Journal of Media, Journalism & Communication Studies

Vol. 1 No. 1 2025

Article **Open Access**

Bayesian Causal Identification and Modeling of Advertising Conversion Paths

Jing Xie [1,*]

[1] Steinhardt School of Culture, Education, and Human Development, New York University, New York, NY 10003, USA
* Correspondence: Jing Xie, Education and Human Development, NYU Steinhardt School of Culture, NYU, NY, 10012, USA

Received: 01 November 2025
Revised: 30 November 2025
Accepted: 12 December 2025
Published: 13 December 2025

Abstract: Within the internet advertising landscape, the intricate interplay between user exposure, clicks, and conversions often eludes precise measurement. Conventional analytical methods typically capture only superficial causal relationships, failing to accurately depict advertising effectiveness. This paper employs Bayesian causal modelling to construct a framework for identifying and predicting advertising conversions, probabilistically characterising influencing factors at each stage. By decomposing evaluations of diverse advertising pathways using predefined antecedents and consequents, the primary pathway is identified. Experimental results demonstrate that this method enables stable inference under incomplete information, provides more rational support for advertising optimisation, and offers a fresh perspective for exploring causal relationships in digital markets.

Keywords: Bayesian inference; causal identification; advertising conversion pathways; model analysis

1. Introduction

Against the backdrop of fully developed digital marketing technologies, enterprises increasingly prioritise the authenticity of outcomes generated by their advertising campaigns. However, due to the presence of various influencing factors, it remains challenging to determine "whether advertising is effective" using traditional statistical methods. Beneath the seemingly straightforward relationship between ad exposure, clicks, and conversions lie intricate causal networks. This paper employs Bayesian causal models to analyse the causal networks underpinning advertising conversions within existing information and data constraints. It seeks to achieve a balance between interpretability and predictive power while proposing a novel research approach for evaluating advertising effectiveness [1].

2. Causal Identification Logic and Research Framework for Advertising Conversion Pathways

2.1. Structural and Causal Characteristics of Advertising Conversion Pathways

The advertising conversion process typically encompasses three primary stages: exposure, click-through, and conversion. Users encounter advertisements while browsing information feeds; some individuals develop interest and click, ultimately completing a purchase or registration. This process is influenced by factors including ad content, placement, audience characteristics, and external environments, with complex causal dependencies existing between variables [2]. Traditional analyses often rely on correlation-based judgements, which may mistakenly interpret "increased click-through rates alongside rising conversion rates" as causation, while overlooking potential confounding factors such as users' inherent purchasing propensity, price discounts, and temporal differences. The core of causal identification lies in distinguishing between "advertisements causing conversions" and "advertisements merely coinciding with conversions," thereby establishing structural models that reflect genuine effects.

Bayesian causal inference methods utilise probabilistic graphical models. By constructing directed edges between variables, they explicitly define causal directions and adjust parameters through posterior mechanisms during data updates. To reveal the causal pathways and interactions across advertising stages, a Bayesian causal inference framework for advertising conversion paths can be established, as illustrated in Figure 1.

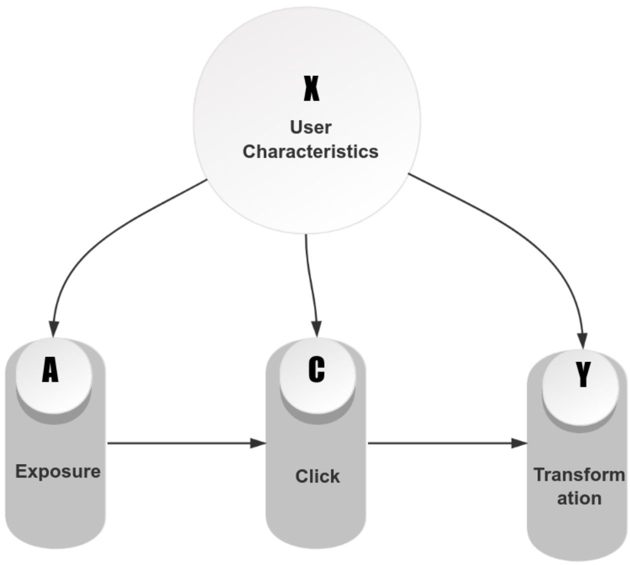

Figure 1. Overall Framework Diagram for Bayesian Causal Identification of Advertising Conversion Pathways.

Nodes represent primary variables (impressions A, clicks C, conversions Y, user characteristics X), while arrows denote causal directionality. User characteristics influence ad impressions, clicks, and conversions respectively. Ad impressions exert an indirect effect on conversions via clicks, whilst also possessing a direct influence pathway. This framework maintains inferential robustness under conditions of limited samples or incomplete information, proving applicable for multi-channel advertising effectiveness analysis and path contribution assessment [3].

2.2. Key Paradoxes and Conversion Logic in Advertising Causality Identification

The primary challenge in advertising causality analysis lies in the counterfactual dilemma. Actual data only records whether users who saw an advertisement converted, without simultaneously observing whether they would have converted without exposure. This missing control information makes direct causal identification difficult. Furthermore, ad exposure often co-varies with external factors such as user interest, marketing activities, or promotional timing, creating confounding effects. Without proper control, model estimates are prone to bias, leading to erroneous causal inferences [4].

Causal structure models, through hierarchical descriptions of variable relationships, can clearly distinguish direct from indirect pathways, thereby identifying the true contribution of each element in the conversion process. Bayesian methods incorporate uncertainty probabilistically into the inference process. By jointly updating prior distributions with sample information, they progressively approximate the true effect of advertising behaviour. This approach not only measures the overall impact of advertising but also reveals the transmission mechanisms of effects across different pathways, providing a basis for advertising optimization [5].

2.3. Overall Conceptual Framework of Bayesian Causal Identification

The causal effects of advertising conversion can be probabilistically characterised within a Bayesian inference framework. Let advertising exposure denote A, conversion outcome denote Y, and control variables denote X. Under intervention conditions, the conversion probability can be expressed as:

$$P(Y|do(A)) = \sum_X P(Y|A,X)\, P(X) \quad (1)$$

Here, P(Y|do(A)) denotes the conversion probability when exposed to the intervention advertisement, while P(Y|A,X) represents the probability under observed conditions. This expression employs the total probability decomposition to eliminate confounding variables, thereby yielding an estimate of the advertisement's net effect.

The core of Bayesian inference lies in expressing parameters as distributions, where models continuously update their posterior distributions with incoming data, enabling robust inference even under conditions of high noise or insufficient samples. This framework not only estimates effect sizes but also reveals causal directions and strength between variables, providing quantitative grounds for advertising resource allocation and strategy evaluation. By introducing hierarchical structures, it enables layered analysis of different ad types and audience segments, thereby enhancing the model's adaptability and interpretability across diverse deployment scenarios [6].

3. Bayesian Causal Modelling Approach and Inference Mechanism

3.1. Model Structure and Variable Definition

Advertising conversion behaviour may be conceptualised as a multi-layered causal system, typically comprising three stages: "advertisement exposure - user click - outcome conversion". To elucidate the causal relationships between these stages, the conversion probability can be represented as a function of advertisement exposure and user characteristics:

$$P(Y|A,X) = log\, i\, t^{-1}(\gamma_0 + \gamma_1 A + \gamma_2 X) \quad (2)$$

Here, P(Y|A,X) denotes the probability of conversion occurring given ad exposure and user characteristics; $logit^{-1}$(.) represents the inverse of the logit function; A denotes the ad exposure status (1 for exposed, 0 for unexposed); Y represents the conversion outcome (1 for converted, 0 for not converted); X is the set of user characteristic variables; γ_0 is the constant term, γ_1 is the effect coefficient for ad exposure, and γ_2 is the coefficient vector for control variables. This model characterises the direct effect of ad exposure on conversion, forming the foundation for subsequent path decomposition and Bayesian inference [7].

The model's logical structure comprises three tiers: the input layer reflects external placements and audience characteristics; the intermediate layer captures user interaction responses; and the output layer corresponds to final conversion outcomes. To facilitate subsequent model expansion and variable stratification, Table 1 summarises the hierarchical structure of the model.

Table 1. Hierarchical Structure of the Model.

Level	Main Content	Example Variables	Function Description
Input Layer	Advertising and audience features	Exposure volume, display frequency, user profiles, etc.	Describe ad reach and audience differences

Middle Layer	User interaction behavior	Clicks, browsing depth, dwell time, etc.	Reflect the immediate response triggered by the ad
Output Layer	Conversion results	Purchases, registrations, inquiries, etc.	Measure the actual effectiveness of the advertisement

This structure clearly delineates the hierarchical levels and functional positions of variables within the advertising conversion process, providing a standardised variable framework for Bayesian causal modelling while laying the groundwork for subsequent prior specification and effect identification.

3.2. Prior Distribution and Posterior Update

Bayesian methods achieve dynamic parameter updates and uncertainty characterisation by integrating prior information with sample data. Prior distributions reflect pre-observation knowledge about parameters, whilst posterior distributions embody how data evidence modifies this knowledge; together they determine the model's estimation outcomes.

Let the model parameter set be Θ\ThetaΘ and the observed data be DDD. According to Bayes' theorem:

$$p(\Theta|D) = \frac{p(D|\Theta)p(\Theta)}{p(D)} \quad (3)$$

Here, p(Θ|D) denotes the posterior distribution, representing the updated parameter values conditional upon the data D; p(D|Θ) is the likelihood function, characterising the extent to which the data supports the parameters; p(Θ) constitutes the prior distribution, embodying empirical knowledge; and p(D) serves as the normalising constant. This relationship indicates that parameter estimation constitutes a probabilistic outcome achieved through balancing empirical data with subjective cognition.

In advertising conversion modelling, probability parameters such as conversion rates are often set with a Beta distribution prior, which flexibly captures the prior uncertainty of successful events. Regression coefficients typically employ a normal distribution prior, reflecting the characteristic that parameters centre around zero while permitting deviation. When strong information is lacking, weak prior information may be selected to ensure estimation stability.

The posterior distribution is typically obtained through Markov Chain Monte Carlo (MCMC) sampling. This method maintains robustness even with limited samples or high noise levels, and can output confidence intervals and uncertainty ranges for parameters. The distribution settings for key parameters are shown in Table 2.

Table 2. Parameter Distribution Type Settings.

Parameter Type	Distribution Form	Applicable Range	Description
Conversion probability parameter	Beta distribution	0-1 interval	Represents the prior uncertainty of event probability
Regression coefficient parameter	Normal distribution	Real number interval	Controls the central tendency of variables or exposure effects
Variance parameter	Gamma distribution	Positive number interval	Describes the intensity of noise or fluctuation

This mechanism provides the core basis for parameter updating in Bayesian causal models, enabling the model to maintain stable and adaptive estimation capabilities under varying sample conditions.

3.3. Path Contribution Identification and Causal Effect Decomposition

The overall effectiveness of advertising campaigns arises from the combined action of multiple causal pathways. To identify the contribution of each pathway, the total effect can be decomposed into direct and indirect effects. The direct effect reflects the independent influence of ad exposure on conversion, while the indirect effect represents the influence generated indirectly through mediating variables such as clicks. This decomposition reveals the true role of advertising at different stages of the communication process.

The total advertising effect can be expressed as:

$$TE = DE + IE \tag{4}$$

TE denotes the total effect, DE represents the direct effect, indicating the independent influence of ad exposure on conversion; IE signifies the indirect effect, reflecting the indirect impact of advertising on conversion via mediating variables. This decomposition enables assessment of the contribution strength across different pathways, providing quantitative grounds for advertising optimisation.

Within advertising conversion scenarios, the indirect effect may be further expressed as:

$$IE = EX[P(Y|A = 1, C, X)] - EX[P(Y|A = 0, C, X)] \tag{5}$$

Here, EX[.] denotes the expected value on the user feature variable X; A represents the ad exposure status; C denotes the click behaviour; and Y signifies the conversion outcome. This expression reflects the average impact of ad exposure on conversion rate changes via the click path, after controlling for user features. A higher IE value indicates that the click path plays a dominant role in ad conversions; conversely, a lower value suggests that direct exposure contributes more significantly to conversions.

The effect values following path decomposition can be calculated from posterior samples and output as means, confidence intervals, or standardised metrics. By comparing effects across different paths, high-value links can be identified, providing quantitative grounds for advertising budget allocation and channel evaluation. Simultaneously, these decomposition results serve as a crucial component in model interpretability validation, ensuring consistency between causal structure assumptions and actual data performance.

3.4. Model Convergence and Uncertainty Assessment

The reliability of Bayesian models primarily manifests in the convergence of parameter posterior distributions and the stability of estimates. Model parameters are typically obtained via Markov chain Monte Carlo sampling; should the sampling chain fail to reach a steady state, inferential results may exhibit bias. To assess the consistency of sampling chains, the Gelman-Rubin criterion may be employed, calculated as follows:

$$\hat{R} = \sqrt{\frac{\hat{V}}{W}} \tag{6}$$

Among these, $\hat{R}$ denotes the multi-chain convergence statistic, $\hat{V}$ represents the inter-chain variance, and W signifies the intra-chain variance. When $\widehat{V}$ approaches 1, it indicates stable sampling results; if exceeding 1.1, extending iterations or adjusting initial values is required to enhance convergence.

Following model convergence, parameter uncertainty should be further assessed. The posterior distribution provides confidence intervals and variance information for parameters, enabling evaluation of estimation robustness. Narrow intervals that do not cross zero indicate significant effects with clear directionality; broad intervals or unstable signs suggest residual model variability, necessitating increased samples or optimised priors to enhance precision. Overall model performance may also be assessed by combining prediction error and goodness-of-fit metrics, commonly including root mean square error (RMSE), coefficient of determination (R^2), and Deviance Information Criterion (DIC). A lower RMSE indicates reduced error, a higher R^2 signifies greater explanatory power, and a smaller DIC reflects a more favourable balance between model precision and complexity. By subjecting Bayesian causal models to dual tests of convergence and uncertainty, the robustness and reliability of their estimation results can be assured.

4. Extensions and Application Value of Bayesian Causal Models

4.1. Hierarchical Extension of Model Structure

Bayesian causal models demonstrate excellent scalability in advertising conversion analysis. To accommodate variations in ad types, channels, and audience characteristics, the foundational model may be structured hierarchically. This approach maintains a balance between overall trends and individual features. By introducing upper-level distributions, the hierarchical structure facilitates information sharing at the aggregate level while preserving local diversity, thereby enhancing estimation stability. Its structural form is as follows:

$$\gamma_{1i} \sim \mathrm{N}\ \left(\mu_{\gamma 1}, \sigma_{\gamma 1}^2\right) \tag{7}$$

Here, γ_{1i} denotes the exposure effect parameter for the i-th category of advertising samples, N represents a normal distribution, $\mu_{\gamma 1}$ signifies the population mean, and $\sigma_{\gamma 1}^2$ indicates the variance. This structure enables the model to share information across different advertising channels while preserving individual differences, making it suitable for advertising effectiveness modelling in multi-source data environments. The hierarchical Bayesian approach not only enhances estimation accuracy but also integrates cross-platform data to achieve unified inference of campaign performance.

4.2. Adaptive Extensions of Inference Methods

Given the vast volume of advertising data and complex variables, traditional MCMC methods face limitations in computational efficiency and convergence. Variational Bayes (VB) or Hybrid Monte Carlo (HMC) techniques can be employed to estimate high-dimensional parameters. Variational inference methods represent the posterior probability density as a series of easily manageable forms, utilising optimisation algorithms to approximate it. This reduces computational costs while yielding approximate results. In contrast, HMC employs gradient information to leap through parameter space, maintaining high-quality sampling under complex conditions. These approaches can be freely selected and applied across varying scales and patterns. In designing advertising conversion workflows, hybrid approaches are typically employed to balance precision and efficiency. Variational methods first provide preliminary estimates, followed by optimisation of key parameter posterior densities via MCMC. This conserves computational time while ensuring model output consistency and determinism.

4.3. Interpretation and Visualisation of Causal Effects

Bayesian causal models not only enhance computational precision but also improve data interpretability through graphical representations. The posterior distributions inferred by the model can be transformed into confidence interval plots, path strength maps, and marginal effect curves to elucidate the magnitude of effects across various advertising stages. Confidence interval plots clarify parameter estimation uncertainty while demonstrating the significance and robustness of advertising effects; path strength plots illustrate the influence and relative impact of each factor; marginal effect curves describe the non-linear relationships between advertising exposure, click-through rates, and conversions. The benefit of visualisation lies in aiding senior management's comprehension of the model inference process, enabling enterprises to make informed decisions in practical operations. Concurrently, causal decomposition results can be integrated into practical marketing strategies to identify high-value customer segments or critical pathway points. By mapping model outputs to advertising budgets and exposure frequencies, enterprises can quantitatively measure the marginal benefits of different channels, achieving optimal allocation of advertising resources. Such applications hold significant potential for expansion in digital marketing, recommendation system optimisation, and user retention analysis.

4.4. Demonstrating the Model's Value in Advertising Scenarios

The Bayesian causal model exhibits robustness and efficacy when applied to advertising. By pre-setting constraints, it addresses issues of imbalanced training data and funnel leakage in advertising, mitigating the impact of anomalous samples on results while preserving overall trends. Combined with post-update methods to adjust parameter values for new advertising data, it enables real-time monitoring and dynamic management of advertising conversion rates. This methodology optimises multi-tiered advertising strategies: for brand marketing, it distinguishes marginal conversion impacts across exposure frequencies and identifies optimal exposure thresholds; for performance advertising, it compares causal effects across channels to guide budget allocation decisions; in remarketing scenarios, it predicts repeat purchase probabilities for different customers and informs personalised touchpoint strategies.

Overall, the Bayesian causal model combines probabilistic inference with hierarchical structures, balancing robustness and flexibility to provide a quantifiable, interpretable toolkit for advertising decision-making. Its application extends beyond ad effectiveness evaluation to audience segmentation, pricing optimisation, and media mix analysis, offering robust support for scientific and refined digital marketing management.

5. Conclusion

Bayesian causal modelling provides a theoretical foundation for constructing advertising efficacy pathways. This approach utilises prior information and posterior update processes to estimate how advertising effectiveness evolves over time under unknown conditions, revealing the true relationships between exposure, clicks, and conversions. It achieves superior predictive performance and interpretability in big data scenarios characterised by small samples, diverse samples, and high noise. Furthermore, hierarchical modelling and variable inference techniques enable the model to strike a balance between computability and accuracy. The Bayesian causal modelling investigated herein can further enhance the scientific rigour of advertising efficacy evaluation and serve as a reference for optimising advertising deployment strategies. Future research may integrate multi-source data with real-time feedback mechanisms to construct adaptive causal learning systems, thereby enabling continuous optimisation of advertising decisions and data-driven intelligent evolution.

References

1. S. Zhao, Z. Zhang, and H. Zhang, "Bayesian inference of dynamic mediation models for longitudinal data," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 31, no. 1, pp. 14-26, 2024. doi: 10.1080/10705511.2023.2230519
2. T. Tanaka, "Evaluating the Bayesian causal inference model of intentional binding through computational modeling," *Scientific Reports*, vol. 14, no. 1, p. 2979, 2024. doi: 10.1038/s41598-024-53071-7
3. F. Castelletti, and G. Consonni, "Bayesian graphical modeling for heterogeneous causal effects," *Statistics in medicine*, vol. 42, no. 1, pp. 15-32, 2023. doi: 10.1002/sim.9599
4. E. V. Orlova, "Hybrid approach to modeling labor productivity factors: Synthesis of randomized controlled experiments and causal Bayesian networks," *Economics and Mathematical Methods*, vol. 60, no. 1, pp. 108-120, 2024. doi: 10.31857/s0424738824010099
5. S. M. Andersen, S. Chen, and R. Miranda, "Significant others and the self," *Self and identity*, vol. 1, no. 2, pp. 159-168, 2002. doi: 10.1080/152988602317319348
6. L. Novelli, K. Friston, and A. Razi, "Spectral dynamic causal modeling: A didactic introduction and its relationship with functional connectivity," *Network Neuroscience*, vol. 8, no. 1, pp. 178-202, 2024. doi: 10.1162/netn_a_00348
7. R. S. Nickerson, "How we know-and sometimes misjudge-what others know: Imputing one's own knowledge to others," *Psychological bulletin*, vol. 125, no. 6, p. 737, 1999.