Primary data were obtained through a semi-structured interview with the key informant, Tax Court Judge HK.01. The interview explored the criteria used to establish material truth, the reliability of the Taxpayer's comparability analysis, the rejection of arguments concerning comparable data, and the process through which judges form their judicial conviction when evaluating complex technical tax information. Secondary data were collected through a structured legal document study comprising Tax Court Decision Number PUT-014532.15/2020/PP/M.XIIIB of 2022 concerning PT X, the Judicial Review memorandum and counter-memorandum, relevant domestic tax laws and regulations, technical circular letters issued by the DGT, and the OECD Transfer Pricing Guidelines. The collection process involved systematically recording and categorizing documentary information according to the principal issues in dispute, information asymmetry, methodological disagreements, and the establishment of material facts. Interview transcripts were anonymized and compared with the documentary evidence to support the interpretation of the case.

3.5. Measurement of Variables and Research Instruments

Because this study adopts a qualitative interpretive approach, it does not operationalize variables through numerical indicators or standardized measurement scales. Instead, the research examines analytical concepts derived from the research questions and theoretical framework. The principal concepts comprise (1) legal, technical-methodological, and evidentiary factors underlying the dispute; (2) the application of the ALP and TNMM in determining the arm's-length profit adjustment; and (3) the implications of the Tax Court's legal reasoning for evaluating potential Judicial Review. These concepts are examined through the tested-party characterization, FAR analysis, the selection of comparable companies, the use of single-year or multiple-year financial data, the treatment of outliers, the determination of the arm's-length range and median, the sufficiency of operational evidence, and the traceability of judicial reasoning.

The primary research instrument was a semi-structured interview guide designed around the research questions. Documentary analysis provided the complementary instrument for systematically examining court decisions, legal submissions, regulations, technical guidance, and the OECD Transfer Pricing Guidelines. The analytical categories were subsequently refined through thematic coding, allowing the interpretation to remain grounded in the empirical and legal materials.

3.6. Data Analysis Techniques

The data were analyzed using Thematic Analysis developed by Braun and Clarke (2006). The procedure involved five interconnected analytical activities: familiarization with the data, initial coding, grouping related codes into subthemes and themes, reviewing and refining the themes, and interpreting

DOI: 10.65246

the established themes in relation to the research questions and theoretical framework. During familiarization, the researcher repeatedly examined the court decision, legal documents, applicable regulations, and interview transcripts to understand the dispute's factual and legal context. Initial codes were assigned to meaningful data concerning legal arguments, transfer pricing methodology, comparability analysis, evidentiary sufficiency, and judicial reasoning. Related codes were then organized into subthemes and broader themes, which were reviewed for internal coherence and consistency with the complete dataset.

The final interpretation integrated the themes with Agency Theory, Information Asymmetry Theory, relevant domestic tax regulations, the OECD Transfer Pricing Guidelines 2022, and previous studies. This process addressed three research questions: the legal, technical-methodological, and evidentiary factors underlying the dispute; the application of the ALP and transfer pricing methods in determining the arm's-length profit adjustment; and the implications of the Tax Court's legal reasoning for evaluating potential Judicial Review. The analysis also distinguished disagreements concerning factual assessment and transfer pricing methodology from legal issues potentially relevant to the statutory grounds for Judicial Review.

Table 1. Thematic Coding Matrix

Research Question		Theme No.	Main Theme	Subtheme No.	Subtheme	Final Codes
RQ 1	What legal, technical-methodological, and evidentiary factors led the DGT to prevail in the dispute?	Theme 3	Commercial Losses That Were Not Operationalized	3.1	Contesting the meaning of losses through causal evidence	M-K01 — Underutilization as an explanation for losses M-B01 — Insufficient evidence of the causal relationship underlying the losses M-B03 — Sufficiency of evidence is more decisive than argumentation
		Theme 3	Commercial Losses That Were Not Operationalized	3.2	Capacity adjustment as an alternative that was not adopted	M-A01 — Adjustment as a distinct analytical stage M-A02 — Capacity adjustment as an alternative
		Theme 4	From Benchmark to Ruling: Judicial Validation of the DGT's Adjustment	4.1	Statutory authority and normative sources underpin the arm's-length examinatio	M-Y01 — The DGT's statutory authority to assess arm's-length compliance M-N01 — Integration of domestic regulations and the OECD Guidelines
		Theme 4	From Benchmark to Ruling: Judicial Validation of the DGT's Adjustment	4.2	Acceptance of the methodology culminates in the full upholding of the adjustment	M-R03 — The DGT's comparability analysis was deemed more consistent with the ALP M-R04 — Full adoption of the DGT's calculation without an alternative reconstruction M-C02 — Complete dismissal of the appeal and its fiscal consequences
RQ 2	How were the ALP and transfer pricing method applied in determining the arm's-length profit adjustment?	Theme 1	The Same Method, a Different Tested Party	1.1	Agreement on the method shifts the dispute to the quality of its application	M-C01 — Implementation of TNMM and OM as the central area of dispute M-F01 — Sequencing of the ALP examination M-F06 — Relationship among FAR, TNMM, and comparable companies

DOI: 10.65246

Research Question	Theme No.	Main Theme	Subtheme No.	Subtheme	Final Codes
	Theme 1	The Same Method, a Different Tested Party	1.2	Business-process reality as the basis for constructing the FAR analysis	M-F02 — Business processes and documents as the basis for the FAR analysis M-F03 — Functions as the central element of the FAR relationship
	Theme 1	The Same Method, a Different Tested Party	1.3	Contestation over the PGF's characterization and the expectation of routine profit	M-F04 — Contested business characterization of the PGF M-F05 — Expectation of routine profit for a limited-function entity
	Theme 2	Arm's-Length Outcomes Constructed through Time, Comparables, and the Median	2.1	The choice of period changes how performance is interpreted	M-W01 — Priority of contemporaneous single-year data M-W02 — Multiple-year data as a conditional approach M-W03 — The need for multiple-year data was not established M-W04 — Multiple-year data used to explain losses M-W05 — Multiple-year data used to construct the profit range
	Theme 2	Arm's-Length Outcomes Constructed through Time, Comparables, and the Median	2.2	The set of comparables is constructed through inclusion and exclusion decisions	M-P01 — Contestation over twelve versus four comparable companies M-P02 — Data completeness as a selection criterion M-P03 — Factual disagreement regarding data availability M-P04 — Elimination of extreme values M-P05 — Candidate comparables versus the final set of comparables M-P06 — Sufficiency of four comparable companies
	Theme 2	Arm's-Length Outcomes Constructed through Time, Comparables, and the Median	2.3	The range determines the non-arm's-length outcome, while the median determines the adjustment	M-R01 — The PGF's OM fell outside the range calculated by the Respondent M-R02 — The 7.49% median was used as the adjustment point without a specific explanation
RQ 3 What are the implications of the Tax Court's legal reasoning for evaluating a potential Judicial Review?	Theme 5	Scope for Judicial Review between the Traceability of Judicial Reasoning and the Dominance of Fact Assessment	5.1	Formal and remedial arguments fall outside the central reasoning of the Panel of Judges	M-Y02 — Argument concerning the invalidity of the adjustment due to its non-attribution to specific transactions M-Y03 — Obstacles to corresponding adjustment, MAP, and the prevention of double taxation M-Y04 — Absence of an explicit response to formal and remedial arguments

DOI: 10.65246

Research Question	Theme No.	Main Theme	Subtheme No.	Subtheme	Final Codes
	Theme 5	Scope for Judicial Review between the Traceability of Judicial Reasoning and the Dominance of Fact Assessment	5.2	The fact–methodology–conclusion chain is not fully articulated	M-F07 — Absence of an explicit determination of the PGF's characterization M-P07 — The analysis of individual comparable companies was not detailed M-B02 — Aggregate assessment of operational evidence M-J01 — The absence of an explicit response does not necessarily mean that an argument was not considered
	Theme 5	Scope for Judicial Review between the Traceability of Judicial Reasoning and the Dominance of Fact Assessment	5.3	The boundary between legal and factual issues constrains the evaluation of Judicial Review	M-J02 — Variations in judicial approaches to problematic adjustments M-PK01 — The need to distinguish issues of law from the assessment of facts M-PK02 — Lack of direct data on the Judicial Review proceedings and newly discovered evidence M-PK03 — The Judicial Review strategy or outcome cannot be predicted from the current dataset

Source: processed by the author.

3.7. Validity, Reliability, and Trustworthiness

The trustworthiness of the findings was supported through data source triangulation. Information from the Tax Court decision was critically compared with the interview transcript, relevant tax laws and regulations, the OECD Transfer Pricing Guidelines, the Judicial Review memorandum and counter-memorandum, and previous studies. This process helped assess the consistency of the evidence and reduced reliance on any single source. The thematic coding process also involved reviewing the boundaries and definitions of codes, subthemes, and themes to maintain analytical coherence. Furthermore, the researcher compared the final interpretations with the underlying documentary and interview evidence to ensure that the conclusions remained traceable to the available data. These procedures support the credibility and dependability of the analysis while recognizing that the interpretation is situated within a single transfer pricing dispute.

3.8. Ethical Considerations

The study addresses confidentiality in the collection and interpretation of primary interview data. The interview participant is identified using the designation HK.01, and the interview transcripts were anonymized to protect confidentiality. The documentary analysis relies on the relevant court decision, legal submissions, regulations, and other materials used for the stated research purposes. The study does not disclose additional identifying information about the informant beyond the designation provided in the research materials. However, the available methodological records do not specify whether formal ethics committee approval, written informed consent, or a separate data-protection protocol was obtained. These details should be reported in the final manuscript if applicable and supported by the actual research records.

3.9. Research Procedure

The research was conducted through a sequence of interconnected stages. First, the researcher established the interpretive framework and formulated the research questions concerning the transfer pricing dispute. Second, the researcher collected and examined the Tax Court decision, Judicial Review memorandum and counter-memorandum, domestic tax regulations, DGT technical circular letters, and the OECD Transfer

DOI: 10.65246

Pricing Guidelines. Third, primary data were collected through a semi-structured interview with the key informant, Tax Court Judge HK.01. Fourth, the documentary materials and interview transcript were examined through repeated readings and systematic coding. Fifth, related codes were grouped into subthemes and themes, reviewed, and refined to establish the final thematic structure. Sixth, the themes were interpreted using Agency Theory, Information Asymmetry Theory, applicable legal provisions, and previous studies. Finally, data source triangulation was conducted, and conclusions were drawn inductively to answer the three research questions and evaluate the legal and methodological implications of the dispute for potential Judicial Review.

3.10. Methodological Limitations

This study has several methodological limitations. First, its focus on a single transfer pricing dispute involving PT X provides an in-depth understanding of the case but limits the transferability of the findings to other disputes, industries, or judicial settings. Second, the primary data rely on one key informant, meaning that the findings may not capture the full range of judicial perspectives or the experiences of other parties involved in transfer pricing litigation. Third, the analysis depends on the availability, completeness, and level of detail of the court decision, legal submissions, interview transcript, and supporting documents. Fourth, although the study examines the implications of the Tax Court's reasoning for potential Judicial Review, the available dataset does not establish the outcome of subsequent Judicial Review proceedings or confirm the existence of newly discovered evidence (*novum*). Therefore, the study evaluates the legal and analytical considerations relevant to a potential review rather than predicting its outcome. These limitations should be considered when interpreting the findings and provide directions for future research involving multiple cases, additional judicial informants, and more extensive documentation of Judicial Review proceedings.

4. Results and Discussion

4.1. Research Results

4.1.1. Sample Description and Descriptive Statistics

This study examines the transfer pricing dispute involving PT X, as documented in Tax Court Decision Number PUT-014532.15/2020/PP/M.XIIIB of 2022. The unit of analysis comprises the legal, technical-methodological, and evidentiary aspects of the dispute, including the application of the Arm's Length Principle (ALP), the Transactional Net Margin Method (TNMM), and the implications of the Tax Court's legal reasoning for potential Judicial Review. The primary data were obtained through a semi-structured interview with one key informant, identified as Tax Court Judge HK.01, while secondary data comprised the court decision, the Judicial Review memorandum and counter-memorandum, relevant tax laws and regulations, technical guidance, and the OECD Transfer Pricing Guidelines.

As a qualitative study, the research does not report statistical descriptive measures such as means, standard deviations, or demographic percentages. Instead, the descriptive findings concern the characteristics of the dispute, the positions of the parties, the transfer pricing calculations, and the judicial considerations.

The Appellant disputed a fiscal adjustment of IDR 76,219,880,536, classified as Other Positive Fiscal Adjustments, because it was not attributed to specific related-party transactions. The Respondent, namely the Directorate General of Taxes (DGT), maintained that its assessment of the Power Generation Factory (PGF) segment was consistent with the ALP. Both parties applied the TNMM using Operating Margin (OM) as the Profit Level Indicator. However, they disagreed over the characterization of PGF, the selection of financial data periods, the number and suitability of comparable companies, the treatment of extreme values, and the adjustment required to establish an arm's-length profit.

4.1.2. Data Quality and Preliminary Analysis

Because this research employs a qualitative interpretive design, statistical tests of normality, multicollinearity, reliability coefficients, and model fit are not applicable. Data quality was assessed through systematic examination of documentary evidence, repeated readings of the interview transcript and court decision, and comparison of the parties' arguments with the applicable legal and technical standards.

The analytical process followed Thematic

DOI: 10.65246

Analysis developed by Braun and Clarke (2006), involving data familiarization, initial coding, grouping codes into subthemes and themes, reviewing the themes, and interpreting the resulting patterns. Data source triangulation was conducted by comparing the Tax Court decision, the interview with HK.01, relevant legal provisions, the OECD Transfer Pricing Guidelines 2022, and previous studies. Particular attention was given to distinguishing the documented contents of the decision from the informant's general explanation of judicial practice and the researcher's own interpretation.

This process identified three principal analytical dimensions: the legal, technical-methodological, and evidentiary factors underlying the DGT's victory; the application of the ALP and TNMM in determining the profit adjustment; and the implications of the Tax Court's reasoning for potential Judicial Review. The findings are presented according to these dimensions.

4.1.3. Main Analytical Results

a. Factors Behind the Directorate General of Taxes' Victory

The findings indicate that the legal factor supporting the Respondent's victory primarily lay in the acceptance of the Directorate General of Taxes' authority to assess the arm's-length nature of related-party transactions, rather than in the explicit rejection of all legal arguments advanced by the Appellant. The Appellant disputed the adjustment of IDR 76,219,880,536, which was classified as Other Positive Fiscal Adjustments because it was not attributed to any specific related-party transaction. According to the Appellant, the lack of clarity regarding the type of transaction, the adjustment amount, the counterparty, and the counterparty's domicile hindered the implementation of a corresponding adjustment and the mutual agreement procedure and potentially resulted in double taxation. The Respondent, by contrast, maintained its authority to conduct an arm's-length assessment of the PGF segment. The Judicial Panel did not provide a specific response to each element of these formal-remedial arguments but focused its considerations on the quality of the comparability analysis and the conformity of the adjustment with the provisions applied in the decision. During the interview, HK.01 also emphasized that the preparation of Transfer Pricing Documentation does not eliminate the authority of the Directorate General of Taxes to assess arm's-length prices or profits. However, HK.01's view represents a general explanation of practice rather than the Judicial Panel's direct reasoning in this case. Therefore, the authority to conduct an assessment and the accuracy of the adjustment must be analytically distinguished. The researcher's interpretation indicates that the legal factor served as the basis for legitimizing the assessment, whereas the victory of the Directorate General of Taxes remained dependent on the Judicial Panel's acceptance of its technical application.

Technical-methodological factors constituted the most apparent determinants in the decision. The Appellant and the Respondent both applied the Transactional Net Margin Method using Operating Margin as the Profit Level Indicator, meaning that the dispute did not concern the selection of the method but rather its application. The Respondent characterized PGF as a contract manufacturer that should reasonably earn a profit, used single-year data for 2017, and retained four comparable companies. This set of comparables produced an interquartile range of 6.55%–9.04%, with a median of 7.49%, whereas PGF recorded a negative Operating Margin of 14.78%. The Appellant rejected the limited-function characterization, used twelve comparable companies, and considered multiple-year data to be more representative. The Judicial Panel accepted the use of 2017 single-year data, considered the Appellant's data to contain extreme values and incomplete data, and stated that four comparables were sufficient. However, the Judicial Panel did not explicitly determine PGF's business characterization and did not provide detailed reasons for the elimination of each company. This issue is consistent with HK.01's explanation that the determination of FAR must begin with an examination of the actual business activities:

> *"The first thing the Judicial Panel asks the parties about is the business process, because it determines the FAR. The second is the contracts, and the third is the deed of establishment."*
>
> (HK.01, translated by the researcher)

HK.01 explained that FAR, the method, and the comparables should be examined sequentially through the business processes, contracts, functions, and actual data. The researcher's interpretation

DOI: 10.65246

indicates that the technical victory of the Directorate General of Taxes primarily resulted from the acceptance of the selected period, the screening of comparable companies, and the use of the median, rather than from explicit agreement with the Respondent's entire FAR construction.

The evidentiary factor was reflected in the Appellant's failure to transform the 2017 loss into a commercial explanation accepted by the Judicial Panel. The Appellant linked PGF's loss to the decline in production hours from 633,065 in 2016 to 484,164 in 2017, while the break-even point was estimated at 499,086 production hours. These data were used to demonstrate that production volume was below the break-even point, meaning that sales were insufficient to cover operating costs. The Appellant also used data from 2015–2017 and proposed a capacity adjustment if the assessment continued to rely on a single year. The Respondent considered that the decline in volume was insufficient to explain the change in margin and maintained the expectation that an entity performing limited functions should earn a profit. The Judicial Panel did not state that the production-hour or break-even-point data were incorrect but considered that the Appellant's arguments, facts, and evidence were insufficient to justify the use of multiple-year data or invalidate the Respondent's benchmark. The considerations also did not include an alternative calculation of the capacity adjustment. The importance of evidentiary quality in this assessment was directly emphasized by HK.01:

> *"Judges look at the evidence: how valid it is, how sufficient it is, and how competent it is. As long as the evidence is competent, sufficient, and valid, that is what is used."*
>
> (HK.01, translated by the researcher)

HK.01 stated that judges assess the validity, sufficiency, and competence of evidence, whereas arguments merely assist in understanding the dispute. The researcher's interpretation indicates that the Appellant's position was weakened not because the operational evidence was declared false, but because the evidence failed to establish a causal relationship and an adjustment mechanism that could be accepted in the final calculation.

Overall, the victory of the Directorate General of Taxes resulted from the interrelationship between legal legitimacy, methodological acceptance, and evidentiary sufficiency. At the legal level, the authority of the Directorate General of Taxes to assess related-party transactions was accepted as the basis for the adjustment, while the Appellant's arguments concerning the non-attribution of transactions, corresponding adjustment, and the mutual agreement procedure did not become the central focus of the written considerations. At the technical level, the Judicial Panel considered that the use of 2017 data, the elimination of extreme values or incomplete data, and the use of four comparables were more consistent with the arm's length principle. This acceptance placed PGF's margin outside the arm's-length range and led to an adjustment based on the median of 7.49%. At the evidentiary level, the data concerning underutilization, the break-even point, multiple-year analysis, and the proposed capacity adjustment were insufficient to alter the outcome of the Respondent's analysis. Nevertheless, conflicting evidence remained: the four comparables originated from the Appellant's Transfer Pricing Documentation, the use of multiple-year data was not rejected absolutely, and PGF's business characterization was not explicitly determined by the Judicial Panel. HK.01 also indicated that judges may recalculate an adjustment or partially grant an appeal in other cases, but this view cannot be attributed to the Judicial Panel in the present case. The researcher's interpretation is that the Directorate General of Taxes prevailed because its comparability framework was considered more reliable, whereas the Appellant's alternative failed to achieve an equivalent level of evidentiary strength and operationalization.

From a legal perspective, the victory of the Directorate General of Taxes cannot be explained solely by the existence of a special relationship, but rather by its success in transforming its authority to make adjustments into an assessment deemed valid and accountable. Article 18 paragraph (3) of UU PPh authorizes the Directorate General of Taxes to redetermine income and deductions based on the arm's length principle, while Article 29 paragraph (1) of UU KUP establishes tax audits as an instrument for assessing compliance. This framework is further reinforced by Article 20 paragraphs (1)–(4) of PER-43/PJ/2010, as amended by PER-32/PJ/2011, which requires consideration of the method and documentation prepared by the Taxpayer and limits adjustments when the Arm's Length Principle has been satisfied. Under an adaptive application of

DOI: 10.65246

Agency Theory, this authority functions as a monitoring mechanism for controlling differences in interests between the state and the Taxpayer, but the theory does not render every adjustment automatically correct. The findings instead demonstrate that legal legitimacy resulted in victory only after the Judicial Panel accepted the quality of its application. This finding supports Tambunan (2020), who found that the Directorate General of Taxes was more reliable when complying with technical and administrative provisions, while extending the study by demonstrating the concrete relationship between authority, analytical quality, judicial assessment, and the ruling that rejected the appeal in full.

From a technical-methodological perspective, the acceptance of the TNMM does not eliminate the need to assess the reliability of all decisions shaping the outcome of the method. Paragraph 2.74 of the OECD Transfer Pricing Guidelines 2022 emphasizes that the TNMM continues to require a robust comparability analysis, while paragraph 3.68 prioritizes contemporaneous comparable data because they reflect the economic environment during the period in which the transaction occurred. However, paragraphs 3.75–3.76 state that multiple-year data remain useful when they provide added value, particularly in understanding the causes of losses. Similarly, paragraph 3.63 prohibits the elimination of extreme results solely because their values differ, while paragraph 3.62 permits the use of the median when unidentified or unquantifiable comparability defects remain. Therefore, the victory of the Directorate General of Taxes was not an automatic consequence of using a single-year approach, four comparables, or the median, but resulted from the Judicial Panel's assessment that this configuration was more reliable than the Appellant's alternative. This finding supports Zulfiqar (2023), regarding the importance of documentation, comparable data, and comparability methods. The difference is that previous research was dominated by decisions favorable to Taxpayers, whereas this case demonstrates circumstances in which the comparable-company selection and arm's-length range constructed by the Directorate General of Taxes were accepted as the basis for the adjustment.

The evidentiary factor explains why the Appellant's commercial explanation failed to alter the technical outcome. Article 76 of Tax Court Law authorizes judges to determine the matters to be proven, the burden of proof, and the assessment of evidence based on at least two items of evidence, while Article 78 requires decisions to be based on the assessment of evidence, tax laws and regulations, and the judges' conviction. By analogy, within the framework of Information Asymmetry Theory, the Taxpayer possesses internal information regarding production capacity, cost structures, and the causes of losses, but this informational advantage only becomes valuable when it can be translated into verifiable evidence and replicable calculations. The proposed capacity adjustment must also satisfy paragraph 3.50 of the OECD Transfer Pricing Guidelines 2022, namely that it is material, supported by high-quality data, and improves the reliability of the comparison. This finding extends Dwindahany and Efendi (2024), who identified weaknesses in the tax authority's substantiation of adjustments as a cause of disputes. In this case, the evidentiary issue operated in both directions: the Directorate General of Taxes was required to substantiate its benchmark, while the Appellant was also required to establish a relationship between underutilization and the loss and operationalize that relationship into an adjustment. The failure at the latter stage meant that the private information was insufficiently robust, measurable, and testable to replace the Respondent's calculation, which was ultimately accepted by the Judicial Panel as the basis for upholding the adjustment.

In synthesis, the victory of the Directorate General of Taxes was formed through three interdependent layers: normative authority provided the basis for monitoring, methodological decisions constructed the benchmark, and evidence determined which benchmark could be sustained. Agency Theory helps explain the monitoring incentives and differences in the parties' interests, but it is insufficient to determine whether single-year data, comparable companies, or the median were appropriately used. Information Asymmetry Theory has more direct explanatory power because the dispute centered on the quality, completeness, comparability, and verifiability of information; nevertheless, this theory also cannot replace the technical standards of the Arm's Length Principle and the law of evidence. Compared with Dwindahany and Efendi (2024), Tambunan (2020), Zulfiqar (2023), this study supports the conclusion

DOI: 10.65246

that the quality of audits, comparables, FAR analysis, and evidence determines the outcome of disputes, but differs in that the decisive weaknesses were found in the Appellant's alternative rather than in the adjustment made by the Directorate General of Taxes. The contribution of this study lies in demonstrating that the tax authority's victory did not arise from a single dominant factor, but from a chain of acceptance: the legal basis legitimized the assessment, the benchmark design generated an indication of non-arm's-length conditions, and the inability of the opposing evidence to displace that design enabled the Judicial Panel to uphold the adjustment in full without independently reconstructing an alternative calculation.

b. Application of the Arm's Length Principle

The application of the Arm's Length Principle in this case began with the parties' agreement to use the Transactional Net Margin Method, with Operating Margin as the Profit Level Indicator, to assess the Power Generation Factory segment. Because neither the method nor the profit indicator was disputed, the disagreement shifted to the determination of PGF's economic characterization and its consequences for the expected level of profit. The Respondent characterized PGF as a contract manufacturer performing limited functions, conducting no or only limited research, design, and product development activities, and ordinarily earning a routine profit. The Appellant rejected this simplified characterization by emphasizing the operational functions, business decisions, and risks that it considered to be broader in scope. This disagreement was important because the functional and risk profile determines the expected profit and the suitability of comparable companies. The decision recorded this disagreement, but the Judicial Panel did not explicitly determine whether PGF was in fact a contract manufacturer; its assessment instead placed greater emphasis on the period and comparable data. HK.01 explained that FAR should be examined through business processes, contracts, deeds, organizational structures, and actual functions before the method and comparables are assessed. Accordingly, the TNMM was accepted as the testing framework, while PGF's characterization remained a disputed foundation that was not conclusively resolved in the final considerations.

The next disagreement concerned the temporal dimension and the construction of the set of comparables. The Appellant used data from 2015–2017 to explain fluctuations in performance while calculating PGF's average Operating Margin of 2 percent, which it considered to fall within the interquartile range of 1.66–9.11 percent generated from twelve comparable companies, with a median of 5.93 percent. The Respondent prioritized 2017 single-year data as the information most contemporaneous with the transactions under examination and retained four companies whose data were considered available. The remaining eight companies were excluded because their 2017 data were declared unavailable, while under the alternative multiple-year analysis, only those four companies were considered to have complete data. The Judicial Panel accepted the single-year approach, considered the Appellant's benchmark to contain extreme values and incomplete data, and stated that four companies were sufficient for the comparability analysis. The relationship between FAR, comparable data, and the period of comparison was also emphasized by HK.01:

> *"This is then linked to the comparable data and the comparable year. When the most appropriate method is an indirect method, we examine whether the comparable companies perform the same functions. Usually, the first priority is the function, and the second is the goods sold or services provided."*
>
> (HK.01, translated by the researcher)

However, the Appellant continued to challenge the elimination of the eight companies, while the decision did not provide individual reasons for accepting or rejecting each company. These choices directly affected the lower quartile, the median, and the position of PGF's margin. Therefore, period selection and comparable-company screening constituted two decisions that jointly shaped the outcome of the TNMM analysis.

After the period and comparables had been determined, the Respondent established a 2017 Operating Margin interquartile range of 6.55–9.04 percent, with a median of 7.49 percent. PGF's actual Operating Margin of negative 14.78 percent fell below the first quartile, leading the related-party transactions to be regarded as inconsistent with the arm's length principle. The Respondent then used the median to convert the operating loss of IDR 50,583,072,030 into an arm's-length operating profit

DOI: 10.65246

of IDR 25,636,808,506 on sales of IDR 342,280,487,403, resulting in a net income adjustment of IDR 76,219,880,536. The Appellant proposed two alternatives to this construction, namely the use of multiple-year data and a capacity adjustment to account for production utilization that it considered abnormal. However, the Judicial Panel did not calculate a capacity adjustment and upheld the outcome of the Respondent's benchmark. The importance of comparables and adjustments in transfer pricing disputes was further reflected in HK.01's explanation:

> *"What the parties usually disagree on is, first, the comparable data. Disputes over the comparable year are relatively rare. Actually, if the adjustment is correct, there is no problem. What is often disputed is the adjustment."*
>
> (HK.01, translated by the researcher)

HK.01 stated that after the comparables are selected, differences in functions, products, contracts, or economic conditions may require adjustments; according to HK.01, comparable data that are inconsistent with the FAR profile or inappropriate adjustments frequently become central points of dispute. The range therefore determined the existence of non-arm's-length conditions, while the median determined the amount of the adjustment. Nevertheless, the decision did not provide a specific explanation as to why the median was selected rather than another point within the range.

Overall, the application of the arm's length principle in this case formed a sequential process: the PGF segment was tested using the TNMM and Operating Margin, its business characterization was disputed, 2017 data were prioritized, twelve candidate comparables were screened down to four, an interquartile range was established, and the actual margin was subsequently adjusted to the median. The Appellant developed an alternative result based on a broader functional characterization, three-year data, twelve comparables, and a proposed capacity adjustment; the Respondent constructed the adjustment based on a contract manufacturer characterization, a single-year approach, four companies with data considered complete, and a median of 7.49 percent. The Judicial Panel accepted the Respondent's framework because it considered the single-year selection more appropriate, found that extreme and incomplete data reduced the reliability of the Appellant's benchmark, and determined that four companies were sufficient for the comparability analysis. Conflicting data remained apparent because the use of multiple-year data was not rejected in absolute terms, the Respondent also conducted an alternative three-year analysis, and the four companies used were drawn from the Appellant's documentation. HK.01's view emphasized that the alignment of the FAR profile, comparables, and adjustments determines the consistency of the profit method's results. Accordingly, the arm's-length profit adjustment did not arise automatically from the TNMM but was shaped by the choices concerning characterization, period, composition of comparables, adjustments, and the adjustment point ultimately accepted by the Judicial Panel.

The application of the Arm's Length Principle in this case demonstrates that the arm's-length level of profit was determined not solely by the selection of the TNMM, but by the methodological and evidentiary consistency among the FAR analysis, the determination of the tested party, and the availability of comparables. PER-43/PJ/2010 positions the comparability analysis, functional analysis, and method selection as a sequential process. Paragraphs 3.18–3.20 of the OECD Transfer Pricing Guidelines 2022 also require the tested party to be consistent with the functional analysis and to be selected as the party to which the method and comparables can be applied most reliably. Within the framework of Agency Theory, the different characterizations of PGF reflect the parties' respective interests in determining the allocation of functions, risks, and remuneration, whereas Information Asymmetry Theory explains why business processes, contracts, organizational structures, and internal data are decisive in verifying such characterization. HK.01 reinforced the view that the FAR analysis must be traced from the realities of the business processes before the comparables are assessed. This finding supports Tambunan and Abbas (2023), who position the FAR analysis as the foundation of comparability analysis, while extending their findings by demonstrating that the Judicial Panel may uphold an adjustment even when PGF's characterization is not explicitly formulated, provided that the Respondent's benchmark construction is considered more reliable.

The selection of the period and comparable

DOI: 10.65246

companies subsequently functioned as a mechanism for translating the economic characterization into a profit range. Paragraph 3.68 of the OECD Transfer Pricing Guidelines 2022 prioritizes contemporaneous independent transactions because they occur within the same economic environment, whereas paragraphs 3.75–3.79 position multiple-year data as a tool to be used when they provide added value, such as in examining losses, business cycles, or anomalies among the comparables. Accordingly, single-year and multiple-year data are not mutually exclusive choices, but serve different analytical functions. Articles 4 and 4A of PER-43/PJ/2010 in conjunction with PER-32/PJ/2011 require either the absence of material differences or the availability of reliable adjustments, while Annex Letter D of PMK 213/PMK.03/2016 requires the reasons for using a multiple-year analysis and the criteria for searching for comparables to be explained. From the perspective of information asymmetry, period selection and company screening constitute vulnerable points because each party possesses or emphasizes different information. This finding is consistent with Tambunan and Abbas (2023), but differs in its outcome: their study considered the analysis of the Directorate General of Taxes inadequate, whereas the Judicial Panel in this case accepted the single-year approach and four comparables because they were regarded as more complete and not distorted by extreme values. This difference confirms that the outcome of the Arm's Length Principle analysis depends on the quality of data operationalization rather than on an abstract preference for a particular period.

Adjustments, the interquartile range, and the median constitute distinct stages that determine both the reliability and the amount of the adjustment. Article 4 of PER-43/PJ/2010 in conjunction with PER-32/PJ/2011 recognizes adjustments intended to eliminate the effects of material differences, while paragraph 3.50 of the OECD Transfer Pricing Guidelines 2022 limits comparability adjustments to circumstances in which the differences are material, the data are of sufficient quality, and the adjustment method improves reliability. Because the Appellant's proposed capacity adjustment was not integrated into the calculation accepted by the Judicial Panel, the underutilization condition remained a commercial explanation rather than a quantitative factor capable of altering the benchmark. Once the set of comparables had been accepted, the interquartile range functioned to determine whether PGF's Operating Margin was within arm's-length conditions, whereas the median became the point used to construct the adjusted profit. Paragraph 3.62 of the OECD Guidelines permits the use of a measure of central tendency when unidentified or unquantifiable comparability defects remain, while paragraph 3.63 requires further examination of extreme results rather than their automatic elimination. This finding extends Turwanto et al. (2022), who demonstrate the use of profit-based methods in manufacturing disputes, by distinguishing the decision concerning the existence of non-arm's-length conditions from the decision concerning the amount of the adjustment. However, because the specific reason for using the median was not explained, the Judicial Panel's acceptance more closely reflected its acceptance of the Respondent's overall benchmark than a separate validation of each statistical stage.

Overall, the adjustment was upheld because each of the Respondent's technical choices formed a mutually reinforcing chain: the characterization supported the use of a one-sided method, the TNMM and Operating Margin converted the transactions into a profit indicator, the period determined the economic context, the comparables formed the distribution of results, the adjustment determined whether differences could be neutralized, and the median translated the non-arm's-length condition into the amount of the adjustment. Agency Theory explains why the parties had interests in selecting constructions that produced different levels of remuneration or tax burdens, but it cannot determine which technical choice was the most reliable. Information Asymmetry Theory more directly explains the importance of internal data, database completeness, and the verifiability of adjustments, although it still requires the Arm's Length Principle as the evaluative standard. This finding supports Permatasari and Husnasari (2022), who identify the Arm's Length Principle as the basis for redetermining profits from related-party transactions, while extending Tambunan and Abbas (2023) and Turwanto et al. (2022) by reconstructing how a sequence of methodological decisions produced an adjustment that remained sustainable before the court. The contribution of this study lies in distinguishing agreement on the method from the

DOI: 10.65246

reliability of its application: the TNMM did not automatically generate the adjustment, but acquired its corrective force only after the FAR analysis, period, comparables, adjustments, range, and adjustment point were accepted as a single construction that was more convincing than the Appellant's alternative.

c. Implications for Judicial Review

The findings for RQ 3 indicate that the scope for further evaluation primarily arises from parts of the considerations that were not specifically elaborated, rather than from the complete absence of a basis for the decision. The Appellant argued that the profit adjustment of IDR 76,219,880,536 was not attributed to specific related-party transactions, leaving the identity of the transactions, the adjustment amount, the counterparties, and their jurisdictions unclear for the purposes of a corresponding adjustment and the mutual agreement procedure. The Respondent focused its arguments primarily on testing the profit of the PGF segment using the TNMM, whereas the Judicial Panel constructed its ratio decidendi around the use of single-year data, extreme values, incomplete data, and the sufficiency of four comparables. The decision did not provide a separate response to the non-attribution argument. In addition, the determination of PGF's business characterization, the reasons for eliminating each of the eight companies, the specific response to the capacity adjustment, and the basis for selecting the median of 7.49% were not individually explained. Nevertheless, the Judicial Panel stated that the Respondent's comparability analysis was considered more consistent with the arm's length principle. Accordingly, the researcher's interpretation identifies these matters as points of traceability that require more thorough and careful examination, particularly regarding their materiality to the ruling and their direct relationship with the amount of the adjustment, rather than as evidence that the Judicial Panel committed a judicial error or disregarded all of the Appellant's arguments.

The evaluation of the decision is also constrained by the predominance of factual, methodological, and evidentiary issues. Disputes concerning whether PGF's loss was caused by underutilization, whether data for particular companies were available, whether a margin constituted an extreme value, and whether four companies were sufficiently comparable essentially require an assessment of documents and evidentiary weight. The Appellant relied on production-hours data, the break-even point, multiple-year performance, and twelve comparables to demonstrate that the 2017 loss could be commercially explained. The Respondent considered that the evidence was insufficient to alter the assessment outcome and retained four companies whose data were considered complete. The Judicial Panel accepted the Respondent's conclusion at an aggregate level but did not separately map the evaluation of each item of evidence and each company. The central role of evidentiary assessment in judicial decision-making was emphasized directly by HK.01:

> "*Judges look at the evidence: how valid it is, how sufficient it is, and how competent it is. As long as the evidence is competent, sufficient, and valid, that is what is used.*"
>
> (HK.01, translated by the researcher)

HK.01 explained that judges decide cases based on "laws and regulations" and the "sufficiency of evidence," while the parties' arguments serve as instruments to assist in understanding the dispute. HK.01's view is general in nature and does not explain the internal deliberative process of the Judicial Panel in this case. The researcher's interpretation indicates that the more an objection depends on reassessing the break-even point, comparable-company data, outliers, actual functions, or the reliability of an adjustment, the more dominant its factual and evidentiary character becomes in any further evaluation of the decision as a whole.

Overall, the considerations that may be evaluated further lie at the intersection between the Appellant's legal arguments and the factual construction accepted by the Judicial Panel. The Appellant emphasized the non-attribution of the adjustment, obstacles to the corresponding adjustment and the mutual agreement procedure, the unclear characterization of PGF, the use of twelve comparables, the underutilization condition, and the capacity adjustment. The Respondent maintained the limited-function characterization, the single-year approach, four comparables, the interquartile range, and the adjustment to the median. The Judicial Panel selected the Respondent's framework because it considered the 2017 data more appropriate, found

DOI: 10.65246

that extreme and incomplete data reduced the reliability of the Appellant's benchmark, and determined that four companies were sufficient; the ruling subsequently upheld the adjustment in full. This position is expressly reflected in the decision:

> *"The Tax Court concluded that the comparability analysis conducted by the Respondent better satisfied the arm's length principle; therefore, the Tax Court decided to uphold the Respondent's adjustment."*
>
> (Tax Court Decision, translated by the researcher)

HK.01 provided context that the dispute should be mapped from the FAR analysis through to the comparables and decided based on the applicable rules and evidence, but this explanation cannot be attributed to the Judicial Panel in this case. Conflicting data remained because multiple-year data were not rejected absolutely, the four comparables originated from the Appellant's documentation, and several analytical steps were not explained in detail. The researcher's interpretation is that any further evaluation must first distinguish questions concerning the application of legal norms from requests to reassess facts, examine the materiality of each issue to the ruling, and acknowledge that the dataset is not yet sufficient to responsibly formulate an actual Judicial Review strategy.

The implications of the decision for Judicial Review must first be assessed based on the exhaustive grounds stipulated in Article 91 of Tax Court Law, rather than solely on dissatisfaction with the extent of the Judicial Panel's considerations. Article 77 paragraph (3) provides for Judicial Review before the Supreme Court, while Article 91 letters d and e respectively require either the existence of a part of the claim that has not been decided without consideration of the underlying reasons or a decision that is manifestly inconsistent with the applicable laws and regulations. Article 84 paragraph (1) also requires a decision to contain an assessment of the evidence, the subject matter of the dispute, and the legal grounds. In this case, the normative dimension is primarily reflected in the argument concerning the non-attribution of the adjustment to specific related-party transactions, its implications for the corresponding adjustment and the mutual agreement procedure, and the traceability of the reasons for using the median and the capacity adjustment. However, the decision still contains a ratio decidendi concerning the single-year approach, data completeness, extreme values, and the sufficiency of four comparables. Therefore, the absence of a separate response to a particular argument does not automatically mean that a "part of the claim" was left undecided as referred to in Article 91 letter d. Such a qualification requires a substantive and measurable examination of the formulation of the petitum, the materiality of the argument to the ruling, and the causal relationship between the deficiency in the considerations and the adjustment that was upheld.

Conversely, most of the technical disputes in the decision predominantly concern the assessment of facts and evidence. The debates concerning the causes of PGF's losses, the level of underutilization, the break-even point, the availability of comparable-company data, the status of extreme values, and the reliability of the capacity adjustment depend on the evaluation of documents, calculations, and the evidentiary weight already examined by the Judicial Panel. Article 76 of Tax Court Law authorizes judges to determine what must be proven, the burden of proof, and the evidentiary value, while Article 78 establishes the assessment of evidence, tax regulations, and judicial conviction as the basis of the decision. This finding is consistent with Sitompul (2025), who positions Judicial Review not as a forum for repeating factual disputes by disguising them as judicial errors or errors of law. In this context, disagreement over the selection of four or twelve comparables does not, in itself, have a normative dimension. The legal dimension becomes more relevant when it can be demonstrated that the Judicial Panel applied a standard contrary to the applicable rules, failed to decide a material part of the claim, or constructed a conclusion unrelated to the facts it assessed. HK.01 also emphasized that judges may adopt different approaches when recalculating an adjustment, meaning that variations in judicial approaches do not automatically indicate a legal defect.

Agency Theory and Information Asymmetry Theory help explain the context of the dispute, but neither can serve as a legal ground for Judicial Review. Agency Theory explains the differences in interests between the authority and the Taxpayer in defending the adjustment or the reported profit, while Information Asymmetry Theory explains the unequal access to information concerning business processes, FAR, comparable data, and adjustment calculations. At the Judicial Review stage, the issue is

DOI: 10.65246

no longer merely which party controls the information, but whether that information has been assessed in accordance with the law and whether the decision satisfies the grounds stipulated in Article 91. The study by Suharsono et al. (2024), reinforces this limitation by demonstrating that the Supreme Court's reasoning in transfer pricing disputes emphasizes substantive truth, transactional evidence, functional analysis, positive law, and the Arm's Length Principle. The findings of this study extend that research by mapping the chain that must be examined: whether the FAR analysis supports the tested party, whether the comparables are consistent with the FAR analysis, whether differences in capacity require an adjustment, and whether the median has a traceable basis. Nevertheless, a break in one part of this chain must still be distinguished among weaknesses in technical arguments, deficiencies in the assessment of evidence, and manifest inconsistency with the applicable regulations; only grounds that can be linked to Article 91 are directly relevant to Judicial Review.

Article 91 letter b of Tax Court Law requires new evidence to be important and decisive, while Article 92 provides a three-month time limit whose commencement differs according to the ground for Judicial Review. Article 89 also limits an application to one submission and provides that its filing does not suspend the enforcement of the decision; the subsequent filing procedures are operationalized through Supreme Court Regulation Number 7 of 2018. This supports Gunawan and Mulyaningrum (2024), who argue that the evaluation of Judicial Review must comply with the statutory grounds and time limits. The contribution of this study is not to determine the likelihood of success, but to provide an initial classification: the non-attribution argument, the traceability of the median, and the scope of the considerations may be examined as normative issues, whereas the break-even point, outliers, company comparability, and evidentiary sufficiency are more closely related to factual matters. Because the dataset does not yet demonstrate the actual grounds for Judicial Review or the availability of new evidence, the strategy that can be formulated remains limited to examining the materiality of each issue, linking it strictly to Article 91, and avoiding the repackaging of differences in factual assessment as errors of law.

4.2. Research Discussion

4.2.1. Interpretation of Key Findings

The findings demonstrate that the DGT's victory cannot be attributed solely to the existence of a related-party relationship or to its statutory authority to conduct an arm's-length assessment. Rather, the outcome resulted from the interaction of legal legitimacy, methodological acceptance, and evidentiary sufficiency.

From a legal perspective, Article 18 paragraph (3) of the Income Tax Law (UU PPh) authorizes the DGT to redetermine income and deductions in accordance with the ALP, while Article 29 paragraph (1) of the General Provisions and Tax Procedures Law (UU KUP) establishes tax audits as an instrument for assessing compliance. This framework is reinforced by Article 20 paragraphs (1)–(4) of PER-43/PJ/2010, as amended by PER-32/PJ/2011, which addresses the application of transfer pricing provisions, consideration of the method and documentation prepared by the Taxpayer, and the treatment of adjustments when the ALP has been satisfied. The findings indicate that the legal authority provided the basis for the assessment, but the DGT prevailed only after the Judicial Panel accepted its technical application.

This interpretation is consistent with Tambunan (2020), who found that the DGT was more reliable when complying with technical and administrative provisions. The present study extends that finding by demonstrating how legal authority, methodological quality, evidentiary assessment, and judicial acceptance interacted to produce a decision upholding the adjustment in full.

From a technical-methodological perspective, Paragraph 2.74 of the OECD Transfer Pricing Guidelines 2022 emphasizes the importance of robust comparability analysis in applying the TNMM. Paragraph 3.68 prioritizes contemporaneous comparable data because they reflect the economic environment of the relevant period, whereas paragraphs 3.75–3.76 recognize the usefulness of multiple-year data when they provide additional explanatory value, particularly in understanding losses. Paragraph 3.63 cautions against eliminating extreme results solely because they differ from other observations, while paragraph 3.62 permits the use of the median when unidentified or unquantifiable comparability defects remain.

Accordingly, the DGT's victory was not an automatic consequence of using single-year data,

DOI: 10.65246

four comparable companies, or the median. It resulted from the Judicial Panel's assessment that the Respondent's overall benchmark was more reliable than the Appellant's alternative. This finding supports Zulfiqar (2023), who emphasized the importance of documentation, comparable data, and comparability methods. However, whereas previous research was dominated by decisions favorable to Taxpayers, this case demonstrates circumstances in which the DGT's selection of comparable companies and construction of the arm's-length range were accepted as the basis for the adjustment.

The evidentiary dimension further explains why the Appellant's commercial explanation did not change the outcome. Article 76 of the Tax Court Law governs the determination of matters to be proven, the burden of proof, and the assessment of evidence, while Article 78 requires the decision to be based on the evidence, applicable tax laws and regulations, and judicial conviction. Under Information Asymmetry Theory, the Taxpayer's access to internal information concerning production capacity, costs, and operational losses does not automatically establish the validity of an alternative benchmark. Such information must be translated into verifiable evidence and replicable calculations.

The proposed capacity adjustment also needed to satisfy paragraph 3.50 of the OECD Transfer Pricing Guidelines 2022, which requires material differences, sufficiently reliable data, and an improvement in the reliability of the comparison. This finding extends Dwindahany and Efendi (2024), who identified weaknesses in substantiating tax adjustments as a source of disputes. In the present case, the evidentiary burden operated in both directions: the DGT had to substantiate its benchmark, while the Appellant had to establish the causal relationship between underutilization and the loss and translate that relationship into an appropriate adjustment. The Judicial Panel accepted the DGT's calculation because the Appellant's alternative did not achieve sufficient evidentiary strength.

4.2.2. Comparison with Previous Studies

The findings support several earlier studies while identifying differences in the direction and outcome of the dispute.

First, the importance of FAR analysis and comparability is consistent with Tambunan and Abbas (2023), who position functional analysis as the foundation of comparability analysis. In the present case, the business characterization of PGF remained disputed and was not explicitly determined by the Judicial Panel. Nevertheless, the DGT's benchmark was upheld because the overall comparability analysis was considered more reliable. This extends previous research by showing that a court may accept a transfer pricing adjustment even when the tested party's characterization is not explicitly formulated in the final considerations.

Second, the findings concerning period selection and comparable-company screening are consistent with Tambunan and Abbas (2023). However, the outcome differs from their study, which considered the DGT's analysis inadequate. In this case, the Judicial Panel accepted the single-year approach and four comparable companies because the Appellant's benchmark was considered to contain extreme values and incomplete data. This difference indicates that the outcome of ALP analysis depends on the quality of the data and their operationalization rather than on an abstract preference for single-year or multiple-year data.

Third, the findings concerning profit-based methods in manufacturing disputes are consistent with Turwanto et al. (2022). The present study extends that research by distinguishing between determining whether the observed margin falls outside the arm's-length range and determining the amount of the resulting adjustment. Although the range indicated non-arm's-length conditions, the median of 7.49% determined the amount used to reconstruct the operating profit. The specific reason for selecting the median, however, was not separately explained in the decision.

The findings also support Permatasari and Husnasari (2022), who identify the ALP as the basis for redetermining profits arising from related-party transactions. They further extend Tambunan and Abbas (2023) and Turwanto et al. (2022) by reconstructing how the tested-party characterization, period selection, comparable-company composition, and adjustment point combined to produce a benchmark that was upheld by the Tax Court.

Regarding Judicial Review, the findings are consistent with Sitompul (2025), who argues that Judicial Review should not be used to repeat factual disputes by presenting them as judicial errors or errors of law. They also support Suharsono et al.

DOI: 10.65246

(2024), whose study demonstrates that Supreme Court reasoning in transfer pricing disputes emphasizes substantive truth, transactional evidence, functional analysis, positive law, and the ALP. Furthermore, the statutory constraints identified in this case support Gunawan and Mulyaningrum (2024), who emphasize compliance with the legal grounds and time limits governing Judicial Review.

Overall, the present study contributes a case-specific explanation of why a transfer pricing adjustment may be upheld despite unresolved methodological questions. It also clarifies why the existence of analytical weaknesses does not, by itself, establish that the statutory requirements for Judicial Review have been satisfied.

4.2.3. Theoretical Contributions

The theoretical contribution of this study lies in integrating Agency Theory, Information Asymmetry Theory, the ALP, and evidentiary analysis to explain the outcome of a transfer pricing dispute.

Agency Theory explains the divergent interests of the Taxpayer and the tax authority. The DGT seeks to ensure that taxable income reflects arm's-length conditions, while the Taxpayer challenges an adjustment that increases its taxable income. These competing interests help explain why the parties advance different characterizations of the tested segment, comparable-company sets, and profitability benchmarks. However, Agency Theory cannot independently establish whether the single-year approach, the selection of four comparable companies, or the use of the median is technically appropriate.

Information Asymmetry Theory provides a more direct explanation of the dispute because its central issues concern the completeness, comparability, and verifiability of information. The Taxpayer possesses internal information about production capacity, costs, and business processes, while the tax authority relies on available documentation and comparable-company data. The findings show that informational advantage becomes persuasive only when the evidence can be verified and translated into a defensible calculation. Nevertheless, Information Asymmetry Theory cannot replace the technical standards of the ALP or the legal rules governing evidence.

The study therefore demonstrates that neither theory alone explains the judicial outcome. Agency Theory explains the monitoring incentives and competing interests, while Information Asymmetry Theory explains the importance of information quality. The ALP provides the technical standard for evaluating related-party transactions, and evidentiary rules determine whether the competing arguments are sufficiently substantiated. Their integration explains the sequence through which legal authority legitimized the assessment, the methodological framework constructed the benchmark, and the insufficiency of the Appellant's alternative enabled the adjustment to be upheld in full.

4.2.4. Practical and Policy Implications

The findings have several implications for tax authorities, taxpayers, and judicial institutions.

For tax authorities, transfer pricing adjustments should be supported by a transparent chain of reasoning linking the tested party's functional profile, the selected method, the comparable-company criteria, the financial period, and the resulting adjustment. Although the DGT prevailed in this case, the lack of detailed explanations concerning individual comparable companies and the selection of the median indicates the importance of documenting each analytical stage.

For taxpayers, Transfer Pricing Documentation should be supported by complete financial data, evidence of actual business processes, and a clear explanation of the relationship between operational conditions and financial outcomes. Claims concerning underutilization, losses, or production constraints should be supported by verifiable evidence and, where appropriate, a reproducible capacity-adjustment calculation. Merely demonstrating that production fell below the break-even point does not establish that the reported loss should be accepted as an arm's-length outcome.

For judicial institutions, decisions would benefit from explicitly connecting the tested party's functional characterization, the treatment of comparable companies, the evidentiary assessment, and the final adjustment. Where an argument is rejected, a clear explanation of its materiality and evidentiary deficiencies would improve the traceability of judicial reasoning and reduce uncertainty regarding the basis of the decision.

For policymakers and practitioners, the findings underscore the importance of consistent

DOI: 10.65246

transfer pricing documentation and clear guidance on comparability analysis. They also demonstrate that potential Judicial Review requires a careful distinction between factual disagreements and legal grounds recognized by the applicable legislation.

4.2.5. Integration with the Research Gap

The study addresses the research gap identified in the literature review by integrating transfer pricing methodology, evidentiary assessment, and judicial reasoning within a single case-based analysis. Rather than examining the TNMM solely as a technical method, it demonstrates how the method's application interacts with the legal authority of the tax administration and the quality of evidence presented in court.

The findings show that agreement on the TNMM did not eliminate disagreement over the tested party, the period of analysis, the selection of comparable companies, or the adjustment point. These decisions collectively shaped the outcome of the dispute. Furthermore, the study distinguishes between weaknesses in the methodological explanation and legal deficiencies that may be relevant to Judicial Review.

This integrated approach extends previous research by demonstrating that the outcome of a transfer pricing dispute depends not only on the formal availability of a legal basis or the selection of an accepted method but also on whether the resulting benchmark is supported by sufficiently reliable evidence and accepted through judicial reasoning. The study's contribution is therefore a structured explanation of the relationship between technical methodology, evidentiary sufficiency, and the legal evaluation of a tax dispute.

4.2.6. Acknowledgement of Study Limitations

The findings should be interpreted in light of the study's focus on a single transfer pricing dispute and its reliance on one key informant. The case provides detailed insight into the application of the TNMM, comparability analysis, and evidentiary assessment, but it does not establish that the same judicial approach will be adopted in other disputes. In addition, several methodological steps in the decision were not explained in detail, including the characterization of PGF, the individual exclusion of comparable companies, and the selection of the median. These omissions limit the extent to which the researcher can reconstruct the Judicial Panel's complete reasoning.

The study also does not establish whether subsequent Judicial Review proceedings occurred, whether new evidence became available, or what the outcome of any such proceedings might have been. Accordingly, the discussion identifies potential legal issues for further examination rather than predicting the likelihood of success in Judicial Review. Future research could compare multiple transfer pricing decisions, examine a broader range of judicial perspectives, and analyze the relationship between technical transfer pricing assessments and the statutory grounds for Judicial Review.

5. Conclusion

5.1. Summary of Key Findings

Based on the research findings and discussion, the victory of the Directorate General of Taxes in the transfer pricing dispute under Tax Court Decision Number PUT-014532.15/2020/PP/M.XIIIB of 2022 resulted from the interrelationship among legal, technical-methodological, and evidentiary aspects. The normative authority of the Directorate General of Taxes provided the basis for reassessing the arm's-length nature of related-party transactions, while the Judicial Panel's acceptance of the TNMM, Operating Margin, 2017 single-year data, four comparable companies, the interquartile range, and the median led the benchmark of the Directorate General of Taxes to be regarded as more reliable than the Appellant's alternative. The explanation concerning underutilization and commercial losses was not successfully operationalized into a measurable capacity adjustment supported by an adequate causal relationship and therefore could not displace the Respondent's calculation, resulting in the arm's-length profit adjustment of IDR 76.21 billion being upheld in full.

5.2. Theoretical Contributions

Empirically, this study demonstrates that the victory of a tax authority is not determined by a single factor, but through a chain of judicial acceptance linking normative authority, benchmark construction, evidentiary quality, and the final ruling. These findings demonstrate that the application of the Arm's Length Principle does not end with the selection of a method, but is shaped through a sequence involving FAR analysis, functional

DOI: 10.65246

characterization, period selection, comparable-company selection, comparability adjustments, range construction, and the determination of the adjustment point. This study also clarifies the distinction between agreement on the method and the reliability of its application and provides an initial classification between legal and factual issues in evaluating Judicial Review.

5.3 Practical and Policy Implications

In practical terms, the findings may serve as a reference for Taxpayers in linking commercial explanations with causal evidence and measurable adjustments, for the Directorate General of Taxes in improving the traceability of FAR analysis, comparable-company selection, and the use of the median, and for parties evaluating Judicial Review in assessing the materiality of each issue based on the legal grounds prescribed in the applicable laws and regulations.

In the context of Judicial Review, the evaluation should be directed toward the traceability of the considerations concerning the non-attribution of the adjustment, corresponding adjustment and mutual agreement procedure mechanisms, PGF's characterization, and the basis for using the median. However, the relevance of each issue must still be assessed by distinguishing questions concerning the application of law from differences in the assessment of facts and evidence and by linking them strictly to the grounds for Judicial Review under Article 91 of the Tax Court Law.

5.4. Limitations of the Study

This study has inherent scope limitations arising from the use of an in-depth case study design. The analysis focused on a single Tax Court decision to allow the relationship among the parties' arguments, methodological choices, evidence, and the Judicial Panel's considerations to be examined in detail. Therefore, the findings are contextual to the characteristics of the PT X case and should be applied cautiously to transfer pricing disputes involving different transaction profiles, business characteristics, and evidentiary conditions. The interview with a Tax Court Judge was used to enrich the understanding of evidentiary practices and judicial considerations in general, without directly attributing the informant's views to the Judicial Panel that decided the case.

5.5. Directions for Future Research

This scope provides opportunities for future research to conduct comparative studies of multiple transfer pricing decisions and to incorporate more diverse perspectives from tax authorities, Taxpayers, tax consultants, academics, and legal practitioners.

6. References

Aisya, S., & Nuryanah, S. (2024). Tax Dispute Analysis of Transfer Pricing Case Study of Tax Court Decisions 2020-2023. *JMKSP (Jurnal Manajemen, Kepemimpinan, Dan Supervisi Pendidikan)*, *9*(1). https://doi.org/10.31851/jmksp.v9i1.14343

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. In *Quarterly Journal of Economics* (Vol. 84, Number 3). https://doi.org/10.2307/1879431

Arabyan, K. (2023). Applicability of the Methods Used in Determining Income (Profi t, Proceeds) for Tax Purposes in Transactions with Related Parties. *Auditor*, *9*(1). https://doi.org/10.12737/1998-0701-2023-9-1-40-44

Arrisman, A., & Fauzi, A. (2025). Analisis Yuridis Pembuktian Sengketa Transfer Pricing. *Syntax Literate ; Jurnal Ilmiah Indonesia*, *10*(2). https://doi.org/10.36418/syntax-literate.v10i2.56912

Bartelsman, E. J., & Beetsma, R. M. W. J. (2003). Why pay more? Corporate tax avoidance through transfer pricing in OECD countries. *Journal of Public Economics*, *87*(9–10). https://doi.org/10.1016/S0047-2727(02)00018-X

Barton, A., Subramaniam, V., & Marino, P. (2021). Corporate Tax Planning: Impact of COVID-19 and Transfer Pricing: Approaches for Comparability Adjustments. *Canadian Tax Journal/Revue Fiscale Canadienne*, *69*(3). https://doi.org/10.32721/ctj.2021.69.3.ctp

Berdonce, J. G., Chand, V., & Ditz, X. (2021). Answering Eight Key Questions about Transfer Pricing during the COVID-19 Era. *International Transfer Pricing Journal*, *2021*(4). https://doi.org/10.59403/34dzkz0

Beuselinck, C., & Pierk, J. (2024). On the dynamics between local and international tax planning in multinational corporations. *Review of Accounting Studies*, *29*(1). https://doi.org/10.1007/s11142-022-09731-y

Braun, V., & Clarke, V. (n.d.). *Using thematic analysis in psychology*.

Brychta, K., De Abreu, M. C., Hudenko, J., Santos, L. S., & Do Valle Poubel, L. C. (2024). The concept of associated persons as a key and

DOI: 10.65246

potentially problematic aspect in transfer pricing. *Financial Internet Quarterly*, *19*(1). https://doi.org/10.2478/fiqf-2023-0003

Chyzhykova, O. (2019). Using statistical techniques for choosing the calculation methodology of arm's length range in the process of transfer pricing documents preparation. *Economy and Entrepreneurship*, (42). https://doi.org/10.33111/ee.2019.42.chyzhykovao

Davies, R. B., Martin, J., Parenti, M., & Toubal, F. (2018). Knocking on Tax Haven's Door: Multinational Firms and Transfer Pricing. *Review of Economics and Statistics*, *100*(1). https://doi.org/10.1162/REST_a_00673

De Robertis, G., & Kondi, D. (2022). How to improve the estimation of intra-firm prices for financial and tax reporting during economic downturns. *Transnational Corporations Review*, *14*(3). https://doi.org/10.1080/19186444.2022.2099216

Dessatniuk, O. (2015). Practical application of the methodology for calculating the arm's length range of prices/profitability for transfer pricing purposes. *European Journal of Economics and Management Sciences*. https://doi.org/10.20534/ejems-15-4-30-33

Dwindahany, C. S., & Efendi, S. (2024). Should Excessive Marketing Expenses Be Remunerated? Lessons From Indonesia's Tax Court Decisions. *Jurisdictie: Jurnal Hukum Dan Syariah*, *15*(1). https://doi.org/10.18860/j.v15i1.26915

Eisenhardt, K. M. (1989). Agency Theory: An Assessment and Review. *Academy of Management Review*, *14*(1). https://doi.org/10.5465/amr.1989.4279003

Ganesh Lal Meena, Neetu Zora. (2025). Qualitative Review of Transfer Pricing Disputes in India. *Journal of Informatics Education and Research*, *5*(1). https://doi.org/10.52783/jier.v5i1.1980

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, *3*(4). https://doi.org/10.1016/0304-405X(76)90026-X

Keen, M., & Slemrod, J. (2017). Optimal tax administration. *Journal of Public Economics*, *152*. https://doi.org/10.1016/j.jpubeco.2017.04.006

Khris, B., & Whiteside, M. (2020). Transfer Pricing: Purpose of Determination and Factors Affecting Transfer Pricing Determination. *Journal Dimensie Management and Public Sector*, *1*(2). https://doi.org/10.48173/jdmps.v1i2.48

Kleven, H., Knudsen, M., Kreiner, C., Pedersen, S., & Saez, E. (2011). Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark. *Econometrica*, *79*(3). https://doi.org/10.3982/ecta9113

Mahmudi, S. (2022). Factors and Functions Affecting Transfer Pricing. *International Journal of Academic Research in Business and Social Sciences*, *12*(8). https://doi.org/10.6007/ijarbss/v12-i8/14518

Martensen, F. (2020). Rethinking Interquartile Ranges. *International Transfer Pricing Journal*, *2020*(1). https://doi.org/10.59403/17st9xw

Matavulj, L. (2021). Calculation of transfer prices: Theoretical and practical implications on the result. *Ekonomika*, *67*(1). https://doi.org/10.5937/ekonomika2101059m

Melega, A., Molociniuc Hritcan, M., & Macovei, A. G. (2022). The transfer pricing file and tax controversies - an empirical study. *Studies and Scientific Researches. Economics Edition*, (36). https://doi.org/10.29358/sceco.v0i36.515

Muhammadi, A. H., Ahmed, Z., & Habib, A. (2016). Multinational transfer pricing of intangible assets: Indonesian tax auditors' perspectives. *Asian Review of Accounting*, *24*(3). https://doi.org/10.1108/ARA-10-2014-0112

Murphy, R., Janský, P., & Shah, A. (2019). BEPS Policy Failure—The Case of EU Country-By-Country Reporting 1 . *Nordic Tax Journal*, *2019*(1). https://doi.org/10.1515/ntaxj-2019-0005

Muyaa, E. (2014). Transfer Pricing Comparability Adjustments: The Pursuit of "Exact" Comparables. *International Transfer Pricing Journal*, *21*(5). https://doi.org/10.59403/1zwphc9

Muzychuk, M. I., & Fomina, O. V. (2021). Functions, assets and risk analysis in the preparation of transfer pricing documentation. *Universal Journal of Accounting and Finance*, *9*(5). https://doi.org/10.13189/ujaf.2021.090505

Mwape, V., Mwange, A., & Hamalengwa, M. (2025). Reassessing the Arm's Length Principle in Transfer Pricing Regulation: A Doctrinal and Practical Analysis from a Zambian Perspective. *African Journal of Commercial Studies*, *6*(1). https://doi.org/10.59413/ajocs/v6.i.1.18

Ökten, R., Schoorl, D., & Kleimann, M. (2025). Comparability Analysis in Practice: Recent Developments from Transfer Pricing Audits and Court Cases. *International Transfer Pricing Journal*, *32*(4). https://doi.org/10.59403/vmean5

Parmalia, I., & Rosid, A. (2023). Analysis Of Tax Disputes On Loans From Shareholders: A Case Study On Tax Court Rulings For The

DOI: 10.65246

Period 2018-2022. *Journal Research of Social Science, Economics, and Management*, *3*(3). https://doi.org/10.59141/jrssem.v3i3.565

Patel, V. T., Dhoot, S., & Savla, K. P. (2023). More transfer pricing legislation rounds through case law. *International Transfer Pricing Journal*, *2023*(5). https://doi.org/10.59403/1f7xz8p

Pelekis, D. (2021). The burden and standard of proof in the tax ruling cases: A practical limit to the eu's arm's length principle? *Journal of European Competition Law and Practice*, *12*(9). https://doi.org/10.1093/jeclap/lpab038

Permatasari, M. I., & Husnasari, F. A. (2022). Penerapan *Arm's Length Principle* Dalam Praktek *Abuse of Transfer Pricing* Perusahaan Multinasional di Indonesia. *Rechtidee*, *17*(2). https://doi.org/10.21107/ri.v17i2.5245

Prayudi, I. C., & Rosid, A. (2025). Tax Disputes Stemming from the Use of Quantitative Techniques in Audits: Implications of Numerical Analysis, Correlation Testing, and Equalization Methods. *E-Jurnal Akuntansi*, *35*(3). https://doi.org/10.24843/eja.2025.v35.i03.p20

Qadri. (2022). Contextualizing the Universe of PSAK 74 [IFRS 17] in Indonesia and Insurance Industry Vigilance Through the Cosmology of "Yoga Kshema." *Journal of Accounting and Strategic Finance*, *5*(1).

Qadri, R. A., & Fatmawati, F. (2021). Storytelling Case Study: How To Win a Tax Dispute Against Tax Authority? *Jurnal Pajak Dan Keuangan Negara (PKN)*, *2*(2). https://doi.org/10.31092/jpkn.v2i2.1118

Sebele-Mpofu, F. Y., Mashiri, E., & Korera, P. (2025). Transfer Pricing Audit Challenges and Dispute Resolution Effectiveness in Developing Countries with Specific Focus on Zimbabwe. *Accounting, Economics and Law: A Convivium*, *15*(2). https://doi.org/10.1515/ael-2021-0026

Sitompul, A. (2025). Consideration of Judges in PPh Correction Disputes Article 23: Review of PK Decision Number 5360/B/PK/CHD/2024. *International Asia Of Law and Money Laundering (IAML)*, *4*(4), 117–123. https://doi.org/10.59712/iaml.v4i4.144

Steens, B., Roques, T., Gonnet, S., Beuselinck, C., & Petutschnig, M. (2022). Transfer pricing comparables: Preferring a close neighbor over a far-away peer? *Journal of International Accounting, Auditing and Taxation*, *47*. https://doi.org/10.1016/j.intaccaudtax.2022.100471

Suharsono, A., Prasetyoningsih, N., & Nur Aida Ikrima. (2024). A Comparison of Supreme Court Justices' Legal Argumentation in Deciding Disputes on Transfer Pricing Management Fees in Indonesia. *Jurnal Jurisprudence*. https://doi.org/10.23917/jurisprudence.v14i2.5874

Sukarno, S. (2022). Apakah Transfer Pricing Documentation Meningkatkan Kepatuhan Pajak? *Jurnal Pajak Dan Keuangan Negara (PKN)*, *4*(1S). https://doi.org/10.31092/jpkn.v4i1s.1864

Sulik-Górecka, A. (2018). Dilemmas of Transfer Pricing Comparability Analysis in Manufacturing Entities. Polish-Czech Case Study. *Management Systems in Production Engineering*, *26*(2). https://doi.org/10.2478/mspe-2018-0012

Tambunan, M. R. U. D. (2020). A note to transfer pricing audit performance by Indonesia's tax administration based on tax court decision 2015-2019 on manufacturing industry. *Jurnal Akuntansi & Auditing Indonesia*, *24*(2). https://doi.org/10.20885/jaai.vol24.iss2.art2

Tambunan, R., & Abbas, Y. (2023). Evaluasi Implementasi Analisis Kesebandingan atas Sengketa Transfer Pricing PT OCI. *Owner*, *7*(4). https://doi.org/10.33395/owner.v7i4.1735

Turwanto, T., Primasari, K., & Firmansyah, A. (2022). Penghindaran Pajak Melalui Transfer Pricing Pada Perusahaan Di Indonesia: Analisis Isi Atas Risalah Putusan Pengadilan Pajak. *Educoretax*, *2*(1). https://doi.org/10.54957/educoretax.v2i1.158

Udoayang, J., Akpanuko, E., & Asuquo, A. (2010). Multinational transfer pricing and international taxation: what, why, how and reporting challenges. *African Research Review*, *3*(5). https://doi.org/10.4314/afrrev.v3i5.51150

Zulfiqar, F. L., Ulupui, I. G. K. A., & Respati, D. K. (2023). A qualitative analysis on transfer pricing tax audit performance in Indonesia. *AKURASI: Jurnal Riset Akuntansi Dan Keuangan*, *5*(1). https://doi.org/10.36407/akurasi.v5i1.805