LEGALITY WITHOUT CONSENT: SENTIMENT ON X AND THE DISCURSIVE LEGITIMACY OF EDUCATION BUDGET RATIONALISATION FOR THE FREE NUTRITIOUS MEAL PROGRAMME IN THE 2026 STATE BUDGET

Ega Nandana Rafif[1], Yuwanto[2]

Email: rafifnandana12345@gmail

Departemen Politik dan Ilmu Pemerintahan

Fakultas Ilmu Sosial dan Ilmu Politik Universitas Diponegoro

Jl. Prof. H. Soedarto, S.H Tembalang, Semarang Kode Pos 1269

Telepon (024) 7465407 Faksimile (024) 7465405

Laman: https: //www.fisip.undip.ac.id Email: fisip@undip.ac.id

ABSTRAK

Kebijakan publik dapat memenuhi seluruh persyaratan legalitas dan pada saat yang sama gagal memperoleh persetujuan publik yang oleh teori legitimasi diperlakukan sebagai unsur konstitutif dari kekuasaan yang sah. Penelitian ini menelaah kesenjangan tersebut melalui kasus APBN 2026, ketika Rp223,5 triliun alokasi Badan Gizi Nasional untuk Program Makan Bergizi Gratis (MBG) diklasifikasikan ke dalam fungsi pendidikan yang dimandatkan konstitusi sebesar 20 persen — sebuah pengaturan yang berdasar hukum pada Undang-Undang Nomor 17 Tahun 2025 namun kini menjadi objek permohonan uji materi di Mahkamah Konstitusi. Penelitian menganalisis 2.037 unggahan berbahasa Indonesia di media sosial X pada Januari 2025 hingga Mei 2026, diklasifikasikan menggunakan InSet Lexicon dan ditafsirkan melalui kerangka persetujuan Beetham (1991) dan legitimasi diskursif Dryzek (2000). Sentimen negatif mencakup 66,76 persen korpus dan mempertahankan status mayoritas pada setiap bulan pengamatan, menguat alih-alih melemah seiring membesarnya volume wacana. Analisis frekuensi kata mengidentifikasi empat klaster tematik, dengan temuan terpenting bahwa kritik tertuju pada mekanisme pembiayaan dan bukan pada tujuan gizi program. Uji validasi terhadap pengodean manual menghasilkan Cohen's Kappa sebesar 0,0647 yang didorong hampir seluruhnya oleh kesalahan klasifikasi kritik ironis sebagai sentimen positif; hasil ini dilaporkan sebagai temuan substantif mengenai keandalan klasifikasi berbasis leksikon pada wacana politik berbahasa Indonesia. Penelitian menyimpulkan adanya defisit legitimasi diskursif dalam wacana yang diteliti, dan bahwa ketahanan pola tersebut lintas waktu lebih bermakna secara analitis daripada besarannya pada satu momen.

Kata kunci: legitimasi kebijakan, persetujuan publik, analisis sentimen, ruang publik digital, politik anggaran, APBN 2026

ABSTRACT

Public policy can satisfy every requirement of legality and still fail to secure the expressed consent that legitimacy theory treats as constitutive of legitimate rule. This study examines that gap through Indonesia's 2026 State Budget, in which Rp223.5 trillion of the National Nutrition Agency's allocation for the Free Nutritious Meal Programme (MBG) was classified under the constitutionally mandated twenty per cent education budget — an arrangement legally grounded in Law Number 17 of 2025 yet currently the subject of a judicial review petition before the Constitutional Court. The study analyses 2,037 Indonesian-language posts published on X between January 2025 and May 2026, classified using the InSet Lexicon and interpreted through Beetham's (1991) account of consent and Dryzek's (2000) account of discursive legitimacy. Negative sentiment accounts for 66.76 per cent of the corpus and retains majority status in every month of observation, intensifying rather than moderating as discourse volume expands. Word-frequency analysis identifies four thematic clusters, of which the most consequential shows that criticism targets the financing mechanism rather than the nutritional objectives of the programme. Validation against manual coding returned a Cohen's Kappa of 0.0647, driven overwhelmingly by the misclassification of ironic criticism as positive sentiment; this result is reported as a substantive finding concerning the reliability of lexicon-based classification in politically charged Indonesian discourse rather than as a disclaimer. The study concludes that the observed pattern constitutes a discursive legitimacy deficit within the discourse examined, and that its persistence across time is more analytically significant than its magnitude at any single moment.

Keywords: *policy legitimacy, public consent, sentiment analysis, digital public sphere, budget politics, 2026 State Budget*

INTRODUCTION

A government decision may be enacted through every procedure the law requires and still be experienced by those it governs as something imposed rather than agreed. Legitimacy theory has long insisted that this distinction matters. In Beetham's (1991) account, power is legitimate only where three conditions hold together: it conforms to established rules, those rules are justifiable by reference to beliefs shared between dominant and subordinate, and there is evidence of express consent on the part of the subordinate. The three conditions are not interchangeable. A policy may satisfy the first and fail the third, and when it does, the resulting condition is not illegality but a legitimacy deficit — a gap between what the state is entitled to do and what the governed are prepared to endorse.

Over the past decade, much research has focused on where such consent can now be observed. The migration of political argument onto digital platforms has produced a large body of work treating social media discourse as a readable trace of public opinion, and a parallel body of work in computational text analysis has supplied the instruments for reading it at scale (Grimmer & Stewart, 2013; Liu, 2015; Pak & Paroubek, 2010). In Indonesia, this has produced a substantial literature applying sentiment classification to contested national policies, including fuel subsidy removal, the Omnibus Law on Job Creation, and pandemic vaccination programmes. That literature has established, repeatedly and consistently, that Indonesian citizens use platforms such as X to evaluate government decisions in evaluatively explicit terms.

It remains unclear, however, what those evaluations amount to. Two gaps are relevant here. The first is interpretive. The prevailing convention in this literature is to report a distribution — so many per cent negative, so many per cent positive — and to stop there. Sentiment is measured but not theorised; the proportion is presented as a finding in itself rather than as evidence bearing on any prior conceptual question. What a majority-negative distribution means for the standing of the policy it concerns is generally left unaddressed. The second gap is methodological. Lexicon-based classification is widely used in Indonesian-language studies because it requires no labelled training data and is computationally transparent, yet studies employing it frequently omit any validation of the classifications against human judgement. The reliability of the numbers circulating in this literature has therefore not been systematically tested, even as those numbers accumulate into an apparent consensus about Indonesian public opinion.

A third gap is substantive and case-specific. In the 2026 State Budget, established by Law Number 17 of 2025, the National Nutrition Agency (Badan Gizi Nasional, BGN) received a base allocation of Rp268 trillion — the largest of any Indonesian ministry or agency, exceeding both the Ministry of Defence (Rp187.10 trillion) and the National

Police (Rp146.05 trillion). Of that allocation, 83.4 per cent, amounting to Rp223.5 trillion, was classified under the education function of the budget, and therefore counted toward the twenty per cent of state expenditure that Article 31 paragraph (4) of the 1945 Constitution reserves for education. The legal basis for this classification is Article 22 paragraph (3) of Law Number 17 of 2025, which brings nutritional provision at educational institutions within the scope of operational funding for educational service delivery. The total education budget in the 2026 State Budget amounts to approximately Rp763 trillion. The arrangement provoked immediate contestation in the legislature and, in February 2026, a judicial review petition before the Constitutional Court brought by a coalition that included honorary teachers. To date this controversy has not been examined through the analytical lens of public sentiment situated within a framework of policy legitimacy.

The purpose of this study is to determine what public discourse on X reveals about the consent dimension of legitimacy in this case, and what the attempt to measure it reveals about the instruments available for doing so. Two questions are addressed. First, what is the distribution, thematic composition, and temporal trajectory of sentiment on X toward the classification of MBG expenditure within the education budget? Second, what does that pattern indicate about the discursive legitimacy of the policy, understood in Beetham's (1991) and Dryzek's (2000) terms?

The study makes two contributions. Theoretically, it operationalises the consent dimension of legitimacy as something empirically observable in naturally occurring discourse, rather than treating legitimacy as a condition to be asserted or inferred. Methodologically, it reports and analyses the failure of lexicon-based classification on this corpus, and argues that this failure is systematic rather than incidental — a property of the encounter between word-level scoring and ironic political speech, and therefore a finding with implications extending beyond the present case.

Legitimacy, Consent, and the Limits of Legality

Beetham's (1991) reconstruction of legitimacy was formulated in explicit contrast to the Weberian tradition, which he argued had collapsed legitimacy into belief — into what people happen to think about the authority they live under. Against this, Beetham proposed that legitimacy is a property of the relationship between power holders and the governed, assessable along three distinct dimensions: conformity to established rules; the justifiability of those rules in terms of beliefs shared by both parties; and the express consent of the subordinate, demonstrated through actions that publicly signal endorsement. The framework's analytical value lies in the independence of the dimensions. A regime or a policy may hold on one and fail on another, and the specific pattern of holding and failing tells us

something that an undifferentiated judgement of legitimacy cannot.

The present case is well suited to that framework precisely because the first dimension is not seriously in dispute. The classification of MBG expenditure under the education function was enacted through statute, elaborated by Presidential Regulation Number 118 of 2025, and defended in the legislature as the product of an agreement between government and parliament. Whether it is constitutionally sound is the question now before the Constitutional Court; that it was legally enacted is not contested. The third dimension is a separate matter, and it is the one this study is equipped to examine. Whether a policy that is legally grounded also enjoys expressed consent is an empirical question, and one that leaves observable traces wherever citizens publicly evaluate the decision.

Dryzek's (2000) account of discursive legitimacy supplies the mechanism connecting individual evaluations to legitimacy as a systemic property. For Dryzek, what bears on legitimacy is not the aggregate of private preferences but the prevailing constellation of discourses in the public sphere — the shared frames within which a decision becomes intelligible and arguable. A policy is discursively legitimate to the extent that the dominant framing of it in public argument is one that accommodates rather than repudiates it. This shifts the analytical object from the count to the configuration: from how many citizens object to how objection is organised, whether it stabilises into a settled frame, and whether governing actors succeed in displacing that frame over time. It is on this basis that the temporal dimension of the present analysis carries interpretive weight independent of the headline proportion.

Two further theoretical resources are used in a supporting capacity. Lippmann's (1922) account of public opinion establishes that citizens respond not to policy as such but to mediated representations of it. This is not a reason to discount expressed sentiment; it is a reason to read it as evidence of how a policy has been publicly constructed rather than as an audit of the policy's actual effects. Habermas's (1989) conception of the public sphere provides the warrant for treating a platform such as X as a site where opinion is discursively formed rather than merely registered, subject to the well-established qualifications concerning access and stratification that subsequent scholarship has documented (Dahlberg, 2001; Papacharissi, 2002; Sassi, 2005).

Social Media Discourse as Evidence of Public Evaluation

The use of platform data as a source of evidence about public opinion rests on a straightforward claim: that publicly posted evaluative speech about a policy constitutes an observable trace of how that policy is being received. Bruns and Stieglitz (2012) established the methodological groundwork for treating platform discourse as quantifiable, and Pak and Paroubek (2010) demonstrated the viability of microblog corpora for

sentiment analysis specifically. Grimmer and Stewart (2013) supplied the field's most influential caution: automated content analysis methods do not replace careful reading, all such methods are wrong in ways that matter, and validation against human judgement is not optional but constitutive of the method's claim to measure anything at all.

Indonesian scholarship applying these methods to national policy controversies is now substantial. Sucahyo, Kurniati, and Harvit (2022) analysed discourse on the Omnibus Law on Job Creation using a Naive Bayes classifier over a corpus crawled with 160 keywords, reporting negative sentiment at 75.77 per cent. Putranta, Rahayudi, and Purnomo (2023) examined fuel subsidy removal using Bernoulli Naive Bayes with N-gram TF-IDF features, reporting a substantially narrower margin at 52 per cent negative. Pandunata et al. (2022) applied comparable methods to COVID-19 vaccination discourse. Closest to the present case, Maharani, Gustriansyah, and Irfani (2025) analysed sentiment toward MBG itself using K-Nearest Neighbors classification over 3,007 tweets, finding negative sentiment dominant and public concern centred on budget allocation and uneven food distribution, while Tadu and Hariadi (2025) examined the 2025 budget efficiency policy using Support Vector Machine classification, reporting negative sentiment at 42.54 per cent with an overall accuracy of 60 per cent.

Two features of this literature are relevant here. The first is that the reported distributions vary widely across cases — from 42.54 to 75.77 per cent negative — which suggests either genuine variation in public response or variation in what the instruments are capturing, and the literature does not currently permit these to be distinguished. The second is that theoretical interpretation is largely absent. These studies establish that sentiment is negative; they do not, for the most part, articulate what negative sentiment toward a legally enacted policy signifies for that policy's standing. The present study addresses both features: it supplies the interpretive frame that the literature lacks, and it treats the reliability of the instrument as a question to be investigated rather than assumed.

RESEARCH METHOD

This study employs a descriptive-computational quantitative design. The analytical task is the systematic description of a discourse and the interpretation of that description against a theoretical framework, not the estimation of population parameters or the testing of causal hypotheses. The design is accordingly oriented toward coverage and classification accuracy rather than toward sampling inference, and the claims it can support are correspondingly bounded: they concern the discourse retrieved, not the Indonesian public.

The data used for this study were collected by automated retrieval from X using tweet-harvest v2.6.1, a Node.js utility executed through Google Colaboratory. Retrieval was governed by five Indonesian-language

keywords selected to capture discourse specifically about the budgetary arrangement rather than about MBG in general: anggaran pendidikan dipotong; dana pendidikan dialihkan MBG; makan bergizi gratis anggaran pendidikan; anggaran pendidikan mbg; and anggaran pendidikan dialihkan mbg. Retrieval was restricted to Indonesian-language content (lang:id), excluded retweets, and excluded protected accounts. The observation window ran from January 2025 to May 2026, a period of seventeen months, with May 2026 partial. Retrieval proceeded in successive batches with a limit of 500 posts per run.

The composition of the keyword set requires explicit acknowledgement. Three of the five keywords contain terms that are themselves evaluatively loaded: dipotong (cut) and dialihkan (diverted) presuppose the critical framing under investigation. This is a design feature with a foreseeable directional consequence — the retrieval is more likely to surface critical discourse than supportive discourse, and the observed proportion is therefore not independent of the instrument that produced it. The issue is revisited in the limitations discussion.

Retrieved posts underwent a two-stage cleaning pipeline. The first stage, conducted in R using the stringr and dplyr packages, removed duplicate entries, converted all text to lowercase, and stripped URLs, mentions, hashtag symbols, non-alphanumeric characters, and excess whitespace. The second stage, conducted in Python using NLTK, applied further regular-expression noise removal, removed Indonesian stopwords using the NLTK Indonesian corpus supplemented by a manually curated list of informal social-media tokens, and excluded tokens of two characters or fewer. Posts containing fewer than five words after preprocessing were removed as insufficiently informative for classification. Removal of 271 duplicate entries reduced the raw corpus of 2,308 posts by 11.74 per cent, yielding a final corpus of 2,037 posts comprising 34,822 tokens.

Classification used the InSet Lexicon (Koto & Rahmaningtyas, 2017), which comprises 3,610 positive and 6,610 negative Indonesian terms, each carrying a polarity weight on a scale from −5 to +5. The lexicon was selected because it requires no labelled training corpus, because its scoring is inspectable at the level of the individual term, and because it was constructed and evaluated specifically for Indonesian microblog text (Musfiroh et al., 2021). Each post was tokenised, each token matched against the lexicon, and the polarity weights of all matched tokens summed to produce a composite score. Posts scoring below zero were classified negative, above zero positive, and exactly zero neutral. The classification rule is deterministic and unthresholded.

A random sample of 100 posts was drawn from the classified corpus using a fixed seed (set.seed(42)) and independently assigned Positive, Negative, or Neutral labels through holistic reading. Agreement between these labels and the automated classifications was quantified using Cohen's Kappa (Cohen, 1960),

computed with the kappa2() function of the irr package and interpreted against the benchmark scale of Landis and Koch (1977). Two properties of this protocol constrain what its result can establish. First, the manual labelling was performed by a single coder who was not blind to the automated output, which means the exercise measures human–machine agreement rather than inter-rater reliability in the conventional sense; the appropriate reporting framework for such a comparison is per-class accuracy, precision, recall, and F1 against a gold standard, of which Kappa is at best a partial summary (McHugh, 2012). Second, a sample of 100 is too small to estimate per-class accuracy with useful precision. The result reported below should therefore be read as establishing the existence and direction of a classification problem rather than its exact magnitude.

RESULTS AND DISCUSSION

Sentiment Distribution

The final corpus comprises 2,037 posts published between January 2025 and May 2026, containing 34,822 tokens after stopword removal. Negative posts number 1,360, or 66.76 per cent of the corpus. Positive posts number 581 (28.52 per cent) and neutral posts 69 (3.39 per cent).

Table 1. Sentiment distribution and score statistics

Category	Count	Share
Negative	1,360	66.76%
Positive	581	28.52%
Neutral	69	3.39%
Mean score	−5.405	Median −4.0
Mean, negative	−10.422	—
Mean, positive	+5.695	—
Score range	−50 to +30	Additive

Source: processed by the author

The aggregate mean of −5.405 with a median of −4.0 indicates that the corpus leans negative in cumulative polarity weight as well as in category counts. The mean score of negative posts (−10.422) is close to twice the absolute magnitude of the mean positive score (+5.695).

This asymmetry requires a qualification the raw figures do not carry. Because scores are additive and unbounded, a longer post containing more lexicon-matched tokens will register a larger absolute score irrespective of the intensity of any individual term. The observed difference in mean magnitude is therefore confounded with post length and cannot on its own support an inference about expressive intensity. A length-normalised score is required to separate the two, and is not yet computed for this corpus. The distributional finding in the upper panel of Table 1 does not depend on this qualification; the intensity claim does.

The neutral proportion of 3.39 per cent warrants comment as an artefact rather than a substantive finding. Under the classification rule specified above, a post is neutral only if it contains no lexicon-matched tokens or if its matched weights sum to exactly zero. No threshold band is applied. A post with a

composite score of +1 is classified positive on the same basis as a post scoring +25. The near-absence of a neutral category is therefore a property of the decision rule, and any comparison with studies employing threshold bands must take this into account.

Classification Validity and Its Implications

Agreement between the automated classifications and manual coding of the 100-post sample yielded a Cohen's Kappa of 0.0647 ($z = 1.07$, $p = 0.284$). On the Landis and Koch (1977) benchmark this falls within the slight agreement band, and the associated p-value indicates that the observed agreement cannot be distinguished from chance.

Table 2. Validation summary

Measure	Value	Reading
Cohen's Kappa	0.0647	Slight agreement
Significance	z = 1.07; p = 0.284	Not distinct from chance
Positive-class error	24 of 30	Irony read as endorsement
Negative, automated	66% of sample	Lower bound
Negative, manual	73% of sample	Direction confirmed

Source: processed by the author

The disagreement is not randomly distributed across classes. Of the 30 sampled posts the lexicon classified as Positive, 24 were assessed as Negative on manual reading — a concentration that identifies the source of failure precisely. The lexicon assigns positive weights to literally positive vocabulary deployed ironically, which in this corpus takes recognisable forms: expressions of mock gratitude toward the government, ostensible congratulation on the budget allocation, and formulaic praise attached to descriptions of adverse consequences. Word-level scoring has no access to the pragmatic reversal that makes such utterances critical. This is the failure mode documented by Lunando and Purwarianti (2013), who found sarcasm markedly more prevalent in Indonesian social media discourse on politically sensitive topics than in general conversation, and by Sykora, Elayan, and Jackson (2020), who report that sarcastic content alone can reduce automated sentiment detection accuracy by up to half.

The directional finding nonetheless survives the failure, and survives it in a specific direction. On the same 100-post sample, the lexicon classified 66 per cent as negative while manual reading assessed 73 per cent as negative. The classification error is asymmetric: it converts negative posts into positive ones and not the reverse. This means the corpus-level figure of 66.76 per cent negative is best understood as a lower bound on the negative share rather than as a point estimate. If the sampled error rate for the positive class were to hold across the corpus, the true negative proportion would be substantially higher; that extrapolation is not made here, because a sample of 30 positive-classified posts cannot support it, but the direction of the bias is established and is not in dispute.

The implication for the Indonesian literature reviewed above is uncomfortable and

should be stated as such. If lexicon-based classification systematically absorbs ironic criticism into the positive category, then studies employing such classification without validation may be reporting negative proportions that are biased downward by an unknown margin. The wide spread of reported distributions across that literature — from 42.54 to 75.77 per cent negative — is at least consistent with instrument variation as an explanation alongside genuine variation in public response. Nothing in the present study permits a determination between these. What it does establish is that the question cannot be settled by studies that do not validate, and that validation is therefore not a methodological courtesy but a precondition for the numbers to bear the interpretive weight routinely placed on them (Grimmer & Stewart, 2013).

Thematic Composition of the Discourse

Word-frequency analysis of the preprocessed corpus was conducted using NLTK's FreqDist function. The highest-frequency substantive terms are anggaran, budget (2,232), mbg (1,345), pendidikan, education (1,331), program (444), dipotong, cut (351), gratis, free (275), makan, meal (264), pemerintah, government (260), dana, funds (253), triliun, trillion (250), negara, state (250), guru, teacher (231), dialihkan, diverted (220), bergizi, nutritious (201), apbn, state budget (183), rakyat, the people (179), and sekolah, school (178). Four thematic clusters were identified through interpretive semantic grouping of high-frequency terms within the policy context, following established practice in qualitative-computational content analysis (Krippendorff, 2004; Grimmer & Stewart, 2013). The clustering is the product of researcher interpretation rather than algorithmic procedure.

The first cluster, fiscal contestation, is organised around anggaran (2,232), dipotong (351), dialihkan (220), dana (253), and triliun (250). Its structure is notable: the vocabulary through which the policy is described is not neutral budgetary language but the language of subtraction. The policy enters the discourse already framed as a removal of resources rather than as a reallocation between functions.

The second cluster, attribution of political responsibility, comprises pemerintah (260), negara (250), and references to the President, the House of Representatives, and BGN. The distribution within this cluster carries interpretive weight: the highest-frequency terms denote the government and the state as institutions, while the agency that administers MBG and holds the contested allocation appears substantially less often. Criticism is directed at the decision and at those who made it, not at the body implementing it.

The third cluster, social impact on vulnerable groups, is constituted by guru (231), rakyat (179), sekolah (178), and the compound guru honorer, honorary teachers. The specific salience of honorary teachers is significant given that this group brought the Constitutional Court petition; the discourse identifies a bounded and sympathetically

framed constituency as bearing the cost of the arrangement.

The fourth cluster, evaluation of the programme itself, is the most consequential for the argument advanced here. Terms associated with MBG — gratis (275), makan (264), bergizi (201) — appear frequently, but they appear without an accompanying vocabulary of nutritional objection. There is no corresponding cluster of terms disputing the value of feeding schoolchildren. The criticism registered in this corpus is directed at the financing mechanism, not at the welfare objective.

Temporal Trajectory and Discursive Hardening

Table 3. Monthly distribution by sentiment category

Period	Neg.	Neut.	Pos.	Total
Jan–Dec 2025	217	17	91	325
January 2026	67	5	36	108
February 2026	269	14	123	406
March 2026	517	19	209	745
April 2026	261	11	101	373
May 2026*	29	3	11	43
Total	1,360	69	581	2,037

**May 2026 data are partial. Source: processed by the author*

Three features of this distribution bear on the analysis. The first is the concentration of volume: February, March, and April 2026 together account for 74.81 per cent of the corpus. The escalation from 108 posts in January 2026 to 406 in February corresponds to the legislative confirmation on 25 February 2026 that MBG funding was drawn from the education allocation, and the March peak of 745 posts coincides with the escalation of the Constitutional Court petition. Discourse volume tracks identifiable policy events rather than fluctuating randomly.

The second feature is the persistence of negative majority status. In every period recorded — including the low-volume months of 2025, which precede the legislative confirmation entirely — negative posts constitute a majority. The proportion in the 2025 period stands at 66.8 per cent, before the controversy achieved national salience. The critical frame was not manufactured by the controversy; it was available in advance of it.

The third feature is the direction of change under expanding participation. The negative proportion moves from 62.0 per cent in January 2026 to 66.3 per cent in February, 69.4 per cent in March, and 70.0 per cent in April. As the number of participants grew roughly sevenfold, the critical framing did not dilute. It consolidated. This is not the pattern one would expect if the early critical majority were an artefact of a small and unrepresentative early cohort; incoming participants would then dilute it. In Dryzek's (2000) terms, what the temporal data records is a discursive constellation consolidating rather than one in contest, and the government's counter-argument — advanced in the legislature through appeals to comparable school meal programmes in Japan, Brazil, and

South Korea — did not displace it within the observation window.

The distinction between this and a simple negative distribution is worth stating plainly. A high negative proportion at one moment is compatible with a transient reaction. A negative proportion that holds across seventeen months, predates the triggering event, and intensifies under expanding participation is a different kind of object. It is the stability rather than the magnitude that supports the legitimacy inference.

Legality Held, Consent Withheld

Read through Beetham's (1991) framework, the case exhibits a clear separation between dimensions. The classification of MBG expenditure within the education budget conforms to established rules: it is enacted by statute, elaborated by presidential regulation, and grounded in a specific statutory provision extending educational operational funding to nutritional provision at educational institutions. Whether that grounding is constitutionally adequate is the matter now before the Constitutional Court, and this study takes no position on it. On the first dimension, the policy stands unless and until the Court determines otherwise.

The third dimension presents differently. What the corpus records is not the absence of expressed consent but the presence of expressed refusal — public, evaluatively explicit, and circulated at scale. Beetham's point is that consent is demonstrated through actions that publicly signal endorsement; in this discourse the corresponding actions signal the opposite, and they do so consistently. The 521,689 retweets and 1,351,930 likes generated by the 2,037 posts matter here for a specific reason: they establish that these evaluations were not private views that happened to be visible but communications that were taken up and propagated by others. Whatever else the corpus shows, it does not show a public that is indifferent.

The theoretical yield is the demonstration that the dimensions genuinely come apart in practice, and that the separation is observable rather than merely assertable. A policy can hold the legality dimension in full while the consent dimension is contested in public, and the second condition is not remedied by the first. This is a modest claim, but it is one the existing Indonesian sentiment literature has not been positioned to make, because reporting a distribution without a theoretical frame leaves no vocabulary in which the distribution can mean anything.

The Object of Criticism: Financing, Not Purpose

The fourth thematic cluster carries the most directly actionable implication, and it is the finding most likely to be of use to those responsible for the policy. Terms denoting the programme appear frequently in the corpus, but no cluster of terms disputing its nutritional purpose accompanies them. Criticism attaches to how the programme is paid for, not to what it does. Read alongside the second cluster — in which responsibility is attributed to the government and the state as decision-makers

rather than to the implementing agency — the corpus describes a public that objects to a budgetary choice, not to a welfare intervention.

This distinction has practical weight because the two objections would call for opposite responses. A public that rejected school feeding on its merits could only be answered by defending or abandoning the programme. A public that accepts the programme but rejects its financing can, in principle, be answered by changing the financing. The discourse examined here is of the second kind. That the government's public argument during the observation window was directed largely at defending MBG's value — a proposition the corpus does not show to be under attack — suggests a mismatch between the objection raised and the answer offered.

Limitations

The corpus is defined by its retrieval parameters and is not a sample of a population. It comprises Indonesian-language posts containing five specified keywords, published on one platform, excluding retweets and protected accounts. Discourse on the same policy conducted in other vocabulary, in images, in reply threads that do not repeat the keywords, or on other platforms is not captured. The findings describe this discourse and do not extend to Indonesian public opinion.

Three of the five keywords are evaluatively loaded. This is the most serious threat to the distributional finding, and it is not remedied by corpus size. The required correction is a robustness retrieval using neutral keywords and a comparison of the resulting proportions against those reported here. That procedure has not been carried out, and until it is, the magnitude of the retrieval effect on the reported 66.76 per cent is unknown. The persistence and temporal-trajectory findings are less exposed to this threat, since a constant retrieval bias cannot by itself produce a rising negative proportion across months, but the level is exposed and should be read accordingly.

The study does not implement bot or coordinated-account detection, and the number of unique accounts contributing to the corpus is not established. Volume concentrated among a small number of high-frequency accounts would substantially alter the interpretation of the proportions, and this possibility cannot currently be excluded. The population represented is also self-selected in two respects: it comprises users of one platform, whose Indonesian user base skews younger, more urban, and more highly educated than the national population, and within that base only those who chose to post about this policy.

Finally, the study operationalises one of Beetham's three dimensions. Legality is currently before the Constitutional Court and its resolution is pending. Normative justifiability — whether the values underlying the policy are genuinely shared between governors and governed — is not accessible through sentiment classification and would require interpretive or deliberative methods. The claim advanced here is correspondingly

bounded: a discursive legitimacy deficit within the discourse examined, not a comprehensive legitimacy assessment.

CONCLUSION

Addressing the first research question, within a corpus of 2,037 posts spanning seventeen months, negative evaluation constitutes 66.76 per cent of classified content and holds majority status in every period observed, including the months preceding the controversy's national salience. The proportion rises rather than falls as participation expands, moving from 62.0 per cent in January 2026 to 70.0 per cent in April. Criticism is organised around four themes, of which the most consequential establishes that objection attaches to the financing mechanism and not to the programme's nutritional purpose. Responsibility is attributed to the government and the state as decision-makers rather than to the implementing agency.

Addressing the second, validation against manual coding returned a Cohen's Kappa of 0.0647, with error concentrated almost entirely in the positive class, where 24 of 30 sampled posts were reclassified as negative on reading. The lexicon fails on irony, irony in this corpus is a critical register, and the resulting bias runs in one direction. The reported negative proportion is therefore a lower bound.

The theoretical contribution is to demonstrate that Beetham's consent dimension can be examined empirically in naturally occurring discourse, and that it separates observably from the legality dimension: this policy holds the first and is contested on the third. The methodological contribution is to document, within a real policy corpus, a systematic and directional failure of an instrument in wide use in Indonesian scholarship, and to argue that unvalidated lexicon results in this domain cannot bear the interpretive weight commonly placed on them.

The practical implication is narrow but concrete. A public that objects to how a programme is financed rather than to what the programme does is answerable by adjusting the financing. Whether that adjustment is warranted is a political and constitutional question outside the scope of this study; that the objection has been misidentified in public argument is a finding within it.

REFERENCES

Beetham, D. (1991). The legitimation of power. Macmillan Education.

Bruns, A., & Stieglitz, S. (2012). Quantifying Twitter discussions. Proceedings of the 6th AAAI International Conference on Weblogs and Social Media, 45–52.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20(1), 37–46.

Dahlberg, L. (2001). The Internet and democratic discourse: Exploring the prospects of online deliberative forums extending the public sphere. Information, Communication & Society, 4(4), 615–633.

Dryzek, J. S. (2000). Deliberative democracy and beyond: Liberals, critics, contestations. Oxford University Press.

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. Political Analysis, 21(3), 267–297.

Habermas, J. (1989). The structural transformation of the public sphere: An inquiry into a category of bourgeois society (T. Burger & F. Lawrence, Trans.). MIT Press.

Koto, F., & Rahmaningtyas, G. Y. (2017). InSet lexicon: Evaluation of a word list for Indonesian sentiment analysis in microblogs. 2017 International Conference on Asian Language Processing (IALP), 391–394.

Krippendorff, K. (2004). Content analysis: An introduction to its methodology (2nd ed.). SAGE Publications.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159–174.

Lippmann, W. (1922). Public opinion. Harcourt, Brace and Company.

Liu, B. (2015). Sentiment analysis: Mining opinions, sentiments and emotions. Cambridge University Press.

Lunando, E., & Purwarianti, A. (2013). Indonesian social media sentiment analysis with sarcasm detection. International Conference on Advanced Computer Science and Information Systems (ICACSIS), 195–198.

Maharani, Gustriansyah, & Irfani. (2025). Analysis of public sentiment towards the Free Nutritious Meal Program in schools based on tweets using the K-Nearest Neighbors method.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. Biochemia Medica, 22(3), 276–282.

Musfiroh, D., Khaira, U., Utomo, P. E. P., & Suratno, T. (2021). Sentiment analysis of online lectures in Indonesia from Twitter dataset using InSet lexicon. MALCOM: Indonesian Journal of Machine Learning and Computer Science, 1(1), 24–33.

Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. Proceedings of LREC 2010, 1320–1326.

Pandunata, P., et al. (2022). Analisis sentimen opini publik terhadap program vaksinasi Covid-19 di Indonesia pada Twitter menggunakan metode Naïve Bayes Classifier. Informatics Journals, 7(3).

Papacharissi, Z. (2002). The virtual sphere: The Internet as a public sphere. New Media & Society, 4(1), 9–27.

Putranta, Rahayudi, & Purnomo. (2023). Analisis sentimen masyarakat terhadap kebijakan penghapusan subsidi BBM pada media sosial Twitter menggunakan algoritma Naive Bayes Classifier dengan ekstraksi fitur N-Gram TF-IDF. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 7(3), 1501–1510.

Republic of Indonesia. (2024). Presidential Regulation Number 83 of 2024 concerning the National Nutrition Agency. State Gazette No. 173.

Republic of Indonesia. (2025). Law Number 17 of 2025 concerning the 2026 State Budget.

Republic of Indonesia. (2025). Presidential Regulation Number 118 of 2025 concerning the details of the 2026 State Budget. State Gazette No. 186.

Republic of Indonesia. (2025). Presidential Regulation Number 115 of 2025 concerning the governance of the Free Nutritious Meal Programme. State Gazette No. 183.

Sassi, S. (2005). Cultural differentiation or social segregation? Four approaches to the digital divide. New Media & Society, 7(5), 684–700.

Sucahyo, N., Kurniati, I., & Harvit, K. (2022). Analisis sentimen masyarakat terhadap UU Cipta Kerja pada media sosial Twitter. Jurnal Rekayasa Informasi Swadharma, 2(1).

Sykora, M., Elayan, S., & Jackson, T. W. (2020). A qualitative analysis of sarcasm, irony and related hashtags on Twitter. Big Data & Society, 7(2).

Tadu, & Hariadi. (2025). Sentiment analysis of the 2025 budget cut policy using the SVM method.