Machine-assisted Content Creation on Peer Production Platforms

Abstract

We examine how AI-powered translation technology shape multilingual content creation on peer production platforms. Drawing on Wikipedia's integration of Google Translate in January 2019, we study both total impact and distributional effects across more than 100 language editions. Our empirical analysis shows that this technological improvement substantially increased translation volume by 149% while maintaining content quality and reader engagement. However, the benefits are unequally shared: languages with large content bases and active communities see much larger increases than smaller editions. Analysis of biographical and geographical content reveals progress in addressing representation gaps, though substantial disparities persist. Our results underline how technological changes and platform structures intersect to shape knowledge production, with implications for platform design and leveraging AI technology for knowledge democratization.

Keywords: machine translation, artificial intelligence, digital language divide, multilingual content, peer production platforms

1 Introduction

Language barriers impose substantial costs on knowledge sharing and information exchange, creating unequal access to online information and hindering offline social interactions (Egger and Lassmann, 2012; Ginsburgh and Weber, 2020). This "language divide" is especially pronounced within collaborative platforms that aim to democratize information access. Despite its role as a major source of open knowledge, Wikipedia exhibits stark content disparities across languages (Miquel-Ribé and Laniado, 2020; Johnson and Lescak, 2022). For example, while English Wikipedia contains over 6 million articles, Bengali Wikipedia—representing the sixth most spoken native language globally—has only about 100,000 articles. Such disparities between major world languages highlight how linguistic barriers restrict both information access and exposure to diverse perspectives, creating persistent gaps in global knowledge sharing (Nagaraj, 2022; Pereira et al., 2024).

The implications of these language barriers extend beyond Wikipedia to fundamental questions of information access in the digital age, spanning social media, e-commerce sites, educational resources, and professional networks. Research demonstrates that differential access to online information affects a wide range of social and economic outcomes, from consumer behavior (Hinnosaar et al., 2023) and financial decision-making (Xu and Zhang, 2013) to scientific knowledge production (Thompson and Hanley, 2018) and judicial decision-making (Thompson et al., 2024). As digital platforms become central channels for information dissemination, linguistic constraints can amplify existing socioeconomic disparities and limit the potential benefits of digital technologies.

Recent AI innovations, particularly in machine translation, have made addressing these pervasive constraints increasingly feasible. Unlike earlier rule-based systems, modern neural machine translation can process semantic context and linguistic nuance, enabling more natural translations at scale. This technological capability to translate content across languages could theoretically reduce information barriers and expand multilingual digital infrastructure

(Goldfarb and Tucker, 2019; Agrawal et al., 2022). However, realizing these benefits in practice may depend on existing platform structures, community resources, and the systematic inequalities embedded within digital ecosystems.

In this study, we examine how machine translation technology shapes multilingual content creation on peer production platforms. We investigate both the aggregate impact and distributional consequences of reduced translation barriers across Wikipedia's diverse language landscape. Our research design leverages Wikipedia's integration of Google Translate into its Content Translation system in January 2019 as a natural experiment, exploiting variation in translation technology access across languages and time.

While machine translation has proven effective for content consumption, its application to content production presents distinct challenges in maintaining accuracy and quality. Wikipedia's integration of Google Translate provides an ideal setting to study these challenges, as it combines volunteer-driven content creation with structured human-machine collaboration in translation workflows. This integration enables us to analyze how combining editorial expertise with machine translation technology affects both the efficiency and quality of content creation in a decentralized platform.

Using a difference-in-differences empirical strategy, we examine two key dimensions of machine translation's impact. First, we analyze overall effects on content supply and demand: measuring changes in translation volume and quality on the supply side while assessing reader engagement patterns on the demand side. Second, we investigate distributional consequences by examining how benefits are shared across languages and content types, with particular attention to systematic coverage gaps.

Our analysis reveals that machine translation has a substantial impact on content creation while uncovering important distributional effects. The integration of Google Translate led to a 149% increase in translation volume, while maintaining content quality and generating sustained reader engagement, with translated articles receiving over 80 monthly views on average. However, these gains are concentrated among languages with larger existing

content bases and active communities. While English increasingly serves as a central source language, we observe greater diversity in target languages receiving translations. Analysis of specific content domains shows progress in addressing representation gaps, particularly in biographical content and geographic coverage of linguistically diverse regions, though substantial disparities persist.

These findings advance our understanding of how AI-powered technology affects decentralized knowledge production. While machine translation reduces content creation costs, its benefits depend heavily on existing platform structures and community resources. This interaction between technological capability and structural inequalities has important implications for platform design and knowledge democratization efforts. Our results suggest that addressing digital divides requires combining technological solutions with targeted support for under-resourced communities.

The paper proceeds as follows: Section 2 reviews background and literature on digital platforms, language barriers, and machine translation technology. Section 3 describes Wikipedia's Content Translation system and the integration of Google Translate as our empirical setting. Section 4 details our data sources and empirical strategy that leverages this natural experiment. Section 5 presents findings on both aggregate effects and distributional consequences across languages and content types. Section 6 concludes with implications for platform design and knowledge democratization efforts.

2 Background and Literature

Our study examines how artificial intelligence affects content creation costs in digital platforms. This focus is particularly timely given recent advances in AI-powered generative capabilities. While existing research has primarily examined machine translation for content consumption (such as reading foreign websites or facilitating cross-border communication),

we analyze its role in content creation - a distinct challenge requiring quality control and integration with human workflows.

Research has extensively documented knowledge gaps in digital content across gender, geographic regions, and knowledge domains (Halavais and Lackaff, 2008; Graham et al., 2014; Wagner et al., 2015, 2016). Studies show that while organized editing campaigns can enhance content coverage, these efforts often struggle to achieve lasting impact (Zhu et al., 2020; Langrock and González-Bailón, 2022). Despite language barriers that present fundamental constraints on content creation and distribution, the linguistic dimension of digital knowledge gaps remains relatively understudied (Miquel-Ribé and Laniado, 2020; Johnson and Lescak, 2022).

The literature on socio-technical systems demonstrates that platform architecture fundamentally shapes information distribution and accessibility. Work by Parker (2016) and Constantinides et al. (2018) shows that digital platforms are not merely technical infrastructure but complex systems where technology interacts with social dynamics. The concept of 'platform affordance' explains how technological features can either democratize knowledge or reinforce existing gaps, with effects often amplified by technical constraints and social structures (Forman et al., 2008; Aaltonen and Seiler, 2016). This framework helps us understand how technological changes like improved translation tools may have heterogeneous effects across different user communities.

Finally, the evolution of machine translation technology, particularly through neural networks, represents a significant advancement in addressing linguistic barriers. Modern approaches using transformer architectures, exemplified by Google Translate, achieve substantially lower error rates by processing entire sentences and incorporating global context (Bahdanau et al., 2014; Vaswani et al., 2017). These improvements in neural machine translation enable more accurate and nuanced translations (Wu et al., 2016), potentially reducing the costs of cross-language content creation. The technical progress extends beyond accuracy improvements to enhance the accessibility and inclusiveness of digital information systems

(Turovsky, 2016).

However, empirical evidence on machine translation's real-world impacts remains limited. Despite extensive documentation of language barriers in international trade and other domains (Egger and Lassmann, 2012, 2015), and widespread adoption across business, legal, and healthcare sectors, real-world analyses of its effects are only beginning to emerge. Two notable studies have examined different aspects of machine translation's impact: Brynjolfsson et al. (2019) shows that eBay's machine translation system significantly increased cross-border e-commerce activities by reducing language barriers, while Yilmaz et al. (2023) demonstrates how Google's neural network-based translation system led to a measurable substitution of human translators in online labor markets.

Our study bridges these literature streams by examining how machine translation technology affects content creation in collaborative platforms. We analyze both aggregate productivity effects and distributional consequences, providing evidence on how technological change interacts with existing platform structures to shape content generation across languages. This investigation advances our understanding of the role that artificial intelligence plays in democratizing the creation of multilingual content.

3 Empirical Setting

Wikipedia's Content Translation system provides an ideal empirical setting to examine how machine translation technology affects multilingual content production. Traditionally, creating Wikipedia articles required editors to write content from scratch - a time-intensive process demanding substantial language expertise. To reduce these barriers, Wikipedia introduced the Content Translation system in 2015, providing a structured environment for translating existing articles across language editions. The system features a side-by-side interface that displays source and target content, enabling editors to work systematically while

preserving article structure and metadata. This design facilitates both manual translation and machine-assisted content creation, offering a more efficient alternative to writing new articles.

The translation workflow comprises several integrated components. Editors begin by selecting an article from one Wikipedia language edition to translate into another language. The system presents a dual-panel interface (Figure 1) with the source article on the left and translation workspace on the right, enabling direct comparison throughout the process. The system automates technical aspects like preserving internal links, maintaining category tags, and transferring media assets, allowing editors to focus on content accuracy and quality.

The system accommodates different translation approaches by offering both fully manual translation and machine-assisted drafts for review and refinement. This flexibility helps maintain Wikipedia's quality standards while supporting diverse editor preferences. The system's impact has been substantial - by late 2024, it had facilitated over 2 million articles across Wikipedia's language editions[1].

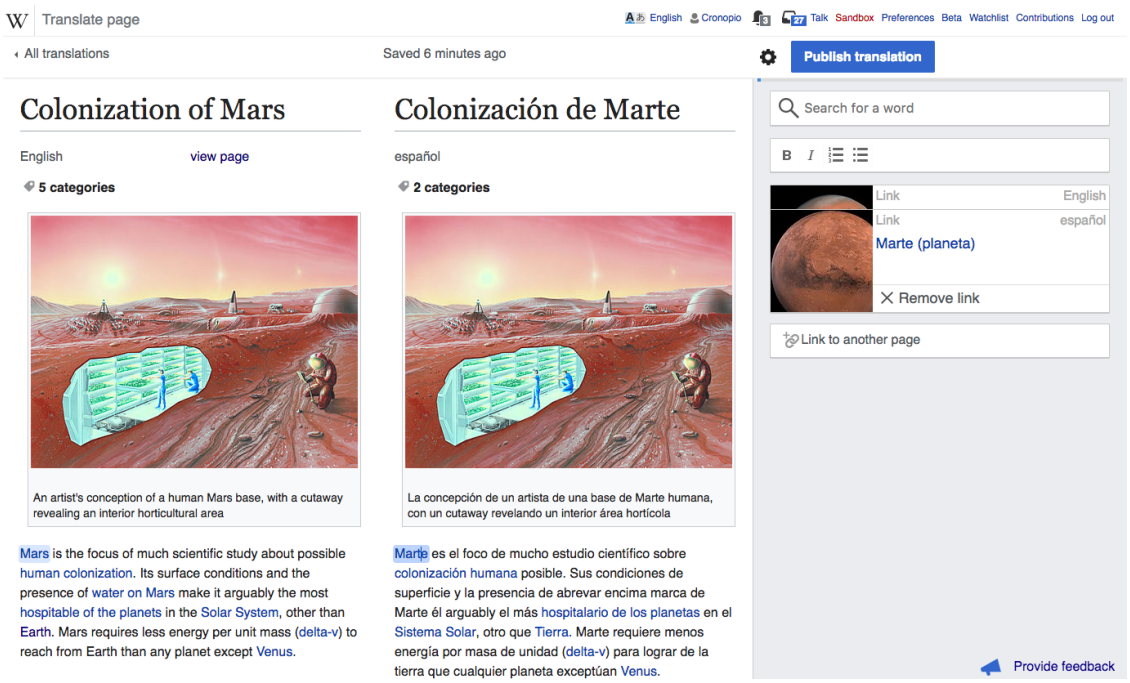
The Content Translation system initially integrated basic machine translation engines including Apertium and Yandex, which offered limited accuracy and language coverage. The integration of Google Translate in January 2019 marked a significant technological advancement (Figure 2), particularly benefiting languages previously served by less sophisticated translation tools[2].

Google Translate's neural machine translation architecture marked a substantial advancement over previous technologies. The system utilizes transformer-based models that process entire sentences holistically, enabling more sophisticated understanding of language context and meaning. Published evaluations demonstrate that Google Neural Machine Translation (GNMT) reduces translation errors by 55-85% across major language pairs compared to earlier approaches. This integration advances Wikipedia's mission of enhancing cross-language

[1]See this page: https://en.wikipedia.org/wiki/Special:ContentTranslationStats

[2]See the announcement from Wikimedia Foundation: https://wikimediafoundation.org/news/2019/01/09/you-can-now-use-google-translate-to-translate-articles-on-wikipedia/

Figure 1: Wikipedia's Content Translation Tool Interface



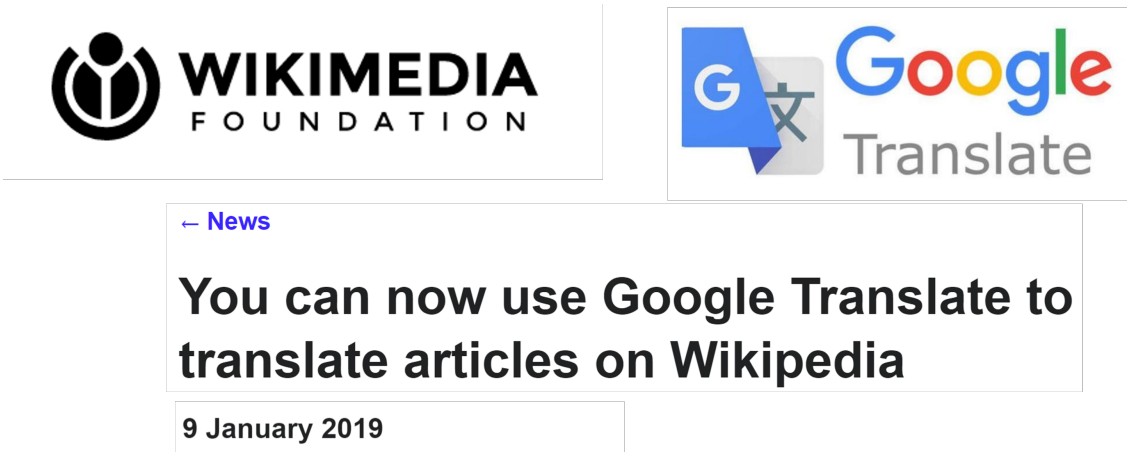
Note: Screenshot showing the side-by-side interface of Wikipedia's Content Translation tool. The left panel displays the source article in English, while the right panel provides the workspace for creating the Spanish translation. The interface preserves article structure and formatting while facilitating translation work.

information accessibility while upholding its commitment to open knowledge principles. All machine-translated content remains under Wikipedia's free license framework, with an implementation that safeguards user privacy by preventing transmission of personal data to external servers.

Importantly, Wikipedia implements Google Translate within a human-in-the-loop framework. Editors utilize machine-generated translations as initial drafts, leveraging their expertise to review, revise, and refine the content. This approach acknowledges that while machine translation can efficiently produce draft translations, human oversight remains critical for ensuring accuracy and preserving cultural nuances.

The Content Translation system provides comprehensive tracking of machine translation usage patterns, offering rich data for empirical analysis. The system logs each translation event, including metadata about language pairs, translation engines used, and creation

Figure 2: Google Translate Integration Announcement



Note: Screenshot of Wikimedia Foundation's January 9, 2019 announcement introducing Google Translate integration into the Content Translation tool. This update provided editors with neural machine translation capabilities while maintaining Wikipedia's open knowledge principles.

timestamps. This setting presents unique advantages for studying how artificial intelligence affects peer production platforms: a clear pre/post technology integration timeline, variation in language coverage, and detailed, comprehensive data enable thorough analysis of both aggregate and distributional effects of machine translation technology. Together, these features create an ideal natural experiment for examining the impact of AI-powered translation tools on collaborative content creation.

4 Data and Method

4.1 Data

Our empirical analysis draws on several complementary data sources. Our primary dataset consists of comprehensive translation records from Wikipedia's Content Translation System API, which documents all article translations between language pairs. Each record contains detailed metadata including source and target article titles, URLs, publication dates, and the

specific machine translation engine used (if any) (see Section A.1 for details). The sample spans January 2017 to December 2020, providing approximately two years of observations both before and after Google Translate's integration.

We supplement this core translation data with two categories of complementary datasets that characterize both articles and languages:

At the article level, we collect: (1) pageview data to analyze content consumption patterns and measure reader engagement with both source and translated articles; (2) language-agnostic topic classifications for consistent categorization across languages; (3) structured metadata from Wikidata capturing attributes like biographical and geographical information; and (4) Wikimedia Foundation's structural quality metrics assessing article quality.

At the language level, we compile: (1) a rich dataset capturing community infrastructure and editor engagement across Wikipedia language editions, including indicators of platform capacity and usage - total number of articles (measuring existing knowledge base), active editor counts (reflecting community capacity), total pageviews in 2019 (indicating readership intensity), and Google Translate performance metrics[3]; and (2) the Knowledge Gap Index measuring baseline content availability across languages and categories (see Section A.2 for details). This information allows us to explore how translation benefits vary systematically across languages and communities.

To construct our analysis sample, we applied several filtering criteria. We included only target languages with at least 20 translated articles over the sample period to ensure sufficient translation activity. We excluded English Wikipedia as a target language due to its specific restrictions on machine translations. We also removed language editions exhibiting abnormal translation spikes[4]. The resulting sample contains 540,488 translated articles across 132 target languages, which we aggregate into a monthly panel from January 2017 to December

[3]Our translation quality metrics come from an independent evaluation of Google Translate accuracy across 107 languages. See: "Empirical Evaluation of Google Translate across 107 Languages" (https://www.teachyoubackwards.com/empirical-evaluation/)

[4]We defined abnormal spikes as instances where a single month's translation volume exceeded 20 times the monthly average, likely indicating coordinated translation campaigns rather than organic activity

2020. Summary statistics are presented in Table 1.

Table 1: Summary Statistics of Wikipedia Translation Activity (2017-2020)

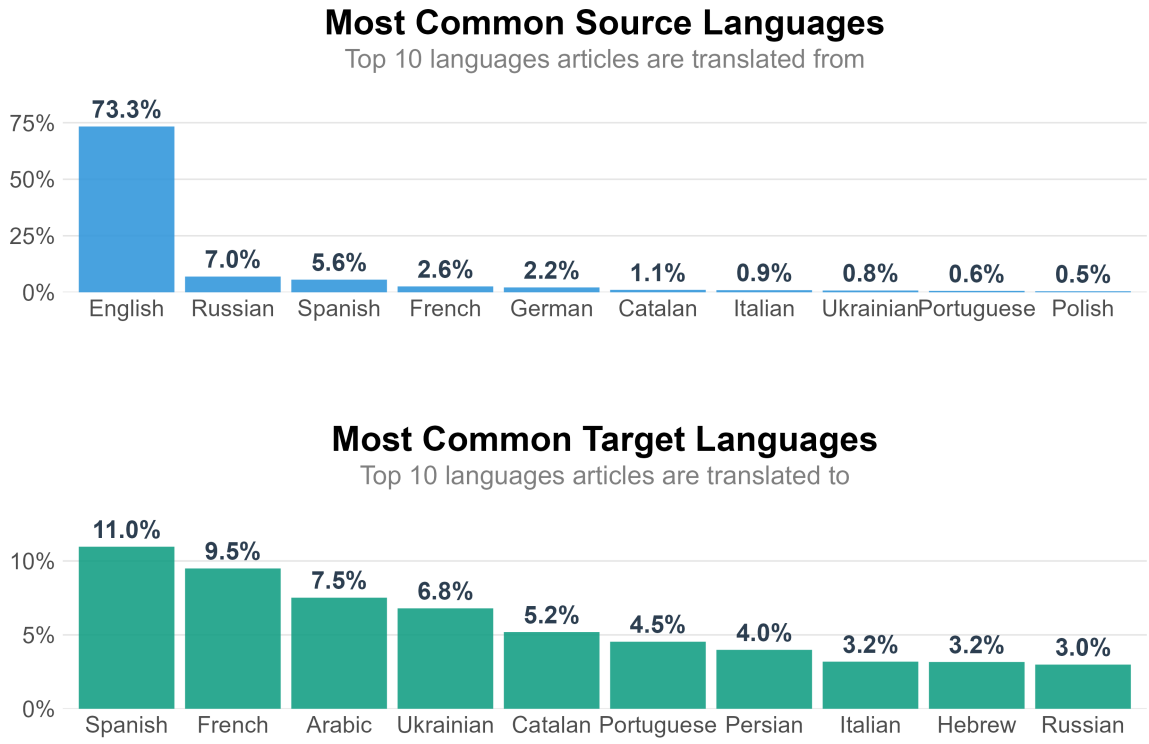
Content Translation Activity Overview	
Key Metrics	**Statistics**
Number of Translations	540,488
Sample Period	2017 – 2020
Number of Target Languages	132
Average Monthly Translations per Language	85

Note: This table presents key metrics from Wikipedia's Content Translation system, showing the scale of translation activity across languages during the study period.

The baseline distribution of translations across languages and topics provides essential context for evaluating Google Translate's impact. Analysis of source and target language patterns reveals significant asymmetries in Wikipedia's translation ecosystem (Figure 3). Source languages show notable concentration, with one language accounting for 73.3% of all translations, while Russian, Spanish, and French play secondary roles. Target languages demonstrate greater diversity - Spanish receives 11% of translations, followed by French, Arabic, and Ukrainian. These patterns motivate our later analysis of whether machine translation technology can foster more balanced cross-language information flows and expand access to underserved language communities.

The topical distribution of translations reveals additional content ecosystem patterns. As shown in Table 2, translations heavily concentrate in Culture (45.12%) and Geography (30.0%), which together comprise three-quarters of all translated articles. STEM (15.3%) and History and Society (9.5%) receive substantially less attention. At a more granular level (detailed in Appendix D), Geography.Regions and Culture.Biography emerge as translation hotspots. Given this concentration in biographical and geographical content, our later analysis examines how machine translation affects representation gaps in these domains,

Figure 3: Source and Target Language Distribution in Wikipedia Translations (2017-2020)



Note: This figure illustrates the distribution of languages in Wikipedia translations. English dominates as a source language (73.3% of translations), while target languages show greater diversity with Spanish (11%), French, Arabic, and Ukrainian being most frequent recipients.

particularly focusing on gender balance in biographies and regional coverage in geographical articles.

Table 2: Distribution of Wikipedia Translations by Topic Category (2017-2020)

Article Topic	**Count**	**Percentage**
Culture	231,830	45.1%
Geography	154,260	30.0%
STEM	78,726	15.3%
History and Society	48,680	9.5%
Total	**513,496**[5]	**100.0%**

Note: This table shows the topical breakdown of translated articles. Culture and Geography dominate, together accounting for 75% of all translations, while STEM and History/Society topics receive less coverage.

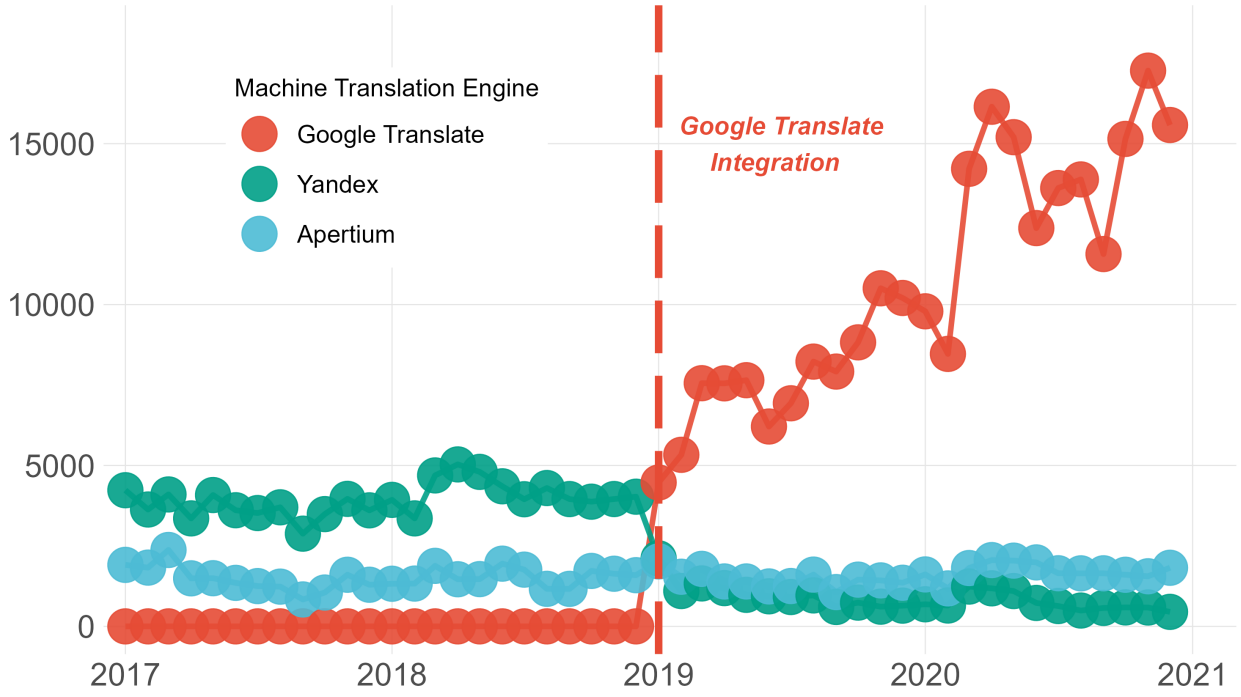
4.2 Model-free Evidence

Before presenting our formal empirical analysis, we examine descriptive patterns in Wikipedia's translation data that provide initial insights into Google Translate's impact. We focus on three key dimensions: the adoption of different machine translation engines, changes in overall translation volume, and the shift from manual to machine-assisted translation methods.

Figure 4 shows the monthly usage trends of machine translation engines from 2017 to 2020. Before Google Translate's integration in January 2019 (marked by the vertical red line), Wikipedia editors primarily used two engines: Yandex and Apertium, both exhibiting stable usage levels. After Google Translate's integration, two clear patterns emerge: rapid adoption of Google Translate and a simultaneous decline in Yandex usage. This suggests both a substitution effect across translation technologies and an overall increase in machine translation adoption.

The introduction of Google Translate coincided with broader changes in translation patterns. Figure 5 tracks the share of articles translated without machine assistance (manual translations). Prior to Google Translate's integration, roughly 20% of translations were completed manually. Following integration, this share dropped dramatically. This sharp decline suggests that editors found machine translation tools sufficiently effective to largely abandon manual translation methods.

Figure 4: Machine Translation Engine Usage Trends in Wikipedia (2017-2020)

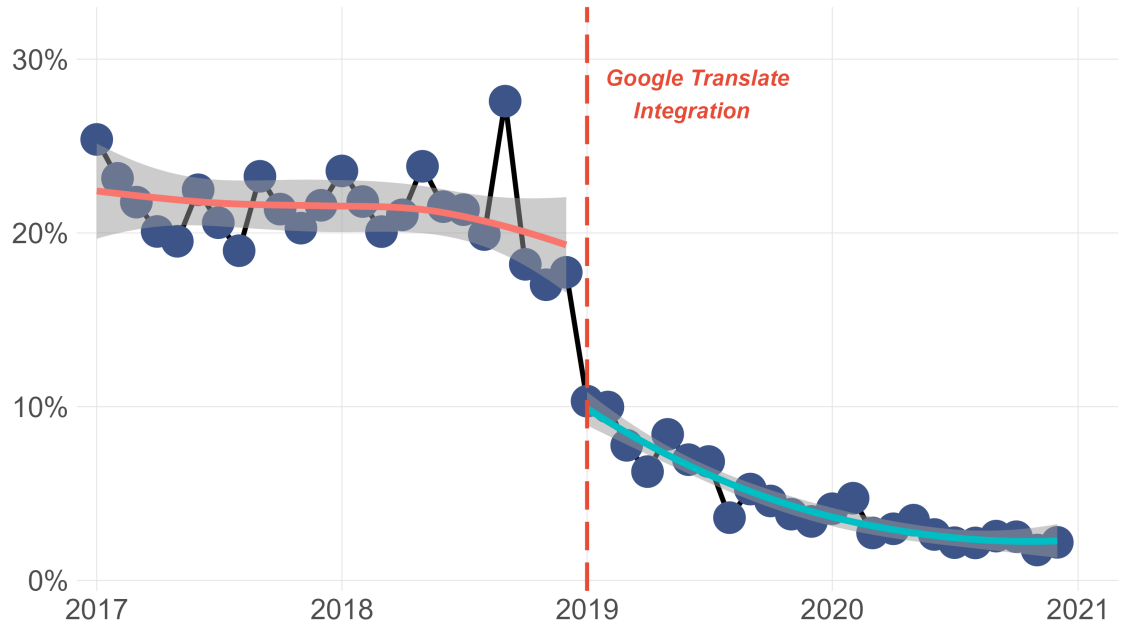


Note: This figure tracks monthly usage of different translation engines. The vertical red line marks Google Translate's integration (January 2019), showing its rapid adoption and displacement of other engines.

These descriptive patterns reveal three key insights. First, Google Translate achieved rapid and widespread adoption within Wikipedia's translation ecosystem, becoming the predominant translation tool. Second, rather than merely redistributing existing translation activity, the new technology appears to have expanded overall translation volume. Third, the dramatic decline in manual translations indicates a fundamental shift in editor behavior, with machine translation becoming central to content creation.

While these patterns suggest compelling evidence of Google Translate's impact, they do not establish causality or fully characterize the transformation of Wikipedia's content creation ecosystem. The following sections employ rigorous empirical strategies to identify causal effects and examine heterogeneous impacts across languages and content domains,

Figure 5: Trend in Manual Translation Following Google Translate Integration (2017-2020)



Note: This figure shows the percentage of articles translated without machine assistance over time. The vertical red line (January 2019) marks Google Translate's integration, after which manual translations declined sharply from 20% to close to zero.

building on these descriptive insights.

4.3 Empirical Strategy

Our empirical strategy leverages Wikipedia's integration of Google Translate as a natural experiment to examine how machine translation technology affects content production in peer production platforms. We implement a difference-in-differences framework that exploits three key sources of variation in Google Translate's rollout across Wikipedia language editions.

First, the timing of Google Translate integration resulted from a top-down partnership between the Wikimedia Foundation and Google, independent of individual language communities' characteristics or needs. This organizational decision provides plausibly exogenous variation in technology adoption. Second, Google Translate's language support capabilities

created natural treatment and control groups, as the system could initially translate content for only about one-third of Wikipedia's language editions. We restrict our analysis to languages ranking in the top 150 by Wikipedia size to ensure comparability between treated and control groups. Third, while most supported languages integrated Google Translate in January 2019, some languages experienced later integration through 2020, providing additional temporal variation in adoption (Table 3).[6]

Table 3: Google Translate Deployment Timeline across Wikipedia Language Editions

Implementation Status	**Deployment Period**	**Languages**
Completed Implementation		
Initial Major Rollout	January 2019	78
Staged Rollout	June 2019 – August 2020	12
Pending Implementation	—	42
Total Coverage		**132**

Note: This table details the phased implementation of Google Translate across Wikipedia languages. The initial major rollout in January 2019 covered 78 languages, followed by 12 more languages in staged releases through August 2020, with 42 languages pending implementation.

Base on the empirical strategy described above, our baseline model specification is as follows:

$$D_{it} = \beta \times Google\ Translate\ Integration_{it} + \gamma_i + \delta_t + \epsilon_{it}$$

where D_{it} represents outcomes (e.g., monthly translated articles) for Wikipedia language

[6]While the staggered rollout raises potential selection concerns, several factors support our identification strategy: (1) the initial January 2019 deployment (87% of treated languages) was driven by plausibly exogenous collaboration rather than language-specific factors; (2) consultations with the Wikimedia Language Team revealed that delayed adoption mainly reflected community governance processes; and (3) a conservative analysis using only January 2019 rollout languages yields similar results, suggesting findings are robust to timing variations.

edition i at month t. $Google\ Translate\ Integration_{it}$ indicates technology availability, while γ_i and δ_t represent language and time fixed effects respectively. The identifying assumption is that treated and control languages would have exhibited parallel trends in translation activity absent Google Translate integration.

To ensure robust inference, we employ multiple contemporary difference-in-differences (DiD) estimators, including traditional two-way fixed effects models, Sun and Abraham interactions, Callaway and Sant'Anna estimators, and a counterfactual imputation framework. Results remain highly consistent across these methodologies, with detailed information and model estimates presented in Appendix C. Our main analysis focuses on the counterfactual imputation method using matrix completion, which offers particular advantages for our setting.

The matrix completion method (Athey et al., 2021) constructs a panel data matrix where rows represent language editions, columns represent time periods, and entries capture outcome variables. The method leverages stable patterns across both dimensions to impute counterfactual outcomes for treatment effect estimation. This approach is particularly well-suited to our context as it accommodates heterogeneous treatment effects across languages and time periods, while utilizing both control units and pre-treatment patterns for counterfactual construction. These features enhance robustness beyond traditional difference-in-differences approaches.

To further validate our findings, we conduct a simplified analysis focusing only on languages that integrated Google Translate in January 2019. This approach, while sacrificing some variation in adoption timing, avoids econometric complications from staggered treatment and captures the majority of treated languages. Results from this simplified specification closely align with our primary findings. We employ this framework particularly for heterogeneity analyses, where its clearer identification supports examination of effect variation across language and topic types.

5 Aggregate Impact on Production and Consumption

We examine the causal impact of Google Translate integration on three key dimensions of Wikipedia's ecosystem: content production, quality maintenance, and reader engagement. Using the collaboration between Google and Wikipedia as a natural experiment, we evaluate how improved machine translation technology affects both content creation (supply side) and reader consumption (demand side) of Wikipedia articles.

For language-level analysis, such as content production, we employ the difference-in-differences framework with matrix completion described above. For article-level analysis of quality maintenance and reader engagement, we use regression discontinuity in time design to compare key metrics before and after Google Translate integration.

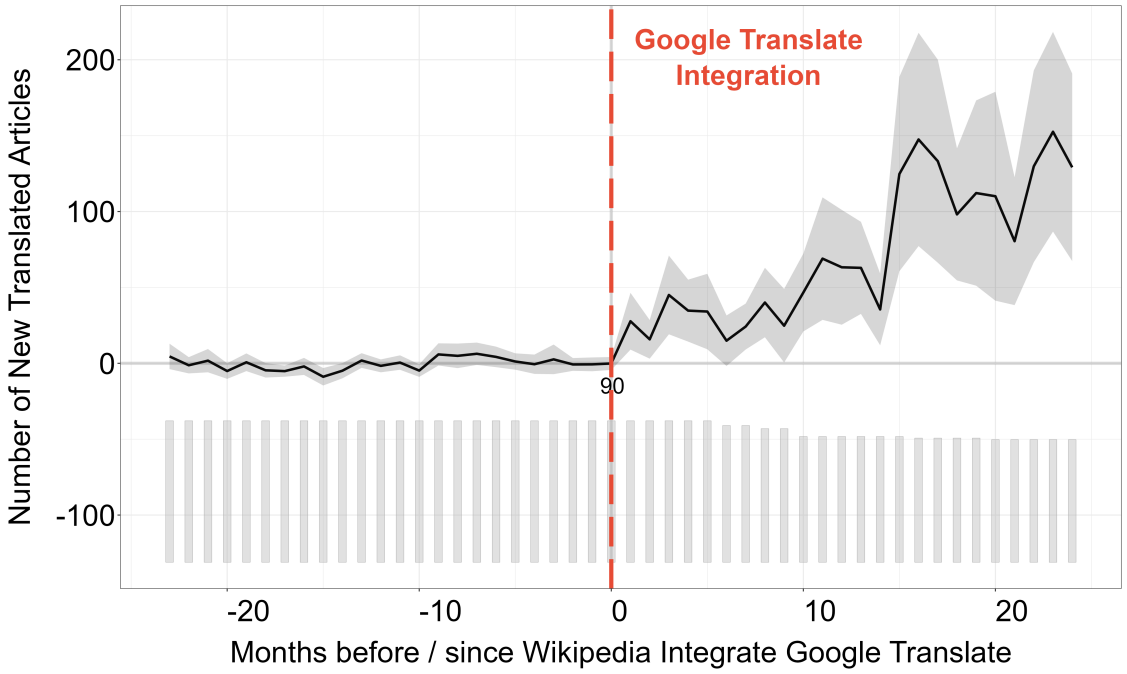
5.1 Content Production

We begin by analyzing how Google Translate integration affects Wikipedia's content translation volume. This analysis sets the stage for our subsequent examinations of quality maintenance and reader engagement.

The event study results (Figure 6) reveal that Google Translate integration triggers an immediate and sustained increase in translation activity. The estimated Average Treatment Effect on the Treated shows an increase of 71.2 translated articles per language edition per month ($p < 0.01$), representing a 149% increase from pre-integration levels. This substantial increase demonstrates how machine translation technology effectively reduces cross-language barriers and empowers editors to create more content.

To strengthen our causal inference, we conduct two key diagnostic tests (Figure 7). First, an equivalence test confirms parallel trends in translation activity before Google Translate integration ($p=0.819$), indicating that treated and control languages followed similar trajectories until 2019. Second, a placebo test that artificially shifts the integration date six months earlier yields no significant changes in translation volume, further supporting our

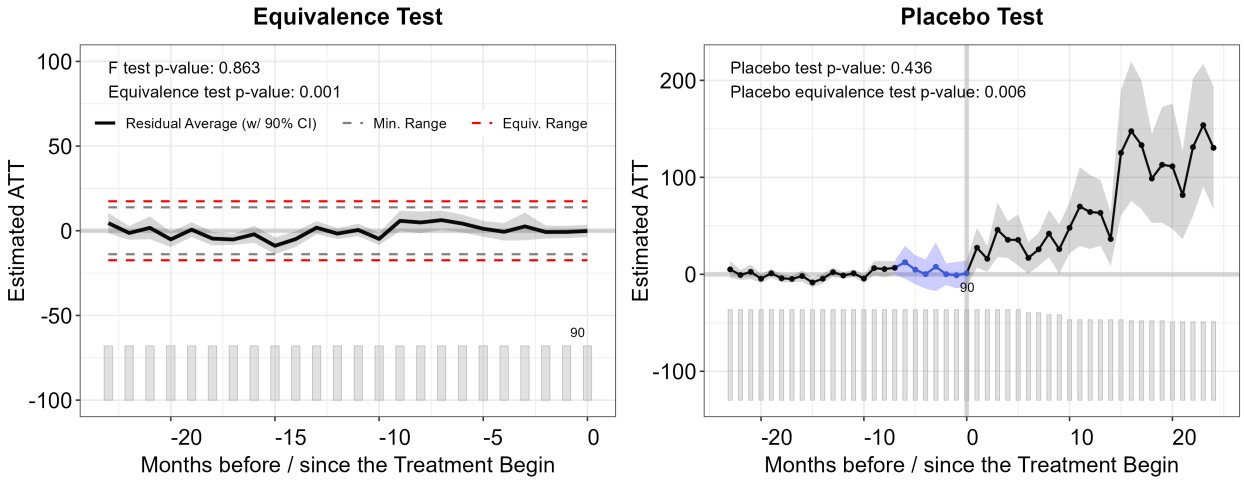
Figure 6: Effect of Google Translate Integration on Monthly Translation Volume



Note: This event study plot shows changes in translation volume relative to Google Translate integration (t=0). The black line represents point estimates with 95% confidence intervals (shaded area), demonstrating a sustained increase in translation activity post-integration.

causal interpretation.

Figure 7: Model Validation: Pre-trend and Placebo Tests



Note: Left panel shows the equivalence test confirming parallel pre-trends between treatment and control groups (F-test p=0.819). Right panel presents placebo test results using an artificial treatment period (blue region), showing no significant effects (p=0.455).

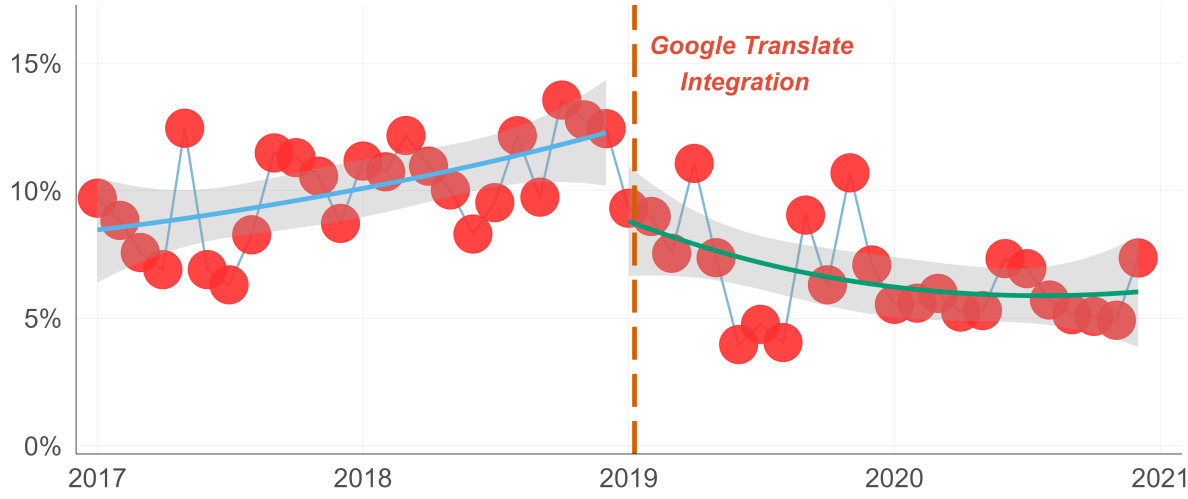
5.2 Quality Maintenance

Given the substantial growth in translation volume, we next assess whether this expansion came at the cost of content quality. We examine two complementary indicators: community-based deletion processes and structural quality metrics based on article organization and references. In the next subsection, we also consider reader engagement with translated content as an additional perspective on content utility. Taken together, these measures address the challenge of evaluating quality across Wikipedia's diverse language editions and provide several viewpoints, including community validation, objective scoring, and user consumption.

Our first quality indicator is Wikipedia's community-based deletion mechanism. Articles deemed to fall short of quality standards are removed through a collaborative process. The deletion ratio (the proportion of translated articles subsequently deleted) is especially informative, as it is the primary quality signal monitored by the Wikimedia Foundation in managing the Content Translation system. While smaller language editions display more variability in deletion ratios due to low monthly translation volumes, aggregating results

across all languages offers a stable monthly metric.[7] Focusing on languages that adopted Google Translate in January 2019, we observe that deletion ratios decline substantially after integration (Figure 8). This pattern suggests more consistent adherence to Wikipedia's quality standards following the introduction of Google Translate.

Figure 8: Article Deletion Rates Before and After Google Translate Integration



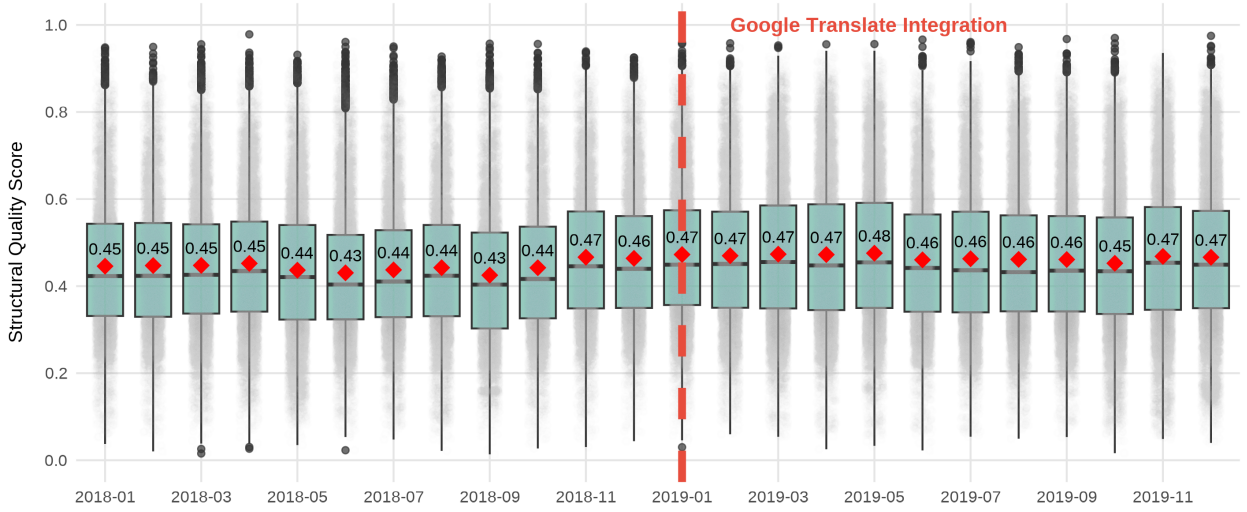
Note: This figure tracks the proportion of translated articles that were subsequently deleted. The vertical red line marks Google Translate integration (January 2019), with LOESS regression curves showing a decline in deletion rates post-integration, suggesting improved content quality.

To complement this community-based assessment, we employ the Wikimedia Foundation's official article quality prediction model as our second indicator. This language-agnostic machine learning system evaluates articles based on structural features including reference count, section organization, article length, and multimedia integration, producing fine-grained quality scores at the article level. As visualized in Figure 9, we observe stable trends in structural article quality during 2018 and 2019, with a modest but statistically significant improvement in articles created after Google Translate integration (0.02 units in

[7]For each translation month M, we calculate: Deletion Ratio(M) = (Number of articles translated in month M that were eventually deleted) / (Total number of articles translated in month M) across all languages.

the structural quality metric, $p<0.01$; see Figure A.20 in Appendix D for details). This enhancement in objective quality metrics indicates that machine translation assists editors in creating more thoroughly structured and well-documented content.

Figure 9: Article Quality Scores Distribution Before and After Google Translate Integration



Note: This figure presents monthly box plots of article quality scores, where t=0 marks Google Translate integration. The modest but significant increase in median scores post-integration indicates that machine translation helps maintain or enhance content standards.

Taken together, these signals present an encouraging picture: machine translation seems to maintain or even enhance content quality. The declining deletion ratios suggest that automated translation assists editors in meeting Wikipedia's standards more reliably. Likewise, improved quality scores point to more organized and better-referenced articles. In the next section, we turn to reader engagement as another lens for assessing the value of translated content, examining how these newly generated articles serve Wikipedia's goal of broad knowledge dissemination.

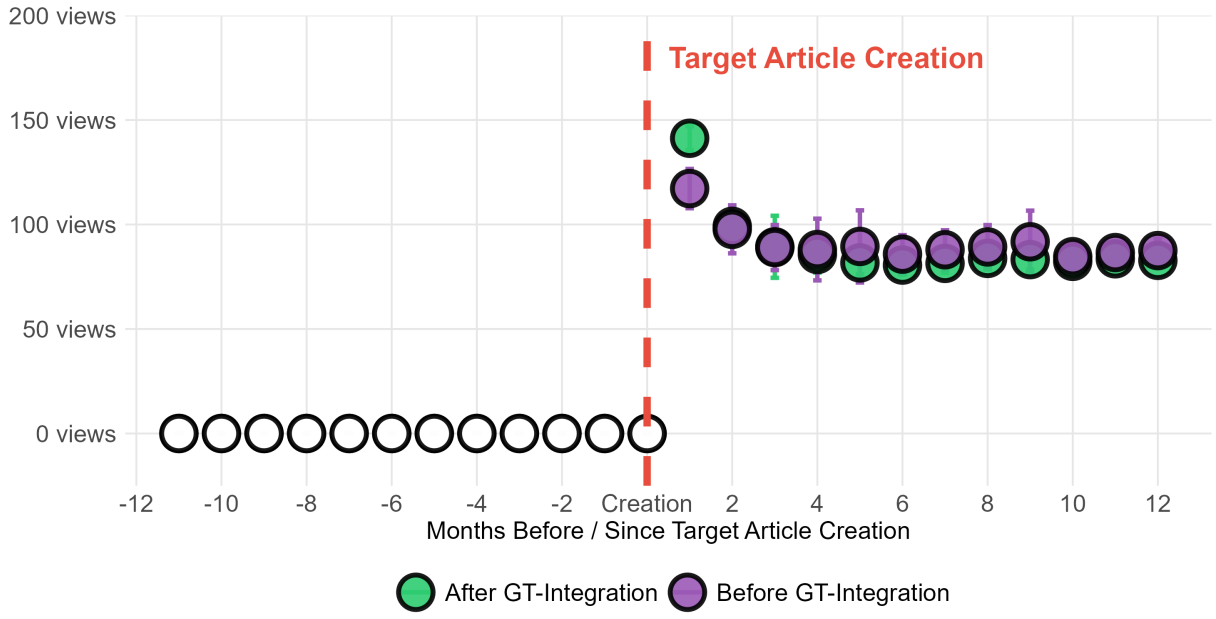
5.3 Reader Engagement

Having established the impact on content creation and article quality, we next examine whether translated articles effectively reach readers and fulfill the goal of expanding infor-

mation access. Our analysis of reader engagement focuses on three key dimensions: the direct effect of Google Translate integration on article readership, changes in combined source-target readership, and engagement patterns across Wikipedia editions of different sizes.

First, we analyze how Google Translate integration affects article pageviews. As shown in Figure 10, translated articles exhibit a characteristic pattern: an initial spike of 140-150 pageviews immediately after creation, followed by stabilization at approximately 80-90 monthly pageviews. Notably, while Google Translate substantially increased translation volume, individual articles maintain consistent readership levels before and after integration. This stability suggests that machine translation enables content expansion while preserving the engagement value of translated articles.

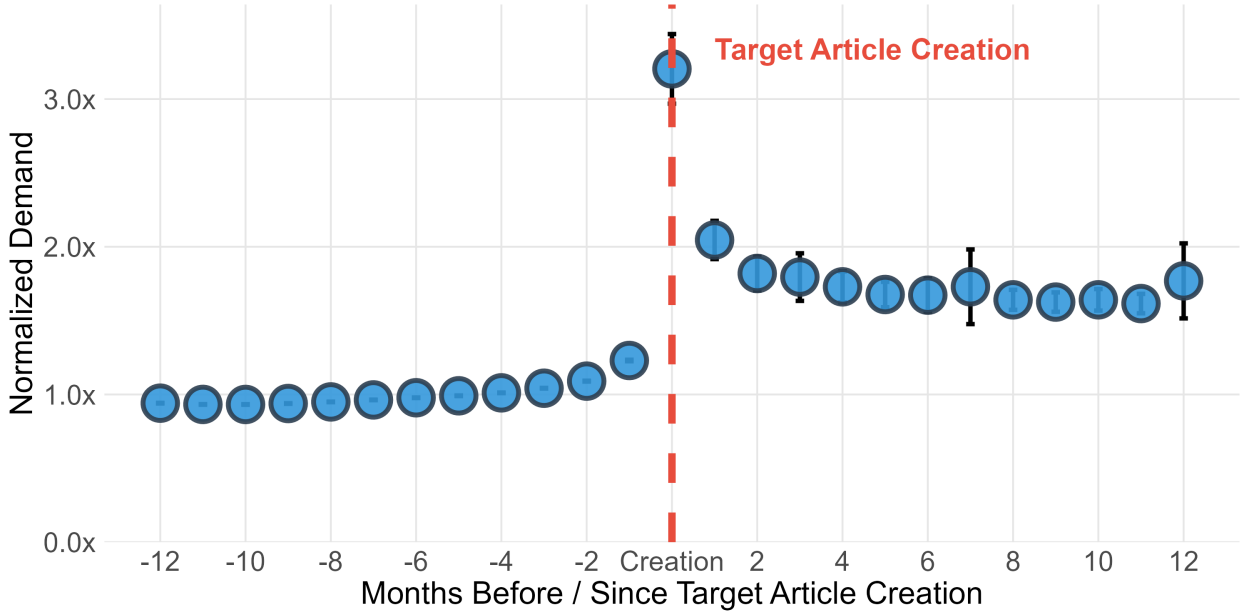
Figure 10: Readership Patterns of Translated Articles Over Time



Note: This figure tracks monthly pageviews of translated articles before and after Google Translate integration in January 2019. Articles show consistent engagement patterns: an initial spike of 140-150 pageviews post-creation, followed by stabilization at 80-90 monthly pageviews. This stability across both periods suggests machine translation maintains article quality and reader interest while enabling increased content production.

An important question is whether translations merely redistribute existing readership from source articles rather than expanding overall reach. To investigate this, we examine combined source-target readership using normalized metrics. In Figure 11, we normalize pageviews relative to a pre-translation baseline (average monthly pageviews of source articles in the 12 months before translation, set to 1.0). The analysis reveals that after translation, combined normalized pageviews increase to approximately 1.6 times the baseline. This elevated engagement persists for at least 12 months, demonstrating that translations genuinely expand Wikipedia's reach rather than simply shifting existing readership.

Figure 11: Combined Readership Impact of Article Translation

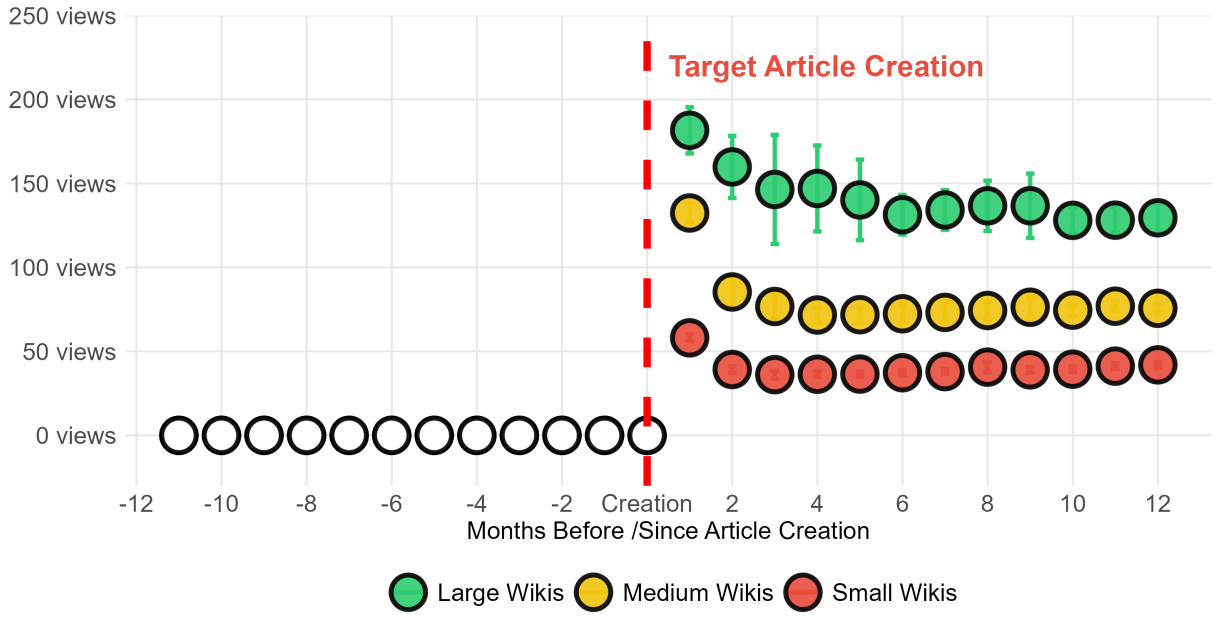


Note: This figure shows normalized pageviews for source articles and combined source-target pairs, with pre-translation source article pageviews as baseline (1.0). The increase to approximately 1.6 in combined normalized views demonstrates that translations expand total readership rather than merely redistributing existing readers. This elevated engagement persists for at least 12 months post-translation.

Finally, we observe significant variation in reader engagement across Wikipedia editions of different sizes. Figure 12 shows that large Wikis (those ranking in the top 10 by article count) consistently attract higher engagement, initially drawing 150-200 monthly pageviews

before stabilizing around 140 views per month. In contrast, smaller Wikipedia editions typically receive more modest engagement of 50-60 monthly pageviews. These disparities likely reflect differences in reader base size and community network effects across language editions, an insight that proves valuable for our subsequent analysis of distributional effects.

Figure 12: Readership Patterns by Wikipedia Edition Size



Note: This figure compares pageview trends across Wikipedia editions of different sizes: large (top 10), medium (ranks 11-50), and small (beyond rank 50). Large editions consistently achieve higher engagement (140-200 monthly views) compared to smaller editions (50-60 views), reflecting the influence of community size and network effects on content visibility.

Our analysis demonstrates that machine translation successfully reduces barriers to multilingual content production while maintaining article quality and fostering substantial reader engagement. However, the observed variations in engagement patterns raise important questions about the distribution of benefits across language editions. The next section examines how these benefits vary across language communities and content domains, investigating whether technological interventions have differential impacts based on existing structural factors.

6 Distributional Effects in Languages and Content

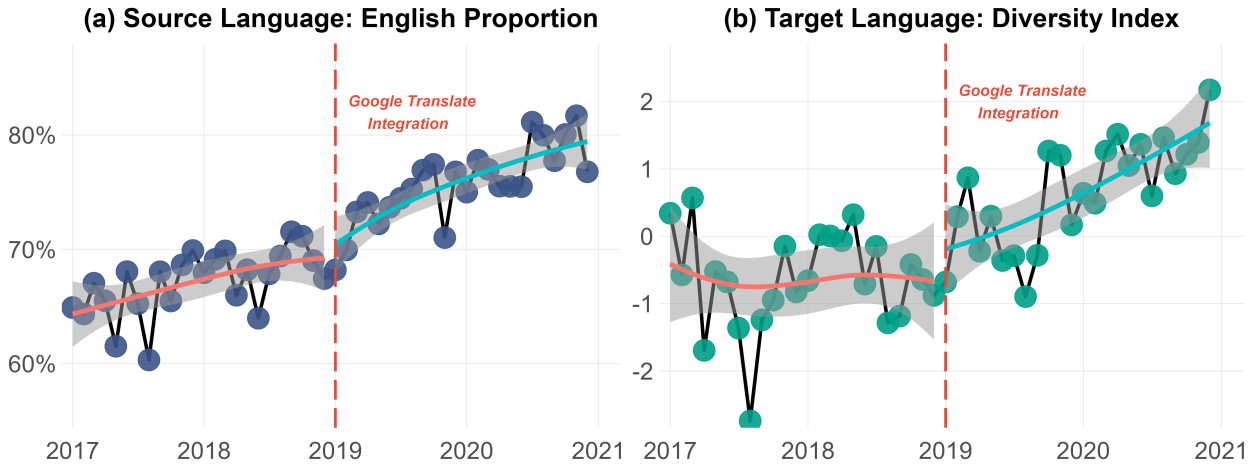
The previous section established that machine translation produces larger-scale article creation, preserving quality and attracting a wider readership. However, it is critical to assess whether this impact is felt equally across all language editions and content domains. In this section, we analyze how gains from Google Translate integration are shared across diverse Wikipedia language communities and topics. We begin with cross-language patterns, then turn to different article categories, including biographical and geographical content.

6.1 Cross-Language Information Flow

Our analysis reveals two distinct trends following Google Translate integration: while English Wikipedia's share of source translations measured as proportion of all translations increases substantially as shown in Panel (a) of Figure 13, there is simultaneously a growing diversity in target languages receiving translations, measured by Shannon entropy index illustrated in Panel (b). This pattern suggests that machine translation technology both reinforces English Wikipedia's role as a primary knowledge repository while democratizing access to that knowledge across a broader range of language communities. The concurrent centralization of source content and diversification of target languages highlights how technological interventions can reshape information flows, potentially both amplifying existing power structures while creating new opportunities for knowledge dissemination.

Building on these aggregate patterns, we next examine how translation benefits vary across language communities based on their characteristics. Our analysis focuses on four key dimensions: existing knowledge base (measured by article count), community capacity (number of active editors), readership (average pageviews), and Google Translate performance. For each dimension, we classify Wikipedia editions into three tiers: high (top 10), medium (ranks 11-50), and low (below rank 50). This tiered classification enables systematic analysis of how existing platform capacity and usage patterns shape the impact of translation

Figure 13: Evolution of Translation Patterns Across Source and Target Languages



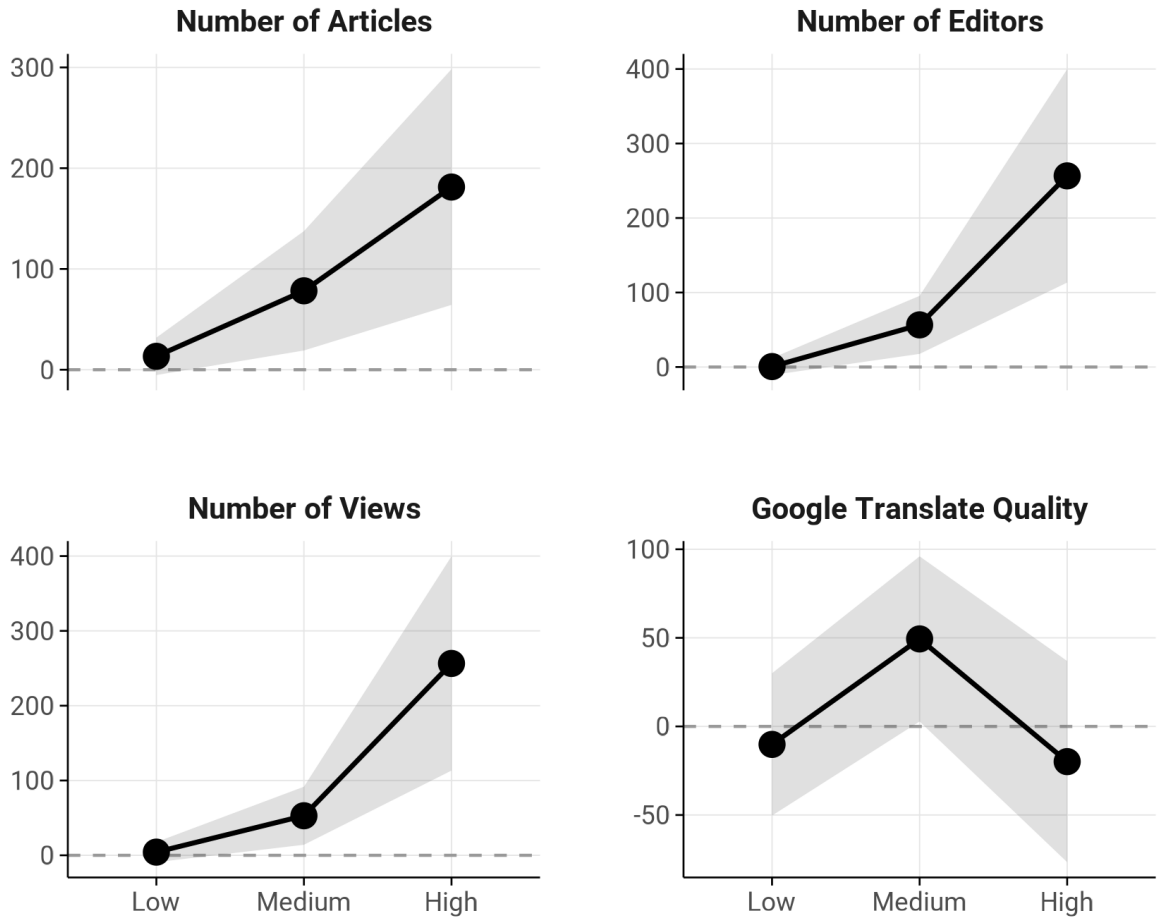
Note: Panel (a) shows English Wikipedia's increasing role as a source for translations, with a steeper trend after Google Translate integration (red line). Panel (b) tracks target language diversity using Shannon entropy index, where higher values indicate more evenly distributed translations across languages. The rising entropy demonstrates growing diversification of target languages despite source centralization.

technology. Figure 14 visualizes model estimates from this analysis, with detailed results provided in Appendix D.

Our analysis reveals consistent patterns across the first three dimensions. Wikipedia language editions with larger existing content bases, more active editors, and higher overall readership demonstrate the strongest gains from translation technology, followed by medium-sized editions. Small Wikipedias show modest increases that often fail to reach statistical significance. These disparities suggest that machine translation technology may amplify the existing advantages of well-established Wikipedia communities. Several factors likely drive this pattern: larger communities have more editors available to effectively leverage the translation tool, higher readership creates stronger incentives for translation, and active communities appear better equipped to integrate new tools into their workflows.

When examining Google Translate performance across languages, we find that medium-performance languages show greater increases in translation activity compared to high- and low- performance languages. However, these estimates carry higher uncertainty since per-

Figure 14: Impact of Google Translate Integration by Wikipedia Community Characteristics



Note: This figure presents difference-in-differences estimates of translation volume changes across Wikipedia editions grouped by four characteristics: knowledge base size, editor community size, readership levels, and translation quality. Each characteristic is divided into high (top 10), medium (ranks 11-50), and low (below rank 50) tiers. Shaded areas represent 95% confidence intervals, revealing consistently stronger benefits for larger, more established communities.

formance metrics are only available for a subset of languages in our sample. This pattern may reflect an optimal balance where translation quality is sufficient to be useful while still addressing significant content gaps that might otherwise remain unfilled.

These findings have important implications for platform design and policy interventions. They suggest that merely providing equal access to translation technology may not be suf-

ficient to achieve equitable outcomes across language communities. Instead, platforms may need to implement targeted support mechanisms for smaller communities, such as additional editorial assistance, specialized quality control processes, or incentive systems that prioritize translations into under-resourced languages. Lastly, the relationship between technological capability and existing community characteristics provides key context for our subsequent analysis of specific knowledge domains, where we examine how these disparities manifest in particular types of content.

6.2 Content Distribution and Representation Gap

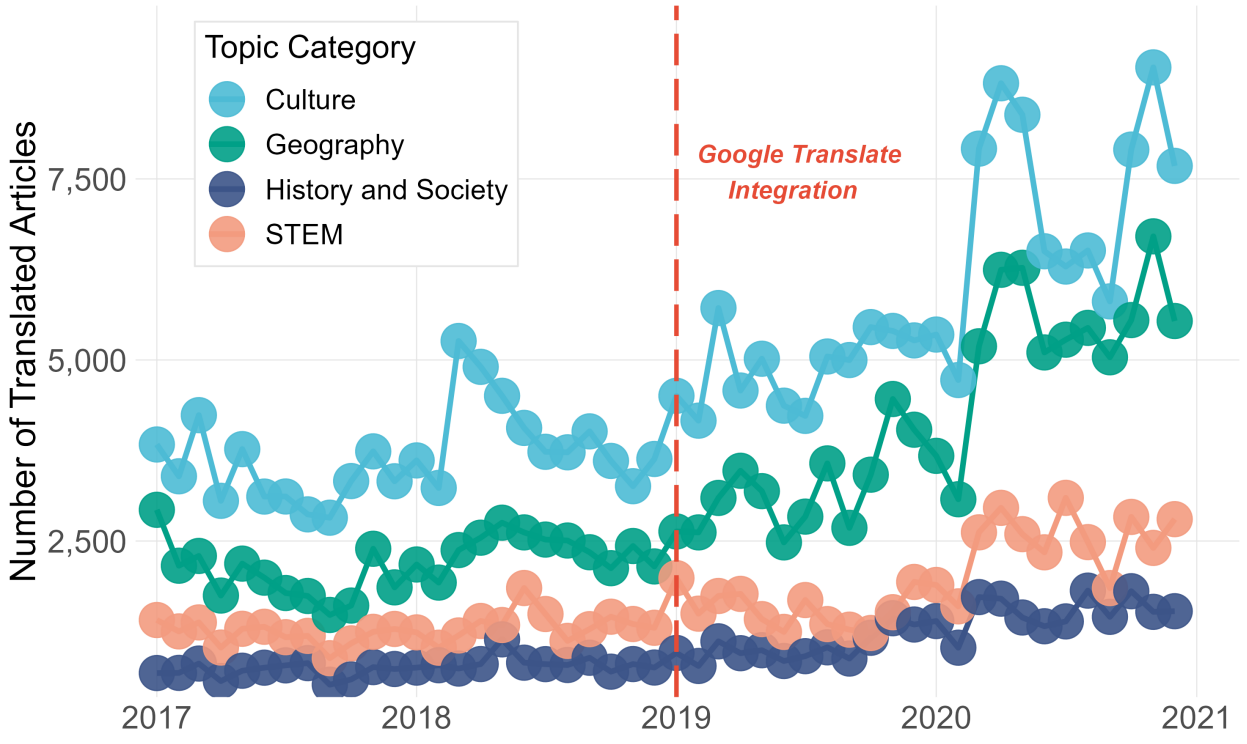
Our previous analyses demonstrated how machine translation shapes content creation and knowledge flows across Wikipedia languages, with benefits concentrated among resource-rich languages. We now examine how this technological intervention affects Wikipedia's systemic content gaps - a fundamental challenge that the Wikimedia Foundation has identified as a key priority. Bridging content gaps across languages and topics is central to Wikipedia's mission of providing comprehensive knowledge coverage. Our investigation examines how translation technology impacts these gaps by analyzing the distribution and representation patterns of articles across knowledge domains.

To motivate our subsequent analyses on systematic content gaps, we first explore how machine translation affects different knowledge domains, building on our earlier finding that impact varies by community resources. Understanding domain-specific effects is crucial given the substantial variation in how different types of content are created, verified, and used across languages. Topic classification in a multilingual setting presents methodological challenges due to linguistic and cultural variations. We address this by leveraging Wikidata as a language-agnostic classification system, following Johnson et al. (2021). This approach maps articles to a standardized 64-category taxonomy based on WikiProject, enabling consistent topic identification across languages. For methodological details, see Section B.1.

Figure 15 reveals heterogeneous growth across content domains following Google Trans-

late integration, with Cultural and Geographical articles showing particularly pronounced increases compared to more modest growth in History and Society as well as STEM topics. Difference-in-differences estimates quantify these patterns: cultural and geographical articles experienced the strongest increases (25.2 and 21.4 articles per month respectively, $p<0.01$), while STEM (7.9 articles, $p<0.01$) and History and Society content (5.6 articles, $p<0.05$) saw more modest effects. This variation suggests that machine translation's effectiveness differs across knowledge domains. Detailed model results are provided in Table A.7 in Appendix D.

Figure 15: Translation Volume by Content Category Before and After Google Translate Integration



Note: This figure tracks monthly translation volume across major content categories from 2017-2020, with Google Translate integration marked by the vertical red line (January 2019). Cultural and geographical content show the strongest growth post-integration, while History/Society and STEM topics exhibit more moderate increases, revealing varying effectiveness of machine translation across knowledge domains.

These topical patterns motivate a focused investigation of two key content domains: biographical articles (Culture) and geographical articles (Geography). These categories collec-

tively represent approximately 50% of all translated content and serve as critical channels for addressing representation gaps. Biographical articles enable diverse human representation across cultures and languages, while geographical content facilitates the sharing of localized knowledge and regional perspectives. Their prominence and fundamental role in knowledge equity make them particularly important for understanding how machine translation impacts Wikipedia's content distribution.

6.2.1 Gender Content Gap

We next examine how machine translation technology affects gender representation in biographical articles. Prior research has extensively documented substantial gender imbalances in Wikipedia's biographical coverage, making this a critical dimension for understanding whether reduced translation barriers can help address historical disparities in representation.

Our analysis leverages linkes between translated articles and Wikidata entries in order to systematically identify gender across languages. This methodology provides a consistent framework for analyzing representation patterns throughout Wikipedia's multilingual ecosystem, avoiding potential inconsistencies from language-specific gender identification approaches. For detailed documentation of our attribute extraction process, see Appendix Section B.2.

Using difference-in-differences estimation, we find that Google Translate integration led to significant increases in biographical translations across all gender categories: female biographies increased by 3.8 articles monthly ($p<0.01$), male biographies by 7.2 articles ($p<0.01$), and non-binary biographies by 0.06 articles ($p<0.01$). While all categories show growth, the larger absolute increase in male biographies suggests that technological interventions alone may not be sufficient to fully address representation gaps. Detailed model results are provided in Table A.8 in Appendix D.

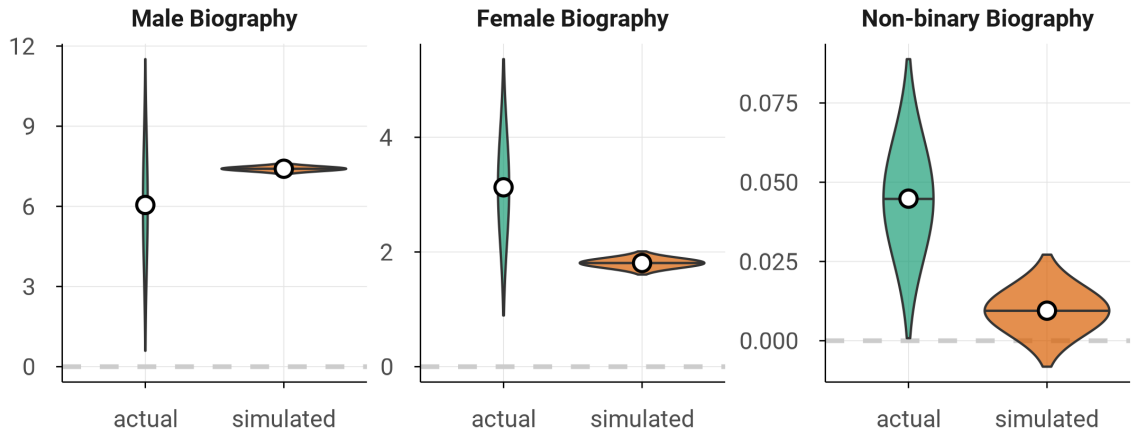
To contextualize the magnitude and significance of these changes, we analyze them rel-

ative to the baseline distribution of biographical content. As of January 2019, English Wikipedia's biographical articles exhibited a stark gender imbalance: 82.2% male (1.29 million articles), 17.7% female (0.28 million articles), and 0.1% non-binary (1,390 articles). However, translation patterns following Google Translate integration (2019-2021) show substantial progress toward better balance, with translations distributed as 62.4% male (30,202 articles), 37.2% female (17,976 articles), and 0.4% non-binary (184 articles).

To rigorously evaluate whether these translation patterns represent meaningful progress in representation, we conduct a simulation analysis that establishes an expected baseline from the existing content distribution. Our simulation models the translation patterns that would emerge if editors randomly selected articles from the English Wikipedia biography pool, creating a null model for comparison. This approach allows us to determine whether the observed shifts toward gender balance reflect intentional editorial choices enabled by reduced translation costs, rather than mirroring the existing content distribution.

Figure 16 reveals that actual translation patterns diverge significantly from simulated benchmarks. Translations of female and non-binary biographies exceed simulated expectations, while male biography translations fall below the simulated benchmark. This deviation from the null model provides strong evidence that translation patterns are actively contributing to improved gender representation, beyond what would be expected from simple mechanical reproduction of existing content distributions. These findings suggest that while technological tools like machine translation cannot automatically eliminate representation gaps, they can effectively support community efforts to achieve more balanced coverage when combined with intentional editorial choices. This pattern aligns with our broader findings about the role of community capacity in shaping translation outcomes, highlighting how technological tools interact with human agency to drive changes in content distribution.

Figure 16: Gender Representation in Translated Biographies: Observed vs. Expected Patterns



Note: This figure compares actual translation patterns with simulated expectations based on existing content distribution. Point estimates with 95% confidence intervals show observed translation coefficients, while violin plots represent distributions from 500 simulation replications. Female and non-binary biography translations significantly exceed simulated benchmarks, while male biography translations fall below expectations, indicating intentional efforts toward gender balance.

6.2.2 Geographic Content Distribution

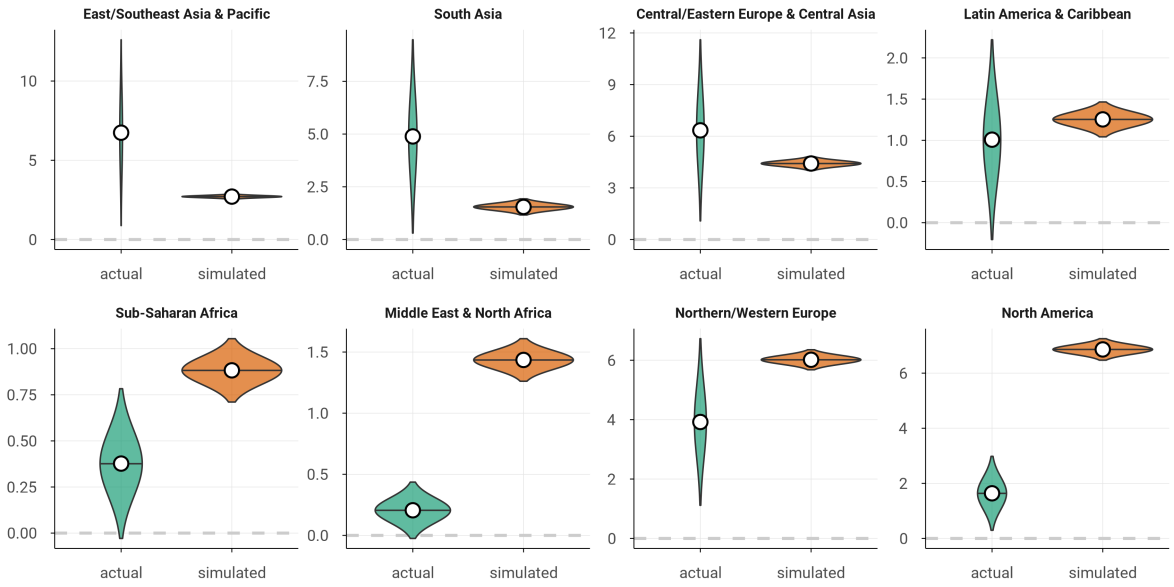
We next examine how machine translation technology affects the distribution of geographical knowledge across Wikipedia. Geographic articles constitute approximately 28% of all translated content and serve as crucial vectors for disseminating local knowledge globally, making them particularly important for understanding how translation tools impact regional representation.

Using the Wikimedia Foundation's eight-region classification system (illustrated in Figure A.21 in Appendix D), we analyze geographic content flows through difference-in-differences estimation. The impact of Google Translate integration varied substantially across regions: Central Eastern Europe Central Asia experienced the largest increase (8.6 articles monthly, $p<0.01$), followed by East/Southeast Asia Pacific (7.3 articles, $p<0.01$) and Northern/Western Europe (5.4 articles, $p<0.01$). However, Sub-Saharan Africa and the Mid-

dle East North Africa saw only minimal gains (0.4 and 0.3 articles respectively, $p<0.05$), highlighting persistent regional disparities in content creation. Detailed model results are available in Table A.10 in Appendix D.

Similar to our gender analysis, we conduct a simulation study establishing an expected baseline by modeling random article selection from existing English Wikipedia geographic content. Figure 17 reveals significant deviations from these simulated benchmarks - content about Asia and Eastern Europe exceeds expectations, while North American and Western European content falls below benchmarks, suggesting machine translation particularly benefits regions with high linguistic diversity.

Figure 17: Geographic Distribution in Translated Geography Articlles: Observed vs. Expected Patterns



Note: This figure compares actual regional content translation patterns against simulated benchmarks derived from existing English Wikipedia geographic content. Point estimates show translation coefficients with 95% confidence intervals, while violin plots represent distributions from 500 simulation replications. Regions like Asia and Eastern Europe exceed expected translation rates, while North American and Western European content falls below simulated benchmarks.

These findings demonstrate machine translation's varied impact on geographic represen-

tation, with the technology proving most effective in linguistically diverse regions. However, minimal gains in regions like Sub-Saharan Africa and the Middle East/North Africa suggest that language barriers are just one of several constraints - others include source content availability, local editorial capacity, and regional infrastructure. Combined with our gender representation analysis, this indicates that while machine translation reduces language barriers, achieving balanced representation requires addressing underlying structural constraints through targeted support for underserved regions.

7 Conclusion

This study provides novel empirical evidence on how artificial intelligence technologies, specifically machine translation, shape content creation and knowledge distribution in peer production platforms. Through a detailed analysis of Wikipedia's Google Translate integration, we uncover interesting dynamics in bridging digital language divides. The integration yielded substantial efficiency gains, driving a 149% increase in translation volume while maintaining content quality and reader engagement. However, these gains were unevenly distributed across language communities, with well-resourced languages capturing disproportionate benefits.

Our findings demonstrate that machine translation both enables and constrains cross-linguistic information flows in systematic ways. While English serves as a central hub for source content, the benefits are asymmetrically distributed based on key structural factors: community size, existing editorial capacity, and historical content development. Languages with large editor communities and established translation practices saw substantial increases in translation volumes, while smaller communities experienced modest gains. The impact on representation gaps varies meaningfully across content domains. For biographical content, translation patterns actively contribute to improved gender balance, with female and

non-binary content translations significantly exceeding simulated benchmarks. Geographic content shows particularly strong growth in linguistically diverse regions like Asia and Eastern Europe, while regions like Sub-Saharan Africa see minimal gains despite high need. These patterns demonstrate how machine translation's effectiveness depends critically on preexisting platform structures and community resources.

Our findings provide key insights into how AI technologies shape digital knowledge production and dissemination. The results demonstrate that, while machine translation can substantially reduce language barriers, its effectiveness is bounded by existing structural inequalities between language communities. Additionally, the impact of AI tools cannot be understood in isolation but must be analyzed within the broader socio-technical context of platform structures, community resources, and institutional capacities. These insights enhance our understanding of technological adoption and digital inequality by highlighting how structural factors mediate the distribution of benefits from AI innovations.

To translate these insights into practical impact, we recommend interventions that specifically target structural barriers in digital knowledge platforms. First, platforms should leverage AI capabilities to systematically identify content gaps in underserved languages and direct translation efforts to areas of greatest need. Second, incentive structures should be redesigned to explicitly reward contributions to low-resource languages, ensuring technological benefits reach underserved communities. Third, targeted capacity-building programs should strengthen smaller language communities' ability to effectively utilize translation tools through training, mentorship, and resource provision. These recommendations aim to create synergies between technological tools and structural reforms, maximizing the democratizing potential of machine translation.

Our study has several limitations that suggest important directions for future research. First, we do not examine how editor behaviors and community dynamics evolve after translation barriers are reduced - understanding these behavioral adaptations could reveal important insights about technology adoption and collaboration patterns. Second, our analysis focuses

on Wikipedia, and comparative studies across other knowledge platforms would help determine whether similar patterns in the distribution of technological benefits emerge in different contexts. As AI tools like machine translation become increasingly central to knowledge creation and sharing, understanding how they interact with platform structures and community dynamics will be crucial for developing policies and frameworks that maximize their benefits for human knowledge, creativity, and collective intelligence.

References

Aaltonen, A. and Seiler, S. (2016). Cumulative growth in user-generated content production: evidence from wikipedia. *Management Science*, 62(7):2054–2069.

Agrawal, A., Gans, J., and Goldfarb, A. (2022). *Prediction Machines, Updated and Expanded: The Simple Economics of Artificial Intelligence*. Harvard Business Press.

Athey, S., Bayati, M., Doudchenko, N., Imbens, G., and Khosravi, K. (2021). Matrix completion methods for causal panel data models. *Journal of the American Statistical Association*, 116(536):1716–1730.

Bahdanau, D., Cho, K., and Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.

Brynjolfsson, E., Hui, X., and Liu, M. (2019). Does machine translation affect international trade? evidence from a large digital platform. *Management Science*, 65(12):5449–5460.

Constantinides, P., Henfridsson, O., and Parker, G. G. (2018). Platforms and infrastructures in the digital age.

Egger, P. H. and Lassmann, A. (2012). The language effect in international trade: A meta-analysis. *Economics Letters*, 116(2):221–224.

Egger, P. H. and Lassmann, A. (2015). The causal impact of common native language on international trade: Evidence from a spatial regression discontinuity design. *The Economic Journal*, 125(584):699–745.

Forman, C., Ghose, A., and Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information systems research*, 19(3):291–313.

Ginsburgh, V. and Weber, S. (2020). The economics of language. *Journal of Economic Literature*, 58(2):348–404.

Goldfarb, A. and Tucker, C. (2019). Digital economics. *Journal of Eonomic Literature*, 57(1):3–43.

Graham, M., Hogan, B., Straumann, R. K., and Medhat, A. (2014). Uneven geographies of user-generated information: Patterns of increasing informational poverty. *Annals of the Association of American Geographers*, 104(4):746–764.

Halavais, A. and Lackaff, D. (2008). An analysis of topical coverage of wikipedia. *Journal of computer-mediated communication*, 13(2):429–440.

Hinnosaar, M., Hinnosaar, T., Kummer, M., and Slivko, O. (2023). Wikipedia matters. *Journal of Economics & Management Strategy*, 32(3):657–669.

Johnson, I., Gerlach, M., and Sáez-Trumper, D. (2021). Language-agnostic topic classification for wikipedia. In *Companion Proceedings of the Web Conference 2021*, pages 594–601.

Johnson, I. and Lescak, E. (2022). Considerations for multilingual wikipedia research. *arXiv preprint arXiv:2204.02483*.

Langrock, I. and González-Bailón, S. (2022). The gender divide in wikipedia: Quantifying and assessing the impact of two feminist interventions. *Journal of Communication*, 72(3):297–321.

Miquel-Ribé, M. and Laniado, D. (2020). The wikipedia diversity observatory: A project to identify and bridge content gaps in wikipedia. In *Proceedings of the 16th International Symposium on Open Collaboration*, pages 1–4.

Nagaraj, A. (2022). The private impact of public data: Landsat satellite maps increased gold discoveries and encouraged entry. *Management Science*, 68(1):564–582.

Parker, G. G. (2016). *Platform revolution: How networked markets are transforming the economy and how to make them work for you*. WW Norton & Company.

Pereira, M., Greenstein, S., Sadun, R., Tambe, P., Darre, L. R., Glazer, T., Kim, A., Dodhia, R., and Ferres, J. L. (2024). The new digital divide. Technical report, National Bureau of Economic Research.

Thompson, N. and Hanley, D. (2018). Science is shaped by wikipedia: evidence from a randomized control trial.

Thompson, N., Luo, X., McKenzie, B., Richardson, E., and Flanagan, B. (2024). User-generated content shapes judicial reasoning: Evidence from a randomized control trial on wikipedia. *Information Systems Research*.

Turovsky, B. (2016). Found in translation: More accurate, fluent sentences in google translate. Google Blog.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.

Wagner, C., Garcia, D., Jadidi, M., and Strohmaier, M. (2015). It's a man's wikipedia? assessing gender inequality in an online encyclopedia. In *Proceedings of the international AAAI conference on web and social media*, volume 9, pages 454–463.

Wagner, C., Graells-Garrido, E., Garcia, D., and Menczer, F. (2016). Women through the glass ceiling: gender asymmetries in wikipedia. *EPJ Data Science*, 5:1–24.

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., et al. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144*.

Xu, S. X. and Zhang, X. (2013). Impact of wikipedia on market information environment: Evidence on management disclosure and investor reaction. *MIS Quarterly*, pages 1043–1068.

Yilmaz, E. D., Naumovska, I., and Aggarwal, V. A. (2023). Ai-driven labor substitution: Evidence from google translate and chatgpt.

Zhu, K., Walker, D., and Muchnik, L. (2020). Content growth and attention contagion in information networks: Addressing information poverty on wikipedia. *Information Systems Research*, 31(2):491–509.

Web Appendix

A Data Source

A.1 Published Translations on Wikipedia

The primary data for our analysis comprises published translations from Wikipedia's Content Translation system. Wikimedia Foundation provides data access through Application Programming Interface (API) that lists published translations, offering comprehensive details essential for our study:

- **API Access**: The API provides lists of all published translations across languages, including details like translation IDs, translation date, source and target titles, source and target languages, URLs of the articles, and translation statistics indicate the extent of human versus machine translation.
- **Parallel Text Dump**: The tool also allows access to parallel text data, encompassing the source content and the final user-translated content.

This data is provided under an open license and represents an ever-improving resource for cross-linguistic research and machine translation development. For more information, refer to the MediaWiki page: Content Translation - Published Translations

A.2 Wikipedia Knowledge Gap Index

The Knowledge Gap Index developed by the Wikimedia Foundation is instrumental in quantifying knowledge disparities across various dimensions on Wikipedia. The dataset is structured to provide a clear mapping of content gaps across languages

- **Content Mapping:** The Knowledge Gap Index employs a content mapping system, categorizing Wikipedia articles into specific domains and themes. This system allows us to assess disparities in content availability and representation across diverse topics and regions.
- **Gender Gap Measurement:** the Knowledge Gap Index addresses the Gender Gap. This measure assesses the balance in representation of different genders across Wikipedia articles. It aids in quantifying and understanding the skewness in gender portrayal,
- **Geographic Gap Measurement:** Another useful aspect of the Knowledge Gap Index is its focus on the Geographic Gap. This measurement evaluates the extent of content coverage pertaining to different geographic regions.
- **Open Access and Continuous Updates:** The Knowledge Gap Index is publicly accessible and is regularly updated, reflecting the dynamic nature of content creation on Wikipedia.

The Knowledge Gap Index offers a detailed perspective on the existing content distribution and the potential areas where machine translation can have a significant impact. For further details about the dataset and its applications, please refer to the Wikimedia Meta-Wiki page on the Knowledge Gap Index and its Geographic Gap Measurement.

B Methodology

B.1 Classifying Topics in Wikipedia Articles

Our approach to classifying article topics across different languages in Wikipedia builds on the language-agnostic methodology developed by Johnson et al. (2021). This method involves several steps:

1. **Topic Taxonomy**: Utilizing a taxonomy of 64 topics derived from WikiProject tags, encompassing a broad range of subjects from science and technology to arts and culture.

2. **Link-Based Representation**: Articles are represented by Wikidata items of contained links, facilitating a language-independent content representation.

3. **Mapping Articles to Topics**: Machine learning algorithms map articles to the predefined topics based on link similarities.

4. **Ensuring Cross-Linguistic Consistency**: The focus on link-based features rather than textual content assures consistent classification across languages.

5. **Validation**: Classification accuracy is validated and refined through comparison with a manually classified article sample.

This methodology enables a comprehensive and consistent analysis of article topics across multiple language editions of Wikipedia, overcoming the limitations of language-specific models and ensuring uniform topic classification.

B.2 Extracting Attributes from Wikidata

We utilized Wikidata to extract demographic and geographical data for our study. We used the Wikidata Query Service, which allows for sophisticated and precise queries, to retrieve the required information efficiently and accurately

- **Biographical Articles**: Information such as gender, sexual orientation, and ethnicity of individuals was systematically extracted from their respective Wikidata entries. This process involved querying Wikidata's structured data to obtain these specific attributes for each biographical article.

- **Geographical Articles**: For articles pertaining to geographical locations, we extracted location data, including coordinates and administrative hierarchies. This was achieved by querying the corresponding Wikidata items linked to each geographical article.

This methodology ensured a standardized and accurate extraction of demographic and geographical information, useful for our analysis.

C Robustness for Difference-in-Differences Modeling

C.1 Matrix Completion Method Estimation

The matrix completion method provides additional validation of our main findings on Google Translate's impact. As shown in Table A.4, The matrix completion method approach estimates an Average Treatment Effect on the Treated (ATT) of 71.2 additional monthly translations ($p<0.01$), with a 95% confidence interval of [38.4, 104.1]. This result aligns with our primary difference-in-differences estimates and the two-way fixed effects model, providing robust evidence that Google Translate integration substantially increased translation volume across Wikipedia languages.

Table A.4: Matrix Completion Method: Estimated Impact of Google Translate Integration on Monthly Translation Volume

ATT	S.E.	CI.lower	CI.upper	P.value
71.2	16.8	38.4	104.1	0.0

Note: This table presents the Average Treatment Effect on the Treated (ATT) estimated using the matrix completion method. The ATT of 71.2 indicates the average increase in monthly translations attributable to Google Translate integration. S.E. represents standard error, CI shows 95% confidence interval bounds, and P.value indicates statistical significance.

C.2 Two-way Fixed Effect

This model estimates the Average Treatment Effect (ATE) of Google Translate's integration on the volume of translated articles. By incorporating both language and time fixed effects, this approach controls for time-invariant language characteristics and common temporal shocks across languages. The statistically significant treated coefficient of 61.4 ($p<0.01$) in Table **??** reinforces the substantial impact of Google Translate on content translation activities in Wikipedia, while accounting for potential confounding factors through the two-way fixed effect structure.

Table A.5: Two-way Fixed Effects Analysis of Google Translate Integration Impact

Dependent Variable:	Number of Translated Articles
Model:	(1)
Google Translate Integration	61.4***
	(14.9)
Fixed-effects	
Language	Yes
Year-Month	Yes
Fit statistics	
Observations	6,336
R^2	0.77

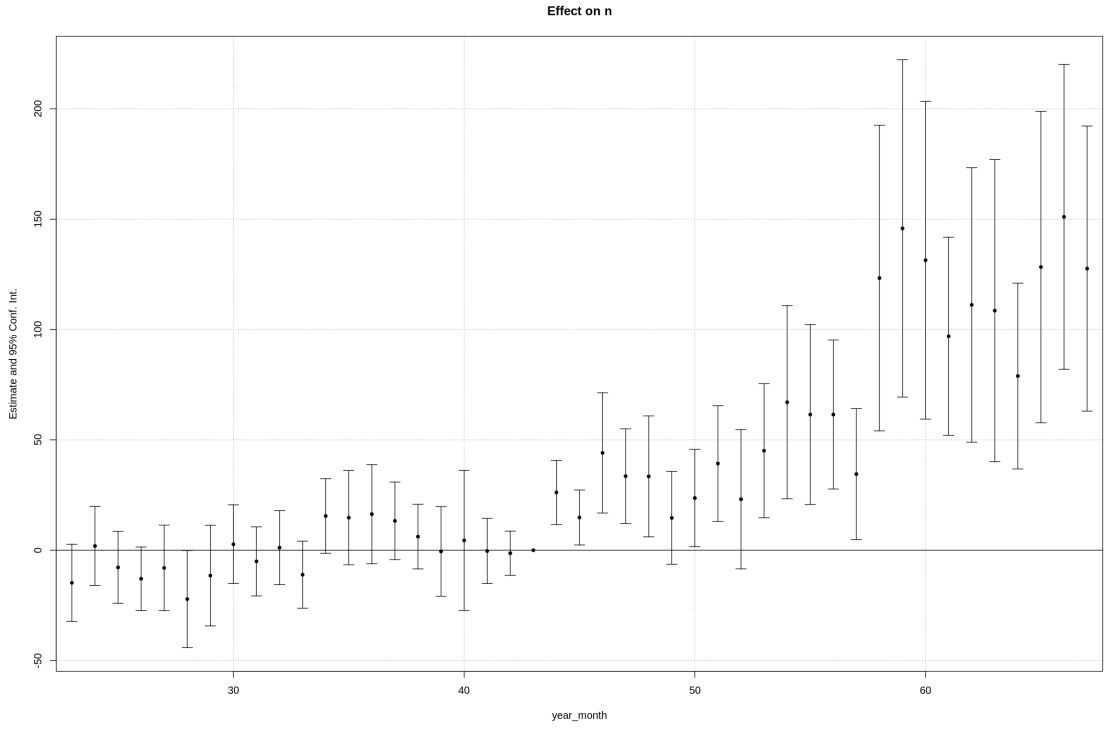
Clustered (Language) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: This table presents results from a two-way fixed effects model examining the impact of Google Translate integration on monthly translation volume. The coefficient of 61.4 represents the average increase in monthly translations after integration, controlling for both language-specific and temporal effects. Standard errors are clustered at the language level.

C.3 Sun and Abraham Weighted Interactions

The Sun and Abraham Weighted Interactions method addresses potential biases in staggered difference-in-differences designs by properly accounting for treatment effect heterogeneity. This approach computes cohort-specific treatment effects and combines them using weights that reflect the relative size of each cohort. The method highlights the temporal dynamics of the treatment effect, indicating a sustained positive impact of Google Translate over time on the translation volume. By explicitly modeling treatment effect heterogeneity across adoption cohorts, this approach provides more reliable estimates of the dynamic treatment effects.

Figure A.18: Dynamic Treatment Effects of Google Translate Integration Using Sun and Abraham Method

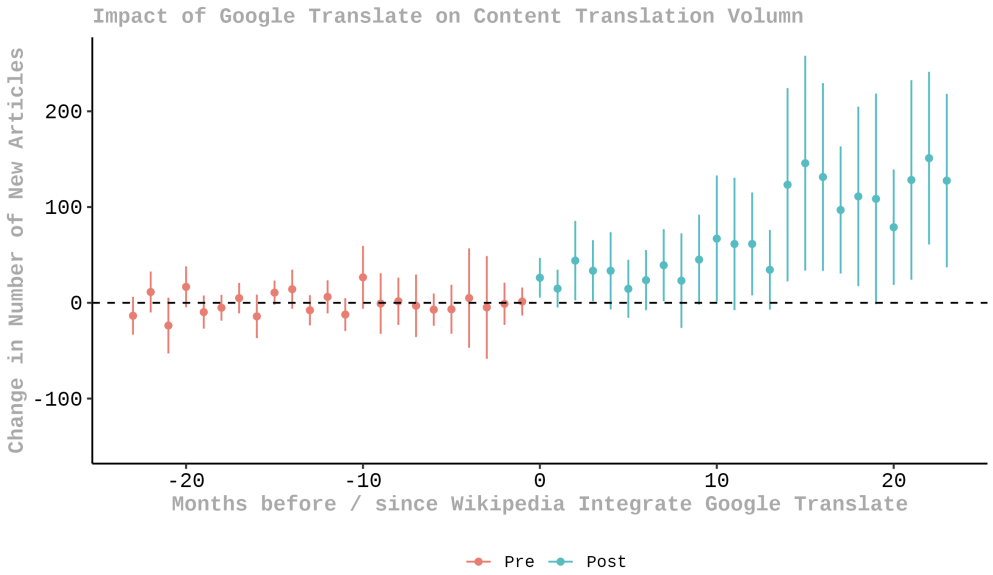


Note: This figure displays the dynamic treatment effects estimated using the Sun and Abraham weighted interactions method. The vertical axis shows the estimated effect on monthly translation volume, while the horizontal axis represents time relative to Google Translate integration. The solid line indicates point estimates, and the shaded area represents 95% confidence intervals. The results demonstrate persistent positive effects of integration across different time periods, accounting for heterogeneous treatment effects across adoption cohorts.

C.4 Callaway and Sant'Anna Estimators

We further assess the robustness of our findings using the Callaway and Sant'Anna Estimators, which are specifically designed for settings with staggered treatment adoption and heterogeneous treatment effects. This methodology offers several advantages: it allows for treatment effect heterogeneity both across groups and over time, remains valid under parallel trends assumptions, and is robust to negative weights that can arise in two-way fixed effects models. The estimators compute group-time average treatment effects and aggregate them into overall effects using carefully chosen weights. The Callaway and Sant'Anna Estimators confirm the robustness of our main findings, showing consistent positive effects of Google Translate's integration on translation volume across various Wikipedia language editions.

Figure A.19: Group-Time Average Treatment Effects Using Callaway and Sant'Anna Estimators



Note: This figure presents the estimated group-time average treatment effects using the Callaway and Sant'Anna method. The vertical axis shows the estimated effect on monthly translation volume, while the horizontal axis represents time relative to Google Translate integration. The solid line shows point estimates, and the shaded area represents 95% confidence intervals. The results demonstrate robust positive effects across different adoption groups and time periods, avoiding potential biases from negative weighting in traditional two-way fixed effects models.

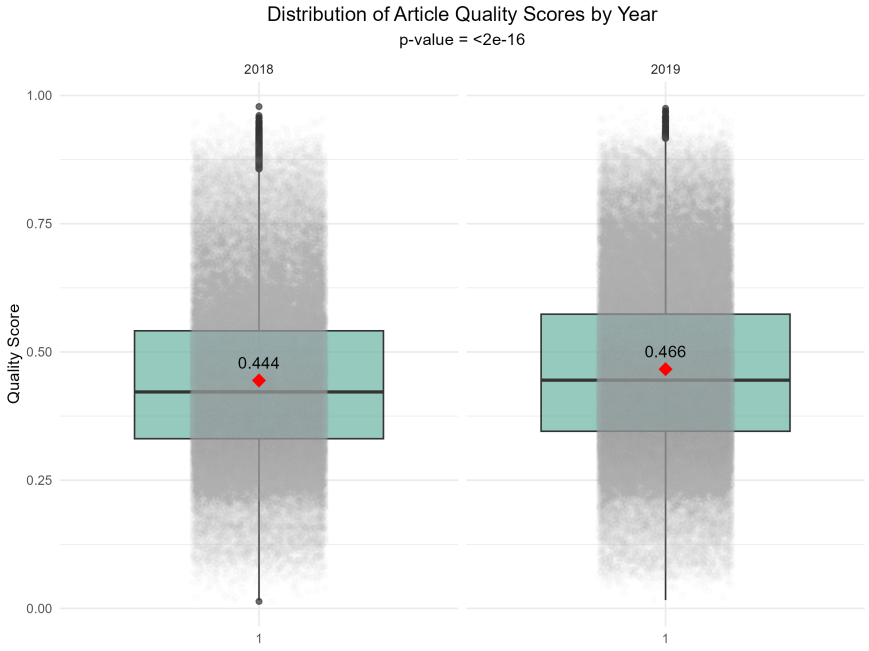
D Additional Supporting Results

The results presented in this section provide detailed model estimates and statistical evidence supporting the key findings discussed in the main paper. While the main text focuses on the high-level implications and interpretation of our analyses, here we present tables showing the full model specifications, robustness checks, and additional statistical tests.

D.1 Article Quality Analysis

We analyze the evolution of article quality between 2018 and 2019. The results show a modest but statistically significant improvement in article quality scores over this period, with the mean quality score increasing from 0.444 in 2018 to 0.466 in 2019. The distribution of scores, as visualized through both box plots and density plots, reveals that while there is considerable variation in article quality within each year, there is a consistent upward shift in the overall quality distribution from 2018 to 2019.

Figure A.20: Distribution of Article Quality Scores by Year (2018-2019)



Note: This figure shows the distribution of article quality scores comparing 2018 and 2019. The box plots display the median, quartiles, and range of quality scores, with the underlying data points shown as a density plot in gray. The red diamond indicates the mean quality score for each year (0.444 in 2018 and 0.466 in 2019). The extremely low p-value (<2e-16) suggests a statistically significant difference between the two years' distributions. The plot demonstrates a slight improvement in average article quality from 2018 to 2019.

D.2 Language Community Characteristics

Using the difference-in-differences estimator described in the Data and Methods section, we examine how the impact of Google Translate varies across different dimensions of language community characteristics. The results in Table A.6 show heterogeneous effects across language tiers based on article count, editor count, view count, and Google Translate quality. When languages are categorized by article count, medium-tier languages saw an increase of 78.3 articles per month and high-tier languages experienced an increase of 181.3 articles compared to the baseline, both statistically significant. Similar patterns emerge when languages are categorized by editor and view counts, with high-tier languages showing the largest gains (256.6 articles for both metrics). Interestingly, when considering Google Translate quality, only medium-tier languages show a significant positive effect of 49.4 articles, while high-tier languages show no significant increase. These findings suggest that community size and engagement are crucial factors in leveraging machine translation technology, potentially more so than the underlying quality of the translation system.

Table A.6: Heterogeneous Effects of Google Translate Integration Across Language Community Characteristics

Dependent Variable:	Number of New Translations			
Language Tier Based On:	Article Count	Editor Count	View Count	GT Quality
Model:	(1)	(2)	(3)	(4)
Variables				
GT Integration × Low-tier	13.3	0.6	4.3	-10.2
	(9.5)	(5.8)	(6.9)	(20.5)
GT Integration × Medium-tier	78.3**	56.6***	52.9***	49.4**
	(30.2)	(19.9)	(19.8)	(23.8)
GT Integration × High-tier	181.3***	256.6***	256.5***	-19.9
	(59.7)	(73.0)	(73.0)	(29.0)
Fixed-effects				
Language	Yes	Yes	Yes	Yes
Year-Month	Yes	Yes	Yes	Yes
Fit statistics				
Observations	6,048	6,048	6,048	3,744
R^2	0.777	0.790	0.789	0.722
Within R^2	0.059	0.112	0.110	0.015

Clustered (Language) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: This table presents difference-in-differences estimates examining how Google Translate's impact varies across language community characteristics. Each column represents a different categorization of languages: by article count, editor count, view count, and translation quality. The coefficients show the differential impact across low, medium, and high-tier languages, with standard errors clustered at the language level. The results demonstrate that community size and engagement significantly influence the effectiveness of machine translation integration.

D.3 Content Domain Analysis

Our difference-in-differences analysis reveals varying impacts across content domains, as shown in Table A.7. Culture and Geography articles saw the largest increases (25.2 and 21.4 articles per month respectively), while STEM and History/Society topics experienced more modest growth (7.9 and 5.6 articles). This pattern suggests that Google Translate may be more effective at handling cultural and geographical content compared to technical or

Table A.7: Differential Impact of Google Translate Integration Across Content Domains

Dependent Variables:	Culture	Geography	STEM	History and Society
Model:	(1)	(2)	(3)	(4)
GT Integration	25.2***	21.4***	7.9***	5.6**
	(5.9)	(5.4)	(2.1)	(2.3)
Fixed-effects				
Language	Yes	Yes	Yes	Yes
Year-Month	Yes	Yes	Yes	Yes
Fit statistics				
Observations	6,336	6,336	6,336	6,336
R^2	0.76	0.64	0.62	0.58
Pre-integration Mean	20.8	12.5	7.2	4.3

Clustered (Language) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: This table presents difference-in-differences estimates showing how Google Translate's impact varies across different content domains. The coefficients represent the average monthly increase in translations for each content type after integration. Standard errors are clustered at the language level, and pre-integration means are provided for context. The results reveal stronger effects for cultural and geographical content compared to technical and historical topics.

historical material.

D.4 Gender Analysis

The difference-in-differences estimates in Table A.8 reveal persistent gender disparities in biographical translations after Google Translate integration. Male biographies increased by 7.2 articles monthly compared to 3.8 for female biographies, while non-binary biographies saw minimal growth (0.06). This suggests that machine translation may reinforce existing gender biases in content creation.

Table A.8: Differential Impact of Google Translate Integration on Biographical Content by Gender

Dependent Variables:	Male Biography	Female Biography	Non-binary Biography
Model:	(1)	(2)	(3)
GT Integration	7.2***	3.8***	0.06***
	(2.5)	(1.3)	(0.02)
Fixed-effects			
Language	Yes	Yes	Yes
Year-Month	Yes	Yes	Yes
Fit statistics			
Observations	6,288	6,288	6,288
R^2	0.72317	0.54092	0.26982
Pre-integration Mean	12.5	7.8	0.06

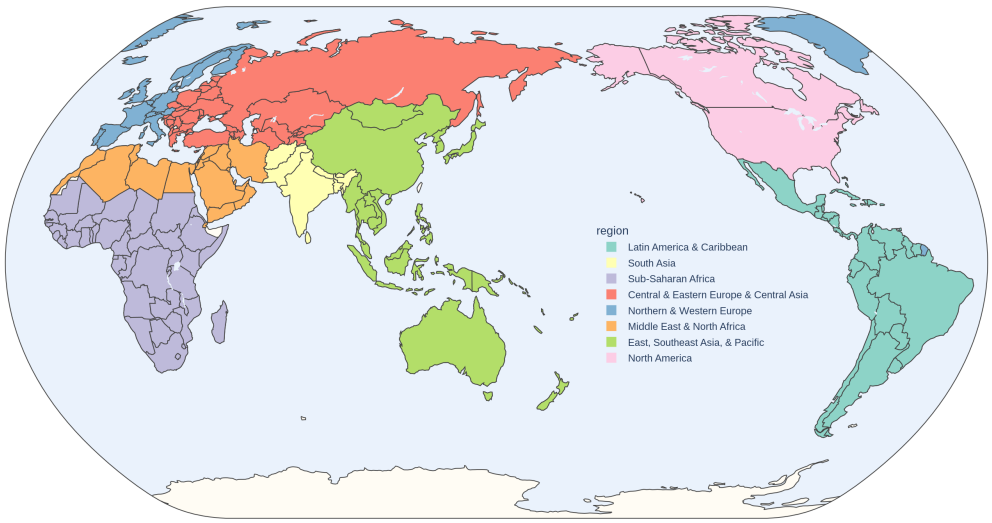
Clustered (Language) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: This table presents difference-in-differences estimates showing how Google Translate integration affects the translation of biographical articles across gender categories. The coefficients represent the average monthly increase in translations for each gender category after integration. Standard errors are clustered at the language level, and pre-integration means are provided for context. The results reveal a persistent gender gap, with male biographies receiving substantially more translations than female or non-binary biographies.

D.5 Geographic Analysis

Figure A.21 illustrates the geographical framework we use to analyze regional patterns in Wikipedia translations. By adopting the Wikimedia Foundation's official regional classification system, which divides the world into eight distinct regions, we can systematically examine how Google Translate's impact varies across different geographical and cultural contexts. This standardized regional breakdown enables us to identify potential disparities in translation activity and technological adoption across diverse global regions.

Figure A.21: Wikimedia Foundation's Regional Classification System for Global Content Analysis



Note: This world map illustrates the eight regional classifications used by the Wikimedia Foundation for content analysis: Latin America & Caribbean (LAC), South Asia (SA), Sub-Saharan Africa (SSA), Central & Eastern Europe & Central Asia (CEECA), Northern & Western Europe (NWE), Middle East & North Africa (MENA), East/Southeast Asia & Pacific (ESEAP), and North America (NA). These regions form the basis for analyzing geographical patterns in Wikipedia translation activity and content distribution.

Using our difference-in-differences framework, we examine geographical coverage across eight major world regions defined by the Wikimedia Foundation (shown in Figure A.21). The results in the following table show that Central & Eastern Europe Central Asia (CEECA) and East/Southeast Asia & Pacific (ESEAP) regions saw the largest increases in geographical article translations (8.6 and 7.3 articles monthly), while regions like Middle East & North Africa (MENA) and Sub-Saharan Africa (SSA) experienced minimal growth (0.3 and 0.4 articles). This regional disparity suggests that machine translation may be reinforcing existing geographical biases in content coverage.

Table A.9: Effects of Google Translate Integration on Monthly Geography Translations by Regions

DVs:	ESEAP	SA	MENA	SSA	NA	SCA	CEECA	NWE
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
GT Integration	7.3***	5.3***	0.3**	0.4**	1.7***	1.3**	8.6***	5.4***
	(2.4)	(1.8)	(0.1)	(0.2)	(0.5)	(0.5)	(2.6)	(1.4)
Fixed-effects								
Language	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Month	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics								
Observations	6,336	6,336	6,336	6,336	6,336	6,336	6,336	6,336
R^2	0.49	0.44	0.54	0.26	0.47	0.24	0.55	0.60
Pre-integration Mean	4.7	2.8	0.2	0.2	1.0	0.6	5.0	3.4

Clustered (Language) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: This table presents difference-in-differences estimates showing the impact of Google Translate integration on monthly article translations across eight Wikimedia regions. ESEAP: East/Southeast Asia & Pacific; SA: South Asia; MENA: Middle East & North Africa; SSA: Sub-Saharan Africa; NA: North America; SCA: Latin America & Caribbean; CEECA: Central & Eastern Europe & Central Asia; NWE: Northern & Western Europe. The coefficients represent the average monthly increase in translations for each region after integration. Standard errors are clustered at the language level, and pre-integration means are provided for context.

D.6 Geographical Wikipedia Articles by Continents

The Table A.10 presents the difference-in-differences analysis results, capturing the impact of Google Translate's roll-out on the volume of articles across different continents. The results show that regions like Asia and Europe experienced a significant increase in article volume, with coefficients of 9.0 and 7.8, respectively, suggesting a more pronounced adoption of machine translation in these areas. America and Africa also saw increases, albeit at a lower magnitude, while Oceania demonstrated a modest but significant rise. These findings reflect the variable impact of machine translation across diverse geographical contexts.

Table A.10: Impact of Google Translate Roll-out on Article Volume by Geography Regions

Dependent Variables:	America	Europe	Asia	Oceania	Africa
Model:	(1)	(2)	(3)	(4)	(5)
Google Translate Integration	3.0***	7.8***	9.0***	0.5***	1.5**
	(1.0)	(2.0)	(3.3)	(0.2)	(0.6)
Fixed-effects					
Language	Yes	Yes	Yes	Yes	Yes
Year-Month	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	6,336	6,336	6,336	6,336	6,336
R^2	0.42505	0.58168	0.42917	0.29379	0.45671
Within R^2	0.00453	0.00887	0.00744	0.00226	0.00466

Clustered (Language) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: This table presents difference-in-differences estimates showing the impact of Google Translate integration on monthly article translations across five continental regions. The coefficients represent the average monthly increase in translations for each continent after integration. Standard errors are clustered at the language level. The results reveal substantial variation in the effect across continents, with Asia and Europe showing the largest increases in translation volume.

E Analysis of Potential Displacement Effects

This section examines potential displacement effects of Google Translate integration on non-translated content creation and consumption. We analyze whether increased machine-translated content affects the production and readership of native content through resource reallocation or shifts in reader attention.

Data Collection We constructed a comprehensive panel dataset tracking monthly article creation across both translated and non-translated content in treated and control Wikipedia language editions from 2017 to 2020. This dataset captures content creation through all methods across more than 100 Wikipedia language communities, extending beyond articles created through the Content Translation system. Additionally, we collected monthly aggregated pageview statistics sourced from billions of records across Wikipedia language editions to analyze content consumption patterns.

Content Creation Analysis Using this rich dataset, we employed two model specifications examining:

1. The relationship between translated and non-translated content using fixed-effects regression controlling for language and time fixed effects

2. The impact of Google Translate integration on non-translated article creation using difference-in-differences, where the treatment effect is measured as the interaction between treatment group and post-integration period

For each specification, we analyzed both standardized article counts and log-transformed article counts to ensure robustness of our findings.

Table A.11: Analysis of Translated vs. Non-Translated Article Creation

	Dependent Variable: Non-Translated Article Creation			
	Fixed Effects Model		Difference-in-Differences	
	Standardized	Log	Standardized	Log
Variables	(1)	(2)	(3)	(4)
Number of Translated Articles	-0.05	0.04		
	(0.04)	(0.03)		
Google Translate Integration			-0.18	-0.30
			(0.18)	(0.18)
Language Fixed Effects	Yes	Yes	Yes	Yes
Year-Month Fixed Effects	Yes	Yes	Yes	Yes
Observations	5,856	5,856	5,856	5,856
R^2	0.20	0.84	0.20	0.85

Notes: Table reports regression results examining whether increased translated content displaces non-translated article creation. Standardized variables have mean zero and standard deviation one. Standard errors (in parentheses) are clustered at the language level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Content Consumption Analysis To examine potential displacement effects in article consumption, we analyzed the relationship between translated and non-translated article pageviews using similar methodological approaches:

1. A fixed-effects regression examining the direct relationship between translated and non-translated article pageviews

2. A difference-in-differences specification analyzing how Google Translate integration affects non-translated article pageviews

The analysis incorporated both standardized and log-transformed pageview counts to ensure robust findings across different specifications.

Table A.12: Analysis of Translated vs. Non-Translated Article Consumption

	Dependent Variable: Non-Translated Article Pageviews			
	Fixed Effects Model		Difference-in-Differences	
	Standardized	Log	Standardized	Log
Variables	(1)	(2)	(3)	(4)
Number of Translated Pageviews	0.12***	0.18***		
	(0.03)	(0.05)		
Google Translate Integration			0.02	0.03
			(0.012)	(0.04)
Language Fixed Effects	Yes	Yes	Yes	Yes
Year-Month Fixed Effects	Yes	Yes	Yes	Yes
Observations	5,856	5,856	5,856	5,856
R^2	0.99	0.99	0.99	0.99

Notes: Table reports regression results examining whether increased translated content affects pageviews of non-translated articles. Standardized variables have mean zero and standard deviation one. Standard errors (in parentheses) are clustered at the language level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Our analysis reveals several findings. For content creation, we find no significant negative effects across all specifications, indicating that increased translated content does not displace native content production. For content consumption, we observe a positive and significant relationship between translated and non-translated article pageviews in the fixed effects model (columns 1-2), suggesting complementarity rather than substitution. The difference-in-differences analysis (columns 3-4) shows no significant negative effect of Google Translate integration on non-translated article pageviews. When examining combined normalized pageviews of source and target articles, we observed an increase from a baseline of 1.0 to approximately 1.5, indicating that translations attract new readers while maintaining

the readership of original content.

These findings demonstrate that the integration of Google Translate expands Wikipedia's content ecosystem rather than creating displacement effects. Machine translation appears to complement rather than substitute for native content creation and consumption, contributing to the overall growth of Wikipedia's knowledge base across languages.

F Increased Participation vs. Editor productivity

We examine individual editor contributions following the integration of Google Translate and examine the dual forces of participation and productivity among Wikipedia editors. This analysis aims to dissect the increase in content production and discern whether it stems from a growth in the editor base or improved editor productivity.

For this investigation, we collect editor IDs for each article created using the Content Translation system in our sample, enabling us to pinpoint the respective editors behind translated articles[8]. We used this information to create a language-month panel and calculate the number of new editors who used translation to create an article for the first time, the number of unique active editors contributing at least one article in a specific language and month, and the productivity of each editor, measured as the total articles created in the given month. For simplicity and robustness, we present results from analysis focused on language editions that integrated Google Translate in January 2019 and utilize two-way fixed effect models to analyze the editor panels in Table A.13.

As detailed in Table A.13, our analysis reveal a 49% growth in monthly new editors (an increase of 2.1 new editors beyond the pre-integration baseline of 4.3), a 66% increase in monthly active editors (an increase of 6.3 from the pre-integration baseline of 9.5), and a 69% increase in number of new translated articles per editor[9] (an increase of 4.5 from the pre-integration baseline of 6.5). Therefore, the magnitude of these effects indicate that the increase in translation volume stems from both existing editors becoming more active and productive as well as an expansion of the editor base. This shift towards greater editor participation and productivity underscores Google Translate's role in empowering Wikipedia editors.

Moreover, Figure A.22 presents the temporal trend of editor productivity, providing a model-free and intuitive visual substantiation to our econometric estimates. The figure

[8]For privacy consideration, only editor ID information is accessible on the website.

[9]Model (3) is analyzed at the editor level, as opposed to target language, and hence has a different number of observations

Table A.13: Editor Engagement and Productivity Analysis Post-Google Translate Integration

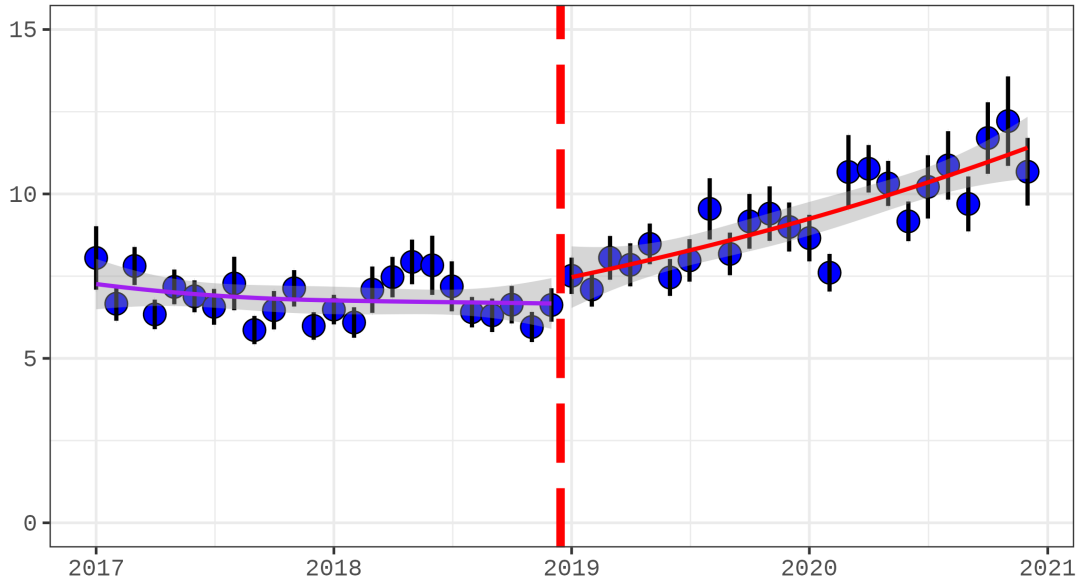
Dependent Variables:	New Editor	Active Editor	Article/Editor
Model:	(1)	(2)	(3)
GT Integration	2.1***	6.3***	4.5***
	(0.6)	(1.5)	(1.3)
Fixed-effects			
Language	Yes	Yes	Yes
Year-Month	Yes	Yes	Yes
Fit statistics			
Observations	6,336	6,336	3,979
R^2	0.83264	0.91247	0.13120
Pre-integration Mean	4.3	9.5	6.5

Clustered (Language) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

illustrates a steady temporal pattern in the monthly number of new translated articles per editor leading up to January 2019, after which there is a discernible upward trend. Individual editor productivity increases from an average of 6.8 article per editor/month in the pre-integration period to around 12 articles toward the end of our sample period. The error bars suggests variability in the data, yet the fitted LOESS curve shows that, on average, editors have been producing more articles each month since the integration of Google Translate.

Interestingly, the Interrupted Time Series analysis reveals a modest 10% increase in the level at the time of interruption for editor productivity ($\beta_2 = 0.7$ more articles relative to a 6.8 article pre-interruption baseline). However, there is a substantive positive shift between the post-interruption slope and pre-interruption slope ($\beta_3 = 0.19$ articles per editor/month). These findings imply that the full extent of the enhancement unfolds gradually as users become more adept with using the technology. Hence, while the tool lowers the skill barrier for editors to contribute, it should be complemented with strategic efforts to improve editor skill and experience with the new technology.

Figure A.22: Evolution of Editor Productivity on Wikipedia Over Time



Notes: This graph presents the trends in monthly number of new translated articles per editor on Wikipedia, illustrating a steady rise in productivity following the rollout of Google Translate in 2019. The fitted lines and shaded areas result from a piecewise LOESS regression.