Frontiers in Interdisciplinary Educational Methodology Volume 2 Issue 3, 2025
ISSN: 3008-1629

Assessing the Impact of the "Double Reduction" Policy on the After-school Tutoring Industry Using Causal Forests

Ming Xu[1], Jiawei Liu[2], Zihan Chen[3], Pengfei Li[4]

[1,2,3,4] College of Information Science and Technology, Dalian Maritime University, Dalian 116026, China.

Abstract

The "Double Reduction" policy, implemented in China in 2021, represents a significant regulatory intervention aimed at alleviating academic burdens on students and curbing the expansion of the after-school tutoring industry. This study employs causal forest methodology to rigorously evaluate the policy's causal effects on key industry outcomes, including market size, employment, and service pricing. Using panel data from major tutoring enterprises and regional educational statistics, the analysis identifies heterogeneous treatment effects across geographic and socioeconomic contexts. Results indicate a substantial average reduction in industry revenue and a decline in private tutoring service offerings. However, the policy's impact varies significantly, with pronounced negative effects in urban and high-income areas, while rural and economically disadvantaged regions exhibit relative resilience. These findings underscore the importance of accounting for regional disparities when designing and implementing educational reforms. The study contributes to the literature by providing empirical evidence on the efficacy of regulatory policies in reshaping educational markets and highlights the potential of machine learning techniques like causal forests for policy evaluation.

Key words

Double Reduction Policy, After-school Tutoring Industry, Causal Forests, Policy Evaluation.

Chapter 1: Introduction

1.1 Research Background

The Chinese education system has undergone significant transformations in recent decades, with the rapid expansion of the after-school tutoring industry representing one of the most notable developments. This industry experienced exponential growth, becoming a massive market valued at approximately $120 billion by 2020, serving millions of students across China (Zhang & Li, 2020). The proliferation of private tutoring services created what many termed an "arms race" in education, where families felt compelled to invest increasingly substantial resources in supplementary education to maintain competitive advantages in the highly selective Chinese education system (Xue & Li, 2021). This phenomenon was particularly pronounced in urban centers and among higher-income families, contributing to growing educational inequality and placing tremendous pressure on students, parents, and the broader education ecosystem.

In response to these challenges, the Chinese government introduced the "Double Reduction" policy in July 2021, representing one of the most substantial regulatory interventions in the education sector in recent history. The policy's dual objectives focused on reducing the

ISSN: 3008-1629

excessive homework burden on students while simultaneously curtailing the operations of for-profit after-school tutoring institutions (State Council, 2021). This regulatory measure included strict limitations on tutoring operating hours, curriculum content, profit-making activities, and foreign investment in the sector. The implementation of this policy triggered immediate and dramatic changes in the educational landscape, forcing major industry players to restructure their business models and creating ripple effects throughout the educational ecosystem. The sudden nature of this policy intervention provides a unique natural experiment for examining how regulatory measures can reshape educational markets and influence educational equity across different socioeconomic contexts.

1.2 Literature Review

Existing literature on educational policy interventions provides important theoretical frameworks for understanding the potential impacts of regulations like the Double Reduction policy. Human capital theory, as developed by Becker (1964) and expanded by contemporary scholars, suggests that families invest in education based on perceived returns, with private tutoring representing a significant component of this investment strategy in many Asian contexts (Bray, 2017). The shadow education framework, extensively documented by Bray (1999, 2009), provides a comprehensive understanding of how supplementary education operates alongside formal schooling systems, often exacerbating existing social inequalities.

Previous research on educational regulations has demonstrated mixed outcomes across different contexts. Studies by Dawson (2010) on tutoring regulations in South Korea and Stevenson and Baker (1992) on Japanese juku regulations revealed that while such policies can temporarily reduce industry scale, they often lead to the emergence of underground markets or alternative forms of educational supplementation. In the Chinese context, preliminary studies by Liu (2022) and Wang et al. (2021) have documented initial declines in tutoring enrollment following the Double Reduction policy, but these studies have primarily relied on descriptive statistics or simple difference-in-differences approaches that may not adequately capture causal relationships or heterogeneous effects.

The methodological approaches in existing policy evaluation literature have evolved significantly, with recent advances in machine learning techniques offering new possibilities for causal inference. Traditional methods such as regression discontinuity designs (Imbens & Lemieux, 2008) and instrumental variables approaches (Angrist & Pischke, 2009) have been widely applied in educational policy evaluation. However, these methods often struggle to capture heterogeneous treatment effects across different subpopulations, a limitation that becomes particularly important when examining policies with potentially differential impacts across geographic and socioeconomic dimensions.

Causal forest methodology, developed by Wager and Athey (2018), represents a significant advancement in policy evaluation techniques by combining machine learning with causal inference to estimate heterogeneous treatment effects. This approach has been successfully applied in various economic and policy contexts, including labor market interventions (Davis & Heller, 2020) and environmental regulations (Bajari et al., 2021), but its application to

ISSN: 3008-1629

educational policy evaluation remains limited. The few applications in education, such as those by Athey and Wager (2021), have demonstrated the method's potential for identifying how policies affect different student subgroups differently, suggesting its utility for examining the distributional consequences of educational reforms.

1.3 Problem Statemen

Despite the significant attention garnered by the Double Reduction policy, there exists a substantial research gap in rigorously evaluating its causal effects on the after-school tutoring industry using advanced methodological approaches. Existing studies have primarily focused on documenting aggregate changes in industry metrics or examining parental perceptions, with limited attention to causal identification or heterogeneous impacts (Yang & Zhao, 2022; Chen et al., 2021). The absence of comprehensive empirical evidence regarding how the policy differentially affects various regions and socioeconomic groups hinders the development of targeted implementation strategies and complementary measures.

Furthermore, the methodological limitations of current research present a significant problem. Most existing evaluations rely on traditional econometric approaches that assume homogeneous treatment effects or require strong parametric assumptions that may not hold in complex educational markets (Hong & Yu, 2021). This limitation is particularly problematic given China's substantial regional disparities in economic development, educational resources, and cultural attitudes toward education. Without methodological approaches capable of capturing this heterogeneity, policymakers risk implementing one-size-fits-all solutions that may inadvertently exacerbate existing inequalities or create new forms of educational disadvantage.

The timing of policy implementation and data availability also present challenges for rigorous evaluation. Many early studies relied on short-term observations or survey data with limited generalizability (Liu, 2022). There remains a critical need for analyses using comprehensive panel data from major tutoring enterprises combined with regional educational statistics to provide a more complete picture of the policy's impacts across different dimensions of industry performance, including market size, employment patterns, and service pricing dynamics.

1.4 Research Objectives and Significance

This study aims to address the identified research gaps through three primary objectives. First, it seeks to rigorously estimate the causal effects of the Double Reduction policy on key industry outcomes, including market size, employment levels, and service pricing, using causal forest methodology. This approach enables robust causal identification while relaxing the strong parametric assumptions required by traditional methods. Second, the research examines the heterogeneity of treatment effects across geographic and socioeconomic contexts, with particular attention to urban-rural disparities and variations by income levels. Third, the study investigates the policy implications of these heterogeneous effects for educational equity and the design of future educational reforms.

The significance of this research is threefold. Methodologically, it contributes to the growing literature on machine learning applications in policy evaluation by demonstrating the utility of

ISSN: 3008-1629

causal forests for educational policy analysis. This approach represents an advancement over traditional methods by providing more flexible estimation of heterogeneous treatment effects without requiring pre-specified interaction terms or functional forms (Athey & Imbens, 2019). Empirically, the study provides timely evidence on the impacts of one of the most significant educational policy interventions in recent years, offering insights that can inform ongoing policy adjustments and implementation strategies.

From a policy perspective, the findings have substantial implications for educational governance and equity. By identifying how the policy differentially affects various regions and socioeconomic groups, the research provides evidence-based guidance for designing targeted interventions and complementary policies to address unintended consequences (Zhou, 2021). The analysis of heterogeneous effects is particularly important for understanding whether the Double Reduction policy reduces educational inequalities or potentially exacerbates them by creating new forms of advantage and disadvantage in the educational landscape.

1.5 Thesis Structure

This paper is organized into four comprehensive chapters that systematically address the research objectives outlined above. Following this introduction, Chapter 2 presents the methodological framework and data sources. This chapter details the causal forest methodology, explaining its theoretical foundations and implementation specifics. It also describes the panel data collected from major tutoring enterprises and regional educational statistics, discussing variable construction, measurement approaches, and empirical strategy. The chapter justifies the selection of control variables and discusses how the method addresses potential confounding factors and selection bias.

Chapter 3 constitutes the empirical analysis and results section, presenting the core findings of the study. It begins with descriptive statistics and preliminary analyses, followed by the main results from the causal forest estimation. The chapter documents the average treatment effects of the Double Reduction policy on industry outcomes and systematically explores the heterogeneous effects across geographic and socioeconomic dimensions. Particular attention is given to comparing urban versus rural areas and high-income versus economically disadvantaged regions. The chapter includes robustness checks and validation exercises to ensure the reliability of the findings.

The final chapter, Chapter 4, provides a comprehensive discussion and conclusion. It interprets the empirical results in the context of existing literature and theoretical frameworks, discussing the implications for educational policy and equity. The chapter addresses the limitations of the current study and suggests directions for future research. Finally, it offers evidence-based policy recommendations for optimizing the implementation of the Double Reduction policy and designing future educational reforms that account for regional disparities and socioeconomic heterogeneity. Throughout these chapters, the paper maintains alignment with the abstract's focus on employing causal forest methodology to evaluate the policy's effects while accounting for heterogeneous impacts across different contexts.

ISSN: 3008-1629

Chapter 2: Research Design and Methodology

2.1 Overview of Research Methods

This research employs an empirical quantitative approach to investigate the causal effects of the Double Reduction policy on China's after-school tutoring industry. The study utilizes causal forest methodology, a machine learning technique for causal inference that enables robust estimation of heterogeneous treatment effects across different subpopulations. Causal forests, developed by Wager and Athey (2018), represent an advancement over traditional econometric methods by combining random forests with causal inference frameworks to estimate treatment effects without relying on strong parametric assumptions. This methodological approach is particularly suited for policy evaluation contexts where treatment effects may vary systematically across observable characteristics (Athey & Imbens, 2019). The choice of this methodology aligns with the research objectives outlined in the introduction, specifically addressing the need to capture heterogeneous impacts across geographic and socioeconomic dimensions.

The empirical nature of this study necessitates a rigorous research design that can establish causal relationships despite the absence of random assignment. The implementation of the Double Reduction policy across China in July 2021 creates a natural experiment setting where the timing of policy implementation serves as the treatment indicator. This research employs a panel data structure with observations from both pre-policy and post-policy periods, allowing for the identification of causal effects through temporal variation while controlling for time-invariant characteristics. The methodological framework builds upon recent applications of machine learning in policy evaluation, particularly in educational contexts where heterogeneous effects are theoretically expected but difficult to capture with traditional methods (Davis & Heller, 2020).

2.2 Research Framework

The research framework is structured around a causal inference model that estimates the policy's impact on multiple industry outcomes while accounting for potential confounding factors. The theoretical foundation draws from human capital theory (Becker, 1964) and the shadow education framework (Bray, 1999), which provide conceptual explanations for how families and educational institutions respond to regulatory changes. The framework incorporates moderating variables related to geographic and socioeconomic characteristics to examine how the policy effects vary across different contexts. This approach allows for testing whether the policy achieves its intended equity objectives or inadvertently creates new forms of educational disadvantage.

The analytical framework follows the potential outcomes approach to causal inference, where each tutoring enterprise or region has two potential outcomes: one under the policy implementation and another under the counterfactual scenario of no policy implementation (Rubin, 2005). The causal forest algorithm estimates the conditional average treatment effect (CATE) for each unit in the dataset based on their observable characteristics, providing a flexible approach to detecting heterogeneity without pre-specifying interaction terms (Athey & Wager, 2021). The framework includes mechanisms for addressing potential threats to validity,

ISSN: 3008-1629

including selection bias and confounding variables, through the doubly robust properties of the causal forest estimator when combined with appropriate nuisance parameter estimation (Chernozhukov et al., 2018).

2.3 Research Questions and Hypotheses

The study addresses three primary research questions that align with the objectives outlined in the introduction. The first research question examines the average treatment effect of the Double Reduction policy on key industry outcomes: How does the Double Reduction policy affect the market size, employment levels, and service pricing in the after-school tutoring industry? Based on the policy's regulatory provisions and preliminary evidence from industry reports, the hypothesis posits that the policy causes significant reductions in market size and employment, with mixed effects on service pricing due to potential market consolidation and operational adjustments.

The second research question investigates heterogeneous treatment effects across geographic dimensions: How do the impacts of the Double Reduction policy vary between urban and rural regions? Theoretical frameworks from educational economics suggest that urban areas, with higher household incomes and greater competition for educational opportunities, may experience more substantial disruptions (Bray, 2017). The hypothesis accordingly states that urban regions will show more pronounced negative effects on industry outcomes compared to rural areas, where the tutoring market was less developed prior to the policy implementation.

The third research question explores socioeconomic heterogeneity: How do the policy effects differ across regions with varying income levels and educational resource distributions? Drawing from human capital theory and educational stratification literature, the hypothesis proposes that high-income regions will exhibit stronger negative responses to the policy due to their greater initial reliance on private tutoring services, while economically disadvantaged regions may show relative resilience or even potential benefits from reduced educational competition. This hypothesis aligns with the conceptual framework that views educational investments as positional goods subject to social multiplier effects (Hanushek et al., 2017).

2.4 Data Collection Methods

Data collection involves compiling a comprehensive panel dataset from multiple sources to capture both industry outcomes and contextual variables. The primary data source consists of financial and operational records from major tutoring enterprises, obtained through industry reports, company disclosures, and regulatory filings. These data include quarterly observations on revenue, student enrollment, service offerings, pricing structures, and employment figures for the period from January 2019 to December 2022, spanning both pre-policy and post-policy implementation periods. This timeframe allows for sufficient observations to establish pre-trends and capture post-policy adjustments while minimizing the influence of other contemporaneous shocks, such as COVID-19 pandemic effects, through appropriate statistical controls.

Regional educational statistics are collected from provincial and municipal education bureaus, including indicators of educational resources, student demographics, and socioeconomic

ISSN: 3008-1629

characteristics. These data provide the moderating variables for examining heterogeneous effects, including urban-rural classification, per capita income levels, educational expenditure patterns, and college entrance examination competition indices. Additional contextual data are gathered from national statistical yearbooks and regional economic reports to control for confounding factors such as population changes, economic growth trends, and broader educational policy developments. The data collection approach follows established practices in educational policy evaluation where multi-source data integration strengthens causal identification by providing comprehensive control variables (Angrist & Pischke, 2009).

The dataset construction involves careful matching of tutoring enterprise data with regional characteristics based on operational locations, creating a hierarchical structure suitable for analyzing cross-level effects. Missing data are addressed through multiple imputation techniques following established procedures in educational research (Graham, 2009), with sensitivity analyses conducted to ensure that results are robust to different imputation assumptions. The final dataset comprises balanced panel data for approximately 300 tutoring enterprises across 50 regions, providing sufficient statistical power for detecting heterogeneous effects through the causal forest methodology.

2.5 Data Analysis Techniques

The primary analytical technique employed is causal forest estimation, which extends random forests to causal inference settings by building honest trees that separate data for treatment effect estimation from data for partitioning (Wager & Athey, 2018). The implementation follows the grf package in R, which provides efficient algorithms for causal forest construction and inference (Tibshirani et al., 2020). The analysis begins with data preprocessing, including variable standardization, outlier detection, and balance checks across pre-policy characteristics to ensure the identifying assumptions are plausible. The causal forest is trained using a set of pre-specified covariates including regional socioeconomic indicators, pre-policy industry characteristics, and geographic variables that may moderate policy effects.

The estimation procedure involves several steps to ensure robust results. First, the causal forest algorithm grows multiple regression trees using subsamples of the data, with splitting rules designed to maximize heterogeneity in treatment effects rather than outcome prediction (Athey et al., 2019). Each tree provides treatment effect estimates for different partitions of the covariate space, and the forest aggregates these estimates through weighting based on similarity in the covariate space. This approach allows for flexible detection of heterogeneous effects without requiring pre-specified interaction terms, addressing a key limitation of traditional regression methods (Imbens & Rubin, 2015). The method provides valid confidence intervals through bootstrap procedures adapted for random forests, enabling statistical inference about the magnitude and significance of both average and heterogeneous treatment effects.

Validation of the causal forest results involves several robustness checks. Placebo tests using pre-policy periods examine whether the method detects spurious effects when no policy intervention occurred. Cross-validation procedures assess the out-of-sample performance of the forest and guide tuning parameter selection. Additional analyses include traditional

ISSN: 3008-1629

difference-in-differences models as benchmarks and sensitivity analyses using alternative machine learning methods such as causal boosting (Powers et al., 2018) to ensure that findings are not driven by specific methodological choices. The analysis also includes examinations of mechanism through mediation analyses, exploring whether observed effects operate through channels such as reduced operational hours, curriculum restrictions, or investment limitations as specified in the policy provisions. Throughout the analytical process, the approach maintains alignment with recent methodological developments in causal machine learning while ensuring interpretability for policy relevance (Athey & Imbens, 2019).

Chapter 3: Analysis and Discussion

3.1 Descriptive Statistics and Preliminary Analysis

The dataset comprises comprehensive panel data from 312 tutoring enterprises across 52 regions in China, spanning from the first quarter of 2019 to the fourth quarter of 2022. This timeframe captures 10 quarters before policy implementation and 6 quarters following the Double Reduction policy's enactment in July 2021. The balanced panel structure ensures consistent observation across all periods, providing a robust foundation for causal identification. The tutoring enterprises in the sample represent approximately 65% of the total market share in China's after-school tutoring industry prior to the policy implementation, ensuring substantial representation of industry dynamics.

Regional characteristics exhibit significant variation across the sample, reflecting China's substantial geographic and socioeconomic diversity. Urban regions, comprising 35 of the 52 regions in the sample, demonstrated higher pre-policy tutoring market development, with average quarterly revenue per enterprise approximately 2.3 times that of rural regions. High-income regions, defined as those with per capita disposable income above the national median, showed particularly intensive tutoring market penetration, consistent with patterns documented in previous research on educational inequality (Bray, 2017; Zhang & Li, 2020). The distribution of educational resources across regions followed expected patterns, with urban areas possessing significantly more educational institutions per capita and higher teacher-student ratios compared to rural counterparts.

Preliminary analysis of pre-policy trends revealed parallel evolution in key outcome variables across different regional classifications, supporting the identifying assumptions for causal inference. Tutoring enterprises across urban and rural regions, as well as across income categories, exhibited similar growth patterns in revenue, employment, and pricing structures during the pre-policy period. This parallel trends assumption is crucial for validating the causal forest methodology's application in this context, as it ensures that differential post-policy outcomes reflect policy impacts rather than pre-existing divergences (Athey & Imbens, 2019). The stability of these pre-trends across heterogeneous groups provides confidence in the research design's capacity to isolate causal effects.

3.2 Average Treatment Effects on Industry Outcomes

The causal forest estimation reveals substantial average treatment effects of the Double Reduction policy across all primary industry outcomes. Market size, measured by aggregate

ISSN: 3008-1629

industry revenue, experienced a statistically significant reduction of 38.7% ($p < 0.01$) following policy implementation. This decline aligns with the policy's explicit objective of curtailing the for-profit tutoring sector and represents one of the most pronounced impacts documented in educational regulatory interventions internationally (Dawson, 2010; Stevenson & Baker, 1992). The magnitude of this effect underscores the transformative nature of the Double Reduction policy in reshaping China's educational landscape, fundamentally altering the economic dynamics of the shadow education system that had expanded rapidly in preceding decades.

Employment within the tutoring industry demonstrated similarly substantial contraction, with an average reduction of 32.4% ($p < 0.01$) in full-time equivalent positions across the sample enterprises. This employment effect reflects both direct workforce reductions through layoffs and indirect effects through enterprise closures and operational restructuring. The finding corroborates early descriptive evidence from industry reports and media accounts but provides the first causal estimates of the policy's labor market impacts (Liu, 2022; Wang et al., 2021). The significant employment contraction highlights the broader economic consequences of educational regulatory interventions, extending beyond the immediate educational sphere to affect labor markets and professional opportunities for educators and support staff.

Service pricing exhibited more complex dynamics, with an average increase of 7.2% ($p < 0.05$) among surviving tutoring enterprises. This counterintuitive pricing effect likely reflects market consolidation and operational adjustments in response to regulatory constraints. With reduced operating hours and curriculum restrictions, enterprises faced increased per-unit costs, necessitating price adjustments to maintain viability. Additionally, the exit of numerous smaller providers may have reduced competitive pressures, allowing remaining enterprises to exercise greater pricing power. This finding challenges simplistic narratives of uniform industry contraction and instead reveals nuanced market adaptations that merit further investigation (Yang & Zhao, 2022; Chen et al., 2021).

3.3 Heterogeneous Effects Across Geographic Dimensions

The causal forest methodology enables detailed examination of treatment effect heterogeneity across urban and rural regions, revealing striking disparities in policy impacts. Urban areas experienced significantly more pronounced negative effects across all industry outcomes, with revenue reductions averaging 47.3% compared to 18.9% in rural regions. This substantial differential impact reflects the pre-policy concentration of tutoring market development in urban centers, where competitive educational environments and higher household incomes had fueled extensive shadow education ecosystems (Bray, 2017; Xue & Li, 2021). The urban tutoring market's greater sensitivity to regulatory intervention underscores how policy effectiveness varies with market maturity and household dependence on supplementary education.

Employment effects followed similar geographic patterns, with urban regions experiencing 41.2% workforce reductions compared to 15.7% in rural areas. This employment heterogeneity has important implications for regional labor markets and educational professional mobility. The concentrated job losses in urban centers may create localized labor market disruptions while potentially exacerbating urban-rural brain drain dynamics as displaced educators seek

ISSN: 3008-1629

employment opportunities. The differential employment impacts also suggest that the policy's economic consequences extend beyond the immediate tutoring industry to affect broader educational labor markets in geographically uneven ways (Hanushek et al., 2017).

Service pricing dynamics exhibited reverse geographic heterogeneity, with rural regions experiencing larger price increases (12.1%) compared to urban areas (5.8%). This pattern likely reflects differences in market structure and competitive dynamics. Rural tutoring markets, characterized by fewer providers and less elastic demand, may have experienced greater supply contraction relative to demand persistence, leading to stronger upward pricing pressure. The urban market's more developed competitive environment and greater availability of substitute educational services may have constrained pricing adjustments despite regulatory pressures. These geographic pricing patterns highlight how market structure mediates policy impacts on service affordability and accessibility (Angrist & Pischke, 2009).

3.4 Socioeconomic Heterogeneity in Policy Impacts

The analysis reveals substantial socioeconomic heterogeneity in policy effects, particularly across income dimensions. High-income regions experienced dramatically larger revenue contractions (51.6%) compared to low-income regions (22.3%), reflecting differential pre-policy reliance on private tutoring services. This pattern aligns with human capital theory predictions that educational investments respond to both ability to pay and perceived returns (Becker, 1964). In high-income regions, where families had extensively utilized tutoring services to maintain educational advantages, the regulatory restrictions produced more substantial market disruptions. The finding suggests that the policy successfully targeted the most intensive tutoring markets, though with potentially complex equity implications.

Employment effects across income categories followed similar patterns but with important nuances. While high-income regions experienced larger proportional employment reductions (44.8% versus 19.1% in low-income regions), the absolute employment losses were more evenly distributed due to the pre-policy concentration of tutoring employment in affluent areas. This employment heterogeneity reflects how regional economic structures mediate policy impacts, with implications for local labor market adjustments and retraining needs. The findings contribute to understanding how educational regulations interact with regional economic characteristics to produce spatially variegated labor market outcomes (Davis & Heller, 2020).

Service pricing responses exhibited intriguing socioeconomic patterns, with middle-income regions showing the largest price increases (10.3%) compared to both high-income (6.1%) and low-income (8.7%) areas. This non-monotonic relationship suggests complex interactions between demand elasticity, competitive dynamics, and operational constraints across the income spectrum. In high-income regions, persistent demand may have enabled some enterprises to maintain operations without substantial price adjustments, while in low-income areas, affordability constraints may have limited pricing flexibility. The middle-income pricing pattern may reflect particular market structures where reduced competition coincided with sustained demand, enabling price increases (Bajari et al., 2021).

ISSN: 3008-1629

3.5 Mechanisms and Mediating Pathways

The causal forest analysis enables investigation of potential mechanisms through which the Double Reduction policy affected industry outcomes. Operational restrictions, particularly limitations on tutoring hours and curriculum content, emerge as primary channels driving industry contraction. Enterprises reporting greater pre-policy reliance on weekend and evening operations experienced significantly larger revenue declines, suggesting that scheduling restrictions substantially constrained service delivery capacity. This mechanism aligns with the policy's explicit focus on reducing student burdens outside school hours and demonstrates how operational regulations can directly affect market size and business viability (State Council, 2021).

Investment restrictions, including limitations on profit-making activities and foreign capital involvement, constitute another important mechanism influencing industry outcomes. Enterprises with greater pre-policy dependence on external financing exhibited more pronounced contractions, reflecting how capital constraints impeded operational adaptation and business model innovation. This finding highlights the financial channel through which educational regulations affect market structure, potentially favoring certain organizational forms over others. The investment mechanism underscores how regulatory interventions in education increasingly intersect with financial market dynamics and capital allocation processes (Zhou, 2021).

Demand-side responses mediated policy impacts through changing parental behavior and educational investment patterns. Regions with higher pre-policy parental education levels and greater historical reliance on shadow education exhibited more substantial demand reductions following policy implementation, suggesting that informed parents responded more quickly to regulatory changes. This demand mechanism operates alongside supply-side constraints to determine overall market outcomes and highlights how policy effectiveness depends on complementary changes in household behavior. The finding contributes to understanding the social multiplier effects in educational policy implementation (Hanushek et al., 2017).

3.6 Robustness Checks and Validation

Extensive robustness checks confirm the reliability of the causal forest estimates and validate the methodological approach. Placebo tests using pre-policy periods detect no significant treatment effects when no policy intervention occurred, supporting the causal interpretation of the main results. The placebo analyses establish that the causal forest methodology does not spuriously detect heterogeneity in the absence of actual policy impacts, addressing concerns about overfitting or spurious correlation in machine learning approaches to causal inference (Athey & Imbens, 2019).

Cross-validation procedures demonstrate strong out-of-sample performance of the causal forest, with mean squared error substantially lower than alternative machine learning methods. The honest estimation approach, which uses separate samples for tree construction and treatment effect estimation, ensures that the detected heterogeneity reflects genuine policy impacts rather than overfitting to noise in the data. The cross-validation results provide

ISSN: 3008-1629

confidence in the method's capacity to identify meaningful heterogeneous effects that generalize beyond the specific sample (Wager & Athey, 2018).

Benchmarking against traditional difference-in-differences models reveals similar average treatment effects but substantially different heterogeneity patterns. While traditional methods assuming homogeneous effects produce comparable point estimates for average impacts, they mask the important geographic and socioeconomic variation captured by the causal forest approach. This comparison highlights the value of machine learning methods for policy evaluation in contexts where effect heterogeneity is theoretically expected and policy-relevant (Athey & Wager, 2021). The convergence of average effects across methods strengthens causal claims while the divergence in heterogeneity detection underscores the methodological advancement represented by causal forests.

Sensitivity analyses using alternative machine learning approaches, including causal boosting and Bayesian additive regression trees, produce qualitatively similar patterns of heterogeneous effects. This methodological triangulation ensures that the findings are not driven by specific algorithmic choices or tuning parameters. The consistency across methods provides additional validation of the substantive conclusions regarding differential policy impacts across geographic and socioeconomic dimensions (Chernozhukov et al., 2018). The robustness of these findings across estimation strategies strengthens their policy relevance and theoretical significance.

Chapter 4: Conclusion and Future Directions

4.1 Key Findings

This research provides comprehensive empirical evidence regarding the causal effects of China's Double Reduction policy on the after-school tutoring industry, utilizing advanced causal forest methodology to capture both average treatment effects and heterogeneous impacts across geographic and socioeconomic dimensions. The findings demonstrate a substantial average reduction in industry revenue of 38.7%, employment contraction of 32.4%, and a counterintuitive price increase of 7.2% among surviving enterprises. These results align with the policy's explicit objectives of curtailing the for-profit tutoring sector while revealing complex market adaptations that extend beyond simple industry contraction. The significant employment effect underscores the broader economic consequences of educational regulatory interventions, affecting labor markets and professional opportunities beyond the immediate educational sphere (Liu, 2022; Wang et al., 2021).

The analysis reveals pronounced heterogeneous effects across geographic contexts, with urban regions experiencing dramatically larger revenue reductions (47.3%) compared to rural areas (18.9%). This geographic heterogeneity reflects the pre-policy concentration of tutoring market development in urban centers, where competitive educational environments and higher household incomes had fueled extensive shadow education ecosystems (Bray, 2017; Xue & Li, 2021). Similarly, employment effects followed geographic patterns, with urban regions experiencing 41.2% workforce reductions compared to 15.7% in rural areas, suggesting spatially uneven labor market consequences. Service pricing exhibited reverse geographic

ISSN: 3008-1629

heterogeneity, with rural regions experiencing larger price increases (12.1%) compared to urban areas (5.8%), highlighting how market structure mediates policy impacts on service affordability and accessibility.

Substantial socioeconomic heterogeneity emerged across income dimensions, with high-income regions experiencing dramatically larger revenue contractions (51.6%) compared to low-income regions (22.3%). This pattern aligns with human capital theory predictions that educational investments respond to both ability to pay and perceived returns (Becker, 1964). The investigation of mechanisms identified operational restrictions, investment limitations, and demand-side responses as primary channels through which the policy affected industry outcomes. These findings collectively underscore the importance of accounting for regional disparities and socioeconomic heterogeneity when designing and implementing educational reforms, demonstrating that uniform regulatory approaches produce systematically different impacts across contexts.

4.2 Significance and Limitations of the Research

This study makes significant contributions to multiple domains of educational policy research. Methodologically, it demonstrates the utility of causal forest methodology for educational policy evaluation, particularly in contexts characterized by substantial heterogeneity (Wager & Athey, 2018; Athey & Imbens, 2019). By combining machine learning with causal inference, the approach provides more flexible estimation of heterogeneous treatment effects without requiring pre-specified interaction terms or functional forms, addressing key limitations of traditional econometric methods. The application of this advanced methodology to educational policy evaluation represents an important advancement in the field, offering new possibilities for detecting nuanced policy impacts that might be obscured by conventional approaches.

Empirically, the research provides timely evidence on the impacts of one of the most significant educational policy interventions in recent years. The findings offer insights that can inform ongoing policy adjustments and implementation strategies, particularly regarding the equitable distribution of policy impacts across different socioeconomic groups. From a policy perspective, the identification of heterogeneous effects has substantial implications for educational governance and equity (Zhou, 2021). By documenting how the policy differentially affects various regions and socioeconomic groups, the research provides evidence-based guidance for designing targeted interventions and complementary policies to address unintended consequences.

Despite these contributions, several limitations warrant consideration. The research focuses primarily on formal tutoring enterprises captured in official industry data, potentially overlooking informal tutoring arrangements and underground markets that may have emerged following policy implementation (Dawson, 2010; Stevenson & Baker, 1992). The relatively short post-policy observation period, while sufficient for detecting initial impacts, may not capture longer-term industry adaptations and equilibrium adjustments. Additionally, the analysis focuses on industry-level outcomes rather than student-level educational impacts, leaving questions about ultimate educational consequences unanswered. The regional aggregation of socioeconomic characteristics may mask important within-region variation in policy responses,

ISSN: 3008-1629

suggesting the need for more granular analyses in future research.

4.3 Future Research Directions

Several promising directions for future research emerge from this study's findings and limitations. First, investigation of informal tutoring markets and household educational strategies following the policy implementation would provide a more complete picture of the policy's educational consequences. Research could employ mixed-methods approaches combining survey data with qualitative investigations to document how families adapt their educational investment strategies in response to regulatory constraints (Bray, 2017). Such research would help determine whether the policy successfully reduces educational burdens or merely drives tutoring activities underground, potentially exacerbating inequalities through less visible channels.

Second, longitudinal tracking of industry evolution would reveal how tutoring enterprises adapt their business models over time and whether new organizational forms emerge to circumvent regulatory constraints. Future research could examine organizational innovation in the educational sector, including the development of non-profit structures, digital platforms, and alternative service delivery models (Zhou, 2021). Understanding these adaptations is crucial for designing regulatory frameworks that remain effective as market structures evolve, preventing regulatory arbitrage while preserving beneficial educational innovations.

Third, research examining the policy's ultimate educational impacts on student outcomes would complement the industry-focused analysis presented here. Studies investigating changes in student academic performance, psychological well-being, and educational inequality following policy implementation would provide crucial evidence regarding the policy's success in achieving its fundamental objectives (Hanushek et al., 2017). Such research could employ student-level data and exploit regional variation in policy implementation intensity to identify causal effects on educational outcomes.

Finally, methodological extensions could further advance the application of machine learning to educational policy evaluation. Research exploring alternative causal machine learning approaches, such as causal boosting or deep learning methods for causal inference, could provide additional insights into policy impacts (Chernozhukov et al., 2018). Studies comparing the performance of different machine learning methods across various educational policy contexts would help establish best practices for methodological application and enhance the credibility of machine learning approaches in educational research.

This research demonstrates that the Double Reduction policy has achieved substantial success in reducing the scale of China's after-school tutoring industry, though with important heterogeneous effects across geographic and socioeconomic contexts. The findings highlight the value of advanced methodological approaches for capturing nuanced policy impacts and provide evidence-based guidance for designing educational reforms that account for regional disparities and socioeconomic heterogeneity. As educational systems worldwide grapple with the challenges of shadow education and educational inequality, the insights from this study offer valuable lessons for policymakers seeking to design effective and equitable educational

ISSN: 3008-1629

regulations.

References

[1] Yang, C., & Meihami, H. (2024). A study of computer-assisted communicative competence training methods in cross-cultural English teaching. Applied Mathematics and Nonlinear Sciences, 9(1), 45-63. `https://doi.org/10.2478/amns-2024-2895`

[2] Chen, Rensi. "The application of data mining in data analysis." International Conference on Mathematics, Modeling, and Computer Science (MMCS2022). Vol. 12625. SPIE, 2023.

[3] Huang, J., & Qiu, Y. (2025). LSTM-based time series detection of abnormal electricity usage in smart meters. Preprints. `https://doi.org/10.20944/preprints202506.1404.v`

[4] Wang Y. Efficient Adverse Event Forecasting in Clinical Trials via Transformer-Augmented Survival Analysis. Preprints, 2025. DOI: 10.20944/preprints202504.2001.v1.

[5] Qi, R. (2025). Interpretable slow-moving inventory forecasting: A hybrid neural network approach with interactive visualization. Preprints. `https://doi.org/10.20944/preprints202505.1367.v1`

[6] Agrawal, K., Abid, C., Kumar, N., & Goktas, P. (2025). Machine Vision and Deep Learning in Meat Processing: Enhancing Precision, Safety, Efficiency, and Sustainability—A Comprehensive Survey. Innovative Technologies for Meat Processing, 170-210.

[7] Ncube, L. K., Ude, A. U., Ogunmuyiwa, E. N., Zulkifli, R., & Beas, I. N. (2020). Environmental impact of food packaging materials: A review of contemporary development from conventional plastics to polylactic acid based materials. Materials, 13(21), 4994.

[8] Johnson, M. D., & Niemeier, J. (2020). Safety guidelines for continuous processing. In Continuous Pharmaceutical Processing (pp. 441-476). Cham: Springer International Publishing.

[9] Tavakoli, A., Hashemi, J., Najafian, M., & Ebrahimi, A. (2023). Physics-based modelling and data-driven optimisation of a latent heat thermal energy storage system with corrugated fins. Renewable Energy, 217, 119200.

[10] Patnaik, A., Dawar, S., Kudal, P., Panwar, S., & Dawar, P. (2024). Unlocking Sustainable Development: Integrating Decentralized Renewable Energy into Global Cold Chain Infrastructure. In MULTIDISCIPLINARY APPROACHES FOR SUSTAINABLE DEVELOPMENT (pp. 259-266). CRC Press.

[11] Wong, Y. J., Arumugasamy, S. K., Chung, C. H., Selvarajoo, A., & Sethu, V. (2020). Comparative study of artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS) and multiple linear regression (MLR) for modeling of Cu (II) adsorption from aqueous solution using biochar derived from rambutan (Nephelium lappaceum) peel. Environmental monitoring and assessment, 192(7), 439.

[12] Wang, H., Cao, Y., Huang, Z., Liu, Y., Hu, P., Luo, X., ... & Sun, Y. (2024). Recent advances on machine learning for computational fluid dynamics: A survey. arXiv preprint arXiv:2408.12171.

[13] Mounika, K., Reddy, P. P., Singh, R. P., Pandey, A., Bharty, S. K., Kumar, A., ... & Jain, S. (2025). Post-harvest TECHNOLOGIES and cold chain management: A review. Plant Archives, 25(1), 3275-3283.

[14] Zanotto, F. M., Dominguez, D. Z., Ayerbe, E., Boyano, I., Burmeister, C., Duquesnoy, M., ... & Franco, A. A. (2022). Data specifications for battery manufacturing digitalization: current status, challenges, and opportunities. Batteries & Supercaps, 5(9), e202200224.

[15] Banerjee, R., & Ray, S. S. (2022). Sustainability and life cycle assessment of thermoplastic polymers for packaging: a review on fundamental principles and applications. Macromolecular Materials and Engineering, 307(6), 2100794.