Does Watching Women Work, Work?

The Effect of Television on Female Labor Outcomes*

Javier Gonzalez

Southern Methodist University
gonzalezjj@smu.edu

October 2025

Latest Available Version

Abstract

I study the effect of female empowerment in media on female labor market outcomes using Latin American telenovelas. Using generative AI, I construct a Female Empowerment Index (FEI) for these TV shows from 1960 to 2024. I show that FEI exposure during the impressionable years increases the likelihood of labor force participation among Latin American women. To identify the causal effect of FEI exposure, I implement an instrumental variables strategy using detailed data on television signal coverage in Mexico. The causal estimates confirm the positive effect of FEI exposure, with an effect size comparable to previous studies. Furthermore, FEI changes account for a significant fraction of the aggregate increase in female labor force participation observed over this period. Moreover, both the emotional framing of narratives and the types of jobs depicted play an important role in shaping the effects.

JEL Clasification: D91, J16, J22, L82, Z13

Keywords: Media, Gender Gap, Social Norms, Latin America, Mass Media, Generative AI, Telenovelas

* Thanks to Ömer Özak, Klaus Desmet, Wookum Kim, Ivan de las Heras, Rouzhi Liang, and Esteban Aleman for their feedback and support. Special thanks to the SMU Economics Department Ph.D. students and staff, as well as the O'Donnell Data Science and Research Computing Institute, for providing access to high-performance computing resources.

1 Introduction

Over the past century, women have made remarkable progress in closing labor market gender gaps. Yet, a small but persistent gap in wages and participation remains. To explain this stubborn gap, a growing body of economic research highlights the powerful role of culture and gender norms as part of the puzzle (Bertrand, 2020). At its core, culture is transmitted through shared narratives that shape beliefs and expectations, whether through direct experiences within one's community or exposure to new ideas and perspectives. While much of this literature examines socialization within families and communities, less is known about how large-scale cultural products, such as television, transmit and reshape these norms (Olivetti, Pan and Petrongolo, 2024). By creating an emotional connection with its audience, television can either reinforce or challenge established gender norms, thereby influencing women's labor market outcomes.

This study examines the cultural influence of television by asking: How does exposure to empowered female characters affect women's labor force participation? To answer this question, I focus on telenovelas, melodramatic television shows from Latin America, due to their cultural significance and widespread viewership (Lopez, 2002; Pastina, Rego and Straubhaar, 2003; Antezana *et al.*, 2022). My approach centers on a novel measure of female empowerment constructed from a new dataset of approximately 2,000 telenovelas aired in Mexico, Brazil, and Chile from 1960 to 2024. Specifically, I utilize Google's Gemini AI to quantify and encode the empowerment level in each show's synopsis based on a modified version of the Bechdel-Wallace test (Appel and Gnambs, 2023) augmented with two questions concerning the employment status of female characters.[1] The resulting Female Empowerment Index (FEI) is the sum of the model's answers for each telenovela. To ensure reliability, I validated the model's performance against my own categorization, finding a strong predictive fit that yielded a large-scale, systematic measure of female empowerment in entertainment media.

I find a strong positive correlation between telenovela exposure during adolescence, a critical period for the formation of gender norms, and women's later labor market participation across several census waves in Latin America.[2]. Furthermore, I show that greater exposure to empowered female characters in telenovelas is linked to more progressive gender attitudes, measured by the World Value Survey. Taken together, these findings suggest that telenovelas play a significant role in shaping gender norms and subsequently influencing women's labor market outcomes in Latin America.

To address potential endogeneity concerns, I exploit geographic and temporal variation in television signal reception across Mexico to identify the causal effect of exposure to empowered female characters. Because telenovelas were historically broadcast over analog airwaves, households farther from transmission towers received weaker signals and thus had poorer reception. Signal reception can also be diminished by obstacles between the household and the tower, such as mountains or dense forests, which block or weaken the broadcast. This naturally occurring variation, combined with the staggered rollout of new broadcasting towers over time, creates plausibly exogenous differences in exposure across cohorts and locations. I use the minimum signal loss from all towers in each location during individuals' teenage years as an instrument for telenovela exposure. This instrument captures the best possible signal quality available in a given area and is uncorrelated with local development or public service expansion, supporting its validity.

[1]The Bechdel-Wallace test was developed in 1985 by Alison Bechdel and Liz Wallace to measure the representation of female characters in films and other media. It consist of 3 questions: (i) are there at least two female characters? (ii) do they talk to each other? (iii) do they talk about something other than men?

[2]Extensive psychological and economic research supports the "impressionable years" hypothesis, which posits that adolescence is a critical period for the formation of values (Krosnick and Alwin, 1989; Malmendier and Nagel, 2011; Etchegaray, Scherman and Valenzuela, 2019). For review on this formative period for gender norms please refer to Ellemers (2018)

To implement this strategy, I focus on Mexico, where three datasets allow me to calculate signal loss at the county level. First, I use individual-level labor market microdata from the Bank of Mexico, which provides detailed information on women's employment outcomes and demographics at the county level. Second, I assemble a panel of telenovela broadcasts that maps their content, empowerment scores, and geographic coverage from 1960 onward. Finally, I merge these data with predicted maps of television signal loss to construct an exogenous measure of exposure during adolescence. This framework effectively compares women from the same birth cohort living in similar labor markets but facing different levels of access to telenovela signals, allowing me to isolate the causal effect of empowered portrayals on later labor market participation.

The causal estimates show that women exposed to higher levels of FEI during their teenage years are more likely to participate in the labor force as adults. A 10% increase in FEI exposure during adolescence leads to a 4% increase in the probability of labor market participation, which corresponds to a 5 percentage point increase. The estimated effects are sizable, with effect sizes comparable to those found for other cultural shocks on female labor force participation (Fernández, Fogli and Olivetti, 2004; Fernández and Fogli, 2009; Charles, Guryan and Pan, 2022). A back-of-the-envelope calculation suggests that FEI explains about 8.4% of the labor force participation changes between 2010 and 2020. The impact of exposure to female-empowered content is most pronounced among older women and those with lower levels of education, with negligible effects on women with a university degree. This exposure significantly increases women's educational attainment, wages, and the likelihood of becoming a manager. Furthermore, FEI leads to a decrease in fertility but has a limited impact on major family decisions like marriage or occupational choices in male-dominated fields.

The way female empowerment is portrayed helps explain what drives these effects. To examine these mechanisms further, I conduct two tests. First, I assess whether the emotions attached to empowered content have differential effects on outcomes. Portrayals of empowerment linked to joy or anger serve as powerful motivators of female labor market participation. In contrast, portrayals that associate empowerment with sadness or fear appear to act as cautionary tales, correlating with lower participation rates. Second, I examine whether the type of job portrayed leads to different impacts. The results suggest that increases in labor participation are primarily driven by exposure to female characters in non-traditional careers (e.g., doctor or lawyer). Exposure to these characters also increases the likelihood that viewers enter male-dominated fields themselves. However, when empowerment is depicted in traditional female jobs (e.g., nurse or teacher), the effects on participation and career choice disappear or even reverse.

This paper provides causal evidence that entertainment media are not merely a reflection of culture, but an active ingredient in its formation. The findings demonstrate that the narratives embedded in popular telenovelas translate into tangible changes in women's economic lives, showing that changing gender norms is not just about changing laws and policies, but about changing the stories we tell.

1.1 Literature Review

This work contributes to the literature that analyzes the role of media in shaping individuals' social attitudes and behavior. Mass media (e.g., news, entertainment, and social media) play a central role in shaping economic perceptions (Soroka, 2014). Pioneering studies have focused on media bias in newspapers and radio and its impact on political outcomes (Besley and Burgess, 2002; Strömberg, 2004; Arceneaux and Johnson, 2013; Drago, Nannicini and Sobbrio, 2014; Gentzkow, Shapiro and Stone, 2015; Martin and Yurukoglu, 2017; Ash and Hansen, 2023). Recently, researchers have explored the effects of television news on various social and political behaviors, such as voter turnout and teenage education outcomes (Gentzkow, 2006; Gentzkow and Shapiro, 2008).

In contrast to traditional media, entertainment media offer a fictionalized but emotionally engaging representation of social reality. Entertainment media can influence political attitudes by immersing viewers in these fictional worlds and creating emotional connections that shape beliefs and opinions (Green, Brock and Kaufman, 2004; Holbrook and Hill, 2005; Morgan and Shanahan, 2010). Economists have shown the effects of entertainment media on the acceptance of domestic violence (Jensen and Oster, 2009), teenage pregnancy (Kearney and Levine, 2015), children's educational outcomes (Kearney and Levine, 2019), poverty reduction (Ferrara, 2016), HIV attitudes (Banerjee, La Ferrara and Orozco-Olvera, 2019), and approval for domestic violence (Banerjee, Ferrara and Orozco, 2019).

Telenovelas offer a unique perspective because of their role in Latin American culture (Lopez, 2002; Green, Brock and Kaufman, 2004; Antezana *et al.*, 2022; 2023). Researchers have focused on the effects of telenovelas on demographic changes and support for minorities. For instance, La Ferrara, Chong and Duryea (2012) and Chong and La Ferrara (2009) show how the entrance of Rede Globo, which essentially introduced telenovelas to viewers in Brazil, decreased fertility and increased divorce filings, respectively. Moreover, Gulesci, Lombardi and Ramos (2024) provides evidence suggesting that exposure to characters from the LGBTIQ+ community in telenovelas reduces the support for said community in Latin American countries. This paper makes two main contributions to this strand of the literature. Unlike prior work, I go beyond measuring exposure to telenovelas and instead analyze which content characteristics drive the effects. This study also reveals a positive long-term causal effects of telenovelas on female labor market outcomes, a dimension that has not been systematically explored in the existing literature.

The mechanisms linking media to economic behavior often overlap with the broader literature on gender norms and labor market inequality. Bertrand (2020) suggests that gender stereotypes play a crucial role in shaping these disparities (Alesina, Giuliano and Nunn, 2013; Olivetti, Pan and Petrongolo, 2024). Several related studies have examined the impact of cultural shocks on female labor market outcomes, highlighting how shifts in cultural context can drive substantial changes in women's participation and economic opportunities. Fernández, Fogli and Olivetti (2004) find that wives whose husbands' mother worked are more likely to join the labor force. Fernández and Fogli (2009) show that changes in the labor force participation of a particular ethnic group, as a proxy for country of origin culture, is related to second-generation Americans female labor force participation. Fernández (2013) argues that gender norms transmitted from mothers to daughters can explain a significant portion of the rise in female labor market participation in the US after the 1950s. Another channel is the introduction of female role models. For instance, Porter and Serra (2020) finds that exposing students to female role models in economics classrooms substantially increases women's likelihood of majoring in the field. These studies underscore the importance of visible role models for changing career expectations, but most examine direct, institutional exposure (for example, in schools or politics). This paper contributes in two ways. First, it develops a new measure of gender empowerment in telenovelas and shows that this measure helps explain variation in women's labor market participation. Second, it highlights the power of indirect role models—fictional women in mass entertainment, as an underexplored but scalable mechanism for shifting gender norms and economic behavior.

Economists have primarily used language as a proxy for culture, grounded in a theory of group identity (Esteban and Ray, 1994; Guiso, Sapienza and Zingales, 2006). This approach views language as a primary carrier of a group's heritage and norms, positing that languages and other cultural traits co-evolve from common ancestral roots (Galor, Özak and Sarid, 2018). Consequently, the structural distance between languages can serve as a quantifiable proxy for broader, unobservable cultural divergence. This linguistic gap creates tangible economic friction, raising barriers to interaction and trust that can impede trade and migration (Duclos, Esteban and Ray, 2004; Michalopoulos, 2012; Desmet, Ortuño-Ortín and Wacziarg, 2017). Recent advances have turned to natural language processing and machine learning to analyze large volumes of text and media content at scale to extract cultural themes and sentiments (Michalopoulos and Xue, 2021; Apel, Blix Grimaldi and Hull, 2022; Haese, 2025; Clayton *et al.*, 2025). There are two papers in this literature that closely relate to this paper. Lippmann and Montalbo (2025) look at short summaries of advertisements in the Nielsen Ratings Data to extract how women are represented. Michalopoulos and Rauh (2024) use LLMs to categorize folklore tales with gender prevalence. This paper builds on previous work and uses a question-based approach to measure a dimension of gender norms: women's empowerment in relation to work.

2 Measuring Female Empowerment

Latin American telenovelas are popular melodramatic television shows, transmitted through open channels. They typically display complicated love stories and interpersonal conflicts, often set against a backdrop of social issues and cultural norms. Telenovelas play a central role in Latin American culture, reflecting the region's societal values and norms. They are widely broadcast across the region and watched by audiences of all ages—including teenagers, who often prefer them over streaming platforms such as Netflix (Lopez, 2002; Pastina, Rego and Straubhaar, 2003; Antezana *et al.*, 2022; 2023). This makes them an ideal setting to test whether exposure to empowerment narratives affects women's economic behavior.

2.1 Telenovela Data

Data on telenovelas was collected using various online sources tailored to each country. From each source, I gathered the written synopsis, airtime, first and last episode dates, number of episodes, broadcasting channel, and genre of the show (e.g., comedy, drama, or melodrama). Additionally, I extracted ratings for each telenovela from the IMDb website, indicating how well-received each telenovela was by its audience. Telenovelas aired in Chile were gathered from chilenovelas, a Wiki-like page that provides comprehensive lists of most telenovelas aired in the country from 1967 to 2024. Similar information for Brazilian telenovelas was sourced from teledramaturgia. For other Mexican telenovelas, lists are available on Wikipedia and users' IMDb lists. The primary input for measuring female empowerment in telenovelas is the written synopsis of each show. The synopses should be self-contained descriptions of each telenovela's plot and themes, including most of the information needed to identify and extract characteristics of the most important characters and their interactions. Each synopsis was translated into English using OpenAI's `gpt-4o-mini model`, which provides good quality translation from Spanish to English and performs reasonably well translating from Portuguese to English (Sanz-Valdivieso and López-Arroyo, 2023; Törnberg, 2023).[3]

The data does not necessarily contain the entire population of TV shows that aired in each country. First, a telenovela in the data is "a TV series, produced locally and broadcast on a local channel for a general audience in open TV via radio waves or satellite". Following the definition, the dataset excludes documentaries, investigative miniseries, or even reality TV shows produced locally. The dataset also excludes imported telenovelas. Most imported telenovelas tend to be adapted for the local audience. The most notable example is Mexico's hit *Rebelde* (2004), which was adapted from *Rebelde Way* which aired in Argentina in 2002. Other examples are Colombia's popular sensation *Yo soy Betty, la fea* (1999), which was adapted in Mexico as *La fea más bella* (2006) with less success. Second, the data may not include less popular or older telenovelas, as these shows are less likely to

[3]Other alternatives like DeepL and Google Translate come at a higher cost. See Hidalgo-Ternero (2020) for a discussion on the Spanish to English translation of these two alternatives.

Table 1: Telenovela Data Description

Panel A: All telenovelas (N = 1855)				
	mean	sd	min	max
Drama (%)	34.39			
Air time	16:38	07:09	10:00	23:15
Number of episodes	129	76	1	1018
Synopsis word count	1719	1681	17	20784
IMDB Rating	6.95	1.12	1.80	9.40

Notes: This table provides descriptive statistics for all collected telenovelas between 1960-present in Chile, Mexico, and Brazil. IMDb ratings represent user ratings for each telenovela, as listed on the IMDB website.

have dedicated fan bases or an online presence during the collection period.[4]

A total of 1855 telenovelas were gathered for Chile, Brazil, and Mexico.[5] Table 1 shows some descriptive statistics for all the telenovelas in my sample. Around 35% of telenovelas are dramas or melodramas, 20% are comedies, and the rest are a mix of comedy and drama. On average, a telenovela airs around 16:38 in the afternoon, consists of 129 episodes, and has an IMDb user rating of around 7 out of 10. The word count for each synopsis varies significantly, with a mean of approximately 2,000 words, but ranges from as few as 66 words to over 10,000 words.

2.1.1 Female Empowerment Index

While exposure to TV and telenovelas may influence gender norms and labor market outcomes, it is important to consider the specific content of these shows. Not all telenovelas are the same; the type of content, particularly the presence of empowered female characters, may play a key role in shaping viewers' gender norms.

To systematically assess how telenovelas portray female empowerment, I introduce the Female Empowerment Index (FEI) using recent advances in generative AI models (e.g., ChatGPT, Gemini, Deepseek) on telenovela summaries. Generative AI models have shown remarkable capabilities in categorizing text into groups not present in their training data (Wang, Pang and Lin, 2023).[6] Therefore, I can apply these tools to read through the telenovela synopsis and extract the relevant information about female characters and their degree of empowerment. Other researchers within the fields of social sciences, psychology, and economics have similarly employed them in different contexts with demonstrated reliability and even outperforming human experts in some cases (Törnberg, 2023; Rathje *et al.*, 2024; Michalopoulos and Rauh, 2024).[7] The synopsis of each telenovela should contain important information on whether there are important female characters in the story, their roles, how they interact with other characters, and importantly if they are employed or not.

I use Google's `Gemini 2.0-flash` API to identify several characteristics of female characters in telenovelas described in their synopses. First, I ask the model to perform the Bechdel-Wallace test

[4]An alternative approach to collecting these data would be to use digitized TV guides from each country. For example, a comprehensive database is publicly available through the *Television Factbook* for the U.S. However, no similar digitization efforts exist for the LATAM region.

[5]Chile, Brazil, and Mexico are the primary focus of the analysis because they have the most comprehensive and accessible online databases of telenovelas. In contrast, data for other countries is more limited and less systematically available.

[6]Text generative AI models, such as OpenAI's gpts, are based on Large Language Models (LLMs) which are probabilistic models trying to predict $P(\text{word}|\text{previous word})$ and create the most likely sentences given some previous words, sentences, or prompts. For a brief overview of these models I recommend 3Blue1Brown's YouTube video "Large Language Models explained briefly", Welch Labs YouTube video The moment I stopped understanding AI and for a more detailed explanation to read Chapter 1.3 and 1.10 of (Jurafsky and Martin, 2024).

[7]For instance, Michalopoulos and Rauh (2024) utilized OpenAI's `gpt-3.5-turbo` to classify film synopses into those with risk-taking attitudes and those with more traditional gender roles.

(Appel and Gnambs, 2023). This test was developed in 1985 by Alison Bechdel and Liz Wallace to measure the representation of female characters in films and other media. The test consists of three questions:

- (Q1) Are there at least two female characters with names?
- (Q2) Do these characters talk to each other?
- (Q3) Do they talk about something other than men?

I then augment the Bechdel-Wallace test with two additional questions regarding work:

- (Q4) Does a female secondary character work?
- (Q5) Does the female protagonist work?

Box 1 shows the prompt used to measure the FEI.[8] The model processes the instruction prompt along with the translated synopses, and outputs an answer for each question, a confidence score (ranging from 50 to 100), and a step-by-step explanation of its reasoning for each question. I repeat this exercise five times for each telenovela to ensure robustness and reliability of the results.[9]

To illustrate how these questions are designed to capture empowered female characters, I use Gabriela Suárez from *La Patrona* as an example.[10] Gabriela is a single mother working in a male-dominated gold mine, where, in the first five minutes of episode 1, she is harassed by colleagues. She immediately fights back against the abuse as shown by Figure 1. This would answer `true` to Q5 since the female protagonist has a job, and in fact is in a male-dominated field. In addition, this telenovela also showcases the strong and feared female antagonist, Antonia Guerra, who kills her husband to gain control over the family business. Both female characters show empowered women in two distinct ways: Gabriela through her resilience and fight against oppression, and Antonia through her ruthless ambition and strategic manipulation. With this information, we know that there are more than 2 female characters (Q1 = `true`) and Antonia has a job as a business owner (Q4 = `true`). Gabriela then fights for control of the mine with Antonia, suggesting that they talk to each other (Q2 = `true`) about the mine and not men (Q3 = `true`).

The Female Empowerment Index (FEI) is then the sum of the most agreed-upon answers for each question out of the five trials. For each question, if the model predicts `true` with over a confidence threshold δ in at least three out of five trials, the response is set to `true`; otherwise, it is set to `false`.[11]

Box 1: FEI prompt

```
You are a helpful research assistant. Read carefully the provided telenovelas summaries and identify the following:
• fem_1: if there are more than 2 female main characters in the show
• fem_2: if these female characters talk to each other
• fem_3: if these female characters talk about something other than men
• fem_4: if there is a female character with a job. Please also specify what job she has, if she does not have
  a job or if there are no female main characters please put 'na'.
• fem_5: if the protagonist is a female and working a job. Please also specify what job she has, if she does
  not have a job or there is a male protagonist please put 'na'.
Provide a confidence level where 50 (uncertain) and 100 (perfectly certain) for fem_1 to fem_7. Please show
your overall reasoning step by step of the categorization.
```

[8]This prompt employs a zero-shot classification approach: the model was not specifically trained to answer these questions and was not provided with any example responses. It relies entirely on its general knowledge and the instructions provided in the prompt.

[9]Generative AI models are stochastic by nature (Chann, 2023), meaning that the models can give you a different output from the same input prompt. Thus, I opted to produce the results 5 times and then get an agreement between the different responses.

[10]A rough English translation of *La Patrona* is *The Female Boss*. You can see the first episode of the telenovela on YouTube.

[11]The threshold $\delta = 75$ was chosen as the one to minimize a combination of MSE and maximize the precision measurement with the validation sample.

Figure 1: Female Empowerment Displayed in *La Patrona*

Notes: This is a frame extracted from the first 10 minutes of episode one of the telenovela *La Patrona* (2013) available on YouTube. It shows Gabriela Suárez, the main character, fighting back against her harassers with a small pickaxe in hand.

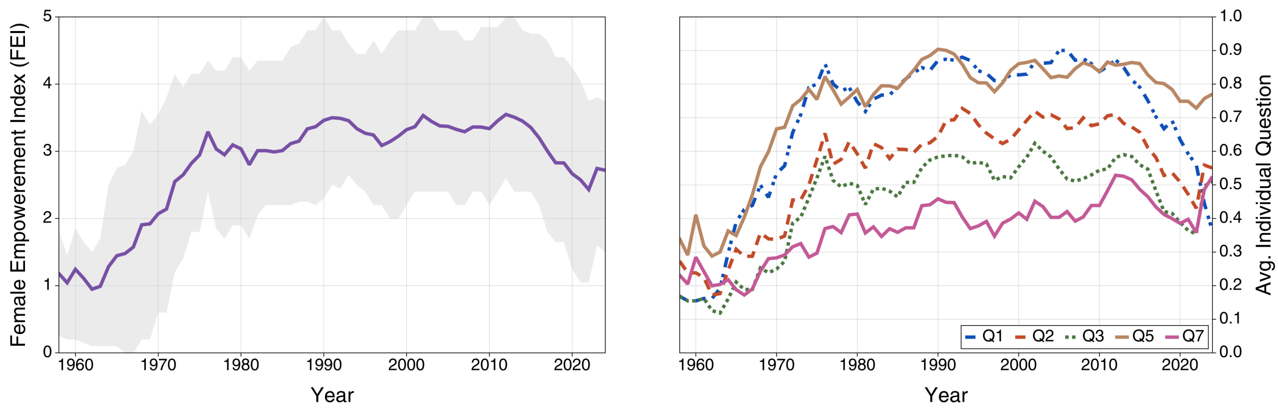
For example, for Q3 in *La Patrona*, the model predicts `true` with confidence scores of 80, 90, 85, 70, and 95 across five trials. Since four out of five scores exceed the $\delta = 75$ threshold, the final response for Q3 is set to `true`. In contrast, for Q3 in *Rebelde (2024)*, the model predicts `true` with confidence scores of 60, 80, and 55 in three trials, and `false` in the other two. Here, only one `true` prediction exceeds the threshold, while the rest are either below 75 or are `false`. After filtering the `true` answers with the threshold, only one `true` remains, four responses are `false`, so the most agreed answer is set to `false`.

The telenovela *La Patrona* scored an expected FEI of 5. In contrast, the telenovela *Rebelde (2024)* scores an FEI of 0. It narrates the story of a new group of aspiring musicians with scarce resources at the Elite Way School who must contend with a secret wealthy society while pursuing their musical dreams. The group is described to have only one female character mentioned in the synopsis (Q1 = `false`), does not seem to talk to other female characters (Q2 = Q3 = `false`), does not have a job (Q4 = `false`), and is not the protagonist (Q5 = `false`).

Figure 2 in Panel A shows the average FEI time trends for telenovelas aired in Latin America from the 1960s to 2024. Telenovelas show an initial increase in FEI from the 1960s to the 1980s, followed by a steady level until the 2010s. Following the 2010s, there is a marked decrease in the FEI in Latin America. Figure 2 shows the share of telenovelas that answered `true` to questions Q1 to Q5 over time in Panel B. The question with the most significant drop during these years was (Q1) are there more than 2 female characters mentioned in the synopsis? In 2010 the average answer to this question was around 90% of telenovelas with a notable drop to almost 40% in 2023. One potential explanation for this trend could be the rise of streaming platforms, which have become increasingly popular in the 2010s. In response, telenovelas may have shifted their focus to more traditional themes, such as romance and family, rather than empowerment. Another potential explanation for this trend could be the start of the #Metoo Movement in the online space, where women started to share their experiences of abuse and harassment online. The movement started in social media around 2006 on Myspace and then spread to Twitter and Facebook, with the hashtag #Metoo gaining momentum worldwide and sparking several protests in 2017. This might suggest that female empowerment in LATAM was a political topic, and producers of telenovelas might have dialed down the representation of women to avoid political associations or controversial topics.[12]

[12]This backlash reaction has some precedent in the literature. Gulesci, Lombardi and Ramos (2024) finds that exposure to LGBTQ+ characters in telenovelas decreased support for the LGBTQ+ community in Latin America. In addition, Gonzalez (2025) finds that telenovelas with social class conflict themes had a negative effect on support for redistribution.

Figure 2: FEI trends in LATAM



Panel A: Overall

Panel B: Individual questions

Notes: Panel A displays the smoothed 5-year moving average FEI score of all telenovelas aired each year, while the grey bands indicate yearly interquartile range for the LATAM sample. Panel B displays the smoothed average per question of the FEI.

2.1.2 Model Validation

To assess the model's ability to identify FEI questions, I compared its predictions to my own manual categorization of 240 telenovelas from Chile and Brazil.[13] I chose four commonly used measures in the Machine Learning literature: precision, recall, accuracy, and f-score (Jurafsky and Martin, 2024). Let tp be the number of true positives, where a true positive happens when both the model and my categorization predict the FEI question label. In addition, let tn, fp, fn, be the true negative, false positive, and false negative, respectively. As an illustration, a false positive (fp) is when the model predicts a `true` answer to a particular question, while I label that question as `false` for a single telenovela. The formal definition of each metric is presented below.

$$\text{accuracy} = \frac{tp + tn}{tp + tn + fp + fn} \qquad \text{precision} = \frac{tp}{tp + fp}$$

$$\text{recall} = \frac{tp}{tp + fn} \qquad \text{F1} = 2 \times \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

Accuracy measures the overall proportion of correct predictions, but it may be misleading if one class (FEI question or not) dominates. Precision evaluates the quality of positive predictions, indicating how many telenovelas classified as having an FEI question are correct, which is important to minimize false positives. Recall assesses the model's ability to identify all telenovelas with FEI questions, reflecting its sensitivity to true positives. The F1 score balances precision and recall, providing a single metric to evaluate the model's effectiveness, particularly when both false positives and false negatives matter or the classes are imbalanced. For this study, precision is the most relevant measure of model performance, as any false positive (misclassifying an FEI answer as `true` when it should be `false`) could bias the estimation of exposure to empowerment-themed content.

The model shows good performance across all questions. Table 2 shows the model performance per question, with an average precision of 0.80, accuracy of 0.77, recall of 0.81, and f-score of 0.80. The model performs particularly well in identifying if there are more than two female characters and if they talk to each other with precision higher than .95. However, the model's performance drops when identifying if the characters work or if they talk about something other than men. For comparison, Gonzalez (2025) shows that the best model to predict whether a telenovela contains

[13] There is an ongoing debate about the best benchmark for model validation. My approach uses my own contextual understanding of telenovelas and their summaries. An alternative is to ask a general audience to assess the classifications and use their consensus as a benchmark (Michalopoulos and Xue, 2021). Lastly, researchers can rely on expert evaluations to validate the model (Törnberg, 2023).

Table 2: Model Performance Validation

	acc.	preci.	recall	F1
Bechdel-Wallace Test				
more than 2 female chars? (Q1)	0.954	0.967	0.981	0.974
do they talk to each other? (Q2)	0.816	0.944	0.812	0.873
do they talk about something other than men? (Q3)	0.640	0.703	0.520	0.598
Working Female Character				
does any female char. work? (Q4)	0.695	0.679	0.980	0.802
does the female protagonist work? (Q5)	0.782	0.696	0.772	0.732
FEI sum				
	0.777	0.798	0.813	0.796

Notes: FEI sum refers to the macro-average performance of all questions for each metric.

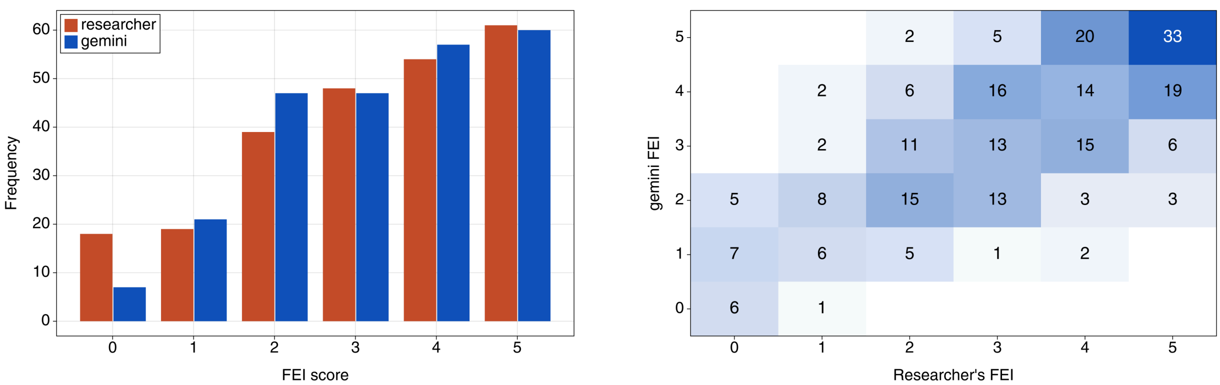
social class conflict or not has around a 0.75 accuracy, 0.82 precision, 0.48 recall, and 0.60 f-score in a similar exercise. Another comparison can be made with Michalopoulos and Rauh (2024), who report that the `gpt-3.5-turbo` model achieves an average accuracy of 0.67 when identifying characteristics (e.g., violent, submissive, intelligent, naive, etc.) of female and male characters in movies.

The correlation between the model's and my FEI scores is approximately 0.71, indicating a strong positive relationship between the model's predictions and my own. Panel A of Figure 3 shows that the distribution of FEI scores produced by the `gemini` model closely resembles that of the researcher's manual scores. Panel B presents the contingency matrix, where each cell indicates the number of telenovelas assigned a particular FEI score by both the model and me. Both pictures suggest that the model tends to overestimate FEI scores: it classifies more telenovelas with a score of 2 than I do, and fewer with a score of 0. However, this does not undermine the fact that the model generally aligns well with manual classification, with a slight positive bias.

2.1.3 Sources of Measurement Error

There are at least 3 sources of measurement error in the FEI. First, the FEI is not predicted perfectly. The model slightly overestimates the empowerment of female characters relative to my own manual categorization, potentially leading to upward bias in FEI assignment. This overestimation may stem from the model's tendency to provide more agreeable or socially desirable responses, a feature common to commercial generative AI models. While the `gemini` API does not record chat context,

Figure 3: Researcher's FEI vs `gemini` FEI



Panel A: Distribution Difference

Panel B: Contingency Matrix

Notes: The researcher's FEI score corresponds to the sum of the researcher's manual categorization of 239 telenovelas from Chile and Brazil, which were exposed to the same questions as the model.

reducing some risks of context contamination, the prompt was also designed without examples to ensure the model does not anchor its responses based on prior interactions. Table A.3 compares `gemini`'s performance to other commercial (e.g., `gpt-4o` or `claude-3-5`) and non-commercial (e.g., `deepseek-8b` or `phi-4`) models and finds similar patterns across all models.

Second, there are substantial differences in synopsis length across telenovelas. While the average synopsis is around 2,000 words, some are under 100 words, offering only a brief plot overview and potentially omitting nuanced depictions of female empowerment. These shorter synopses are likely to focus on early episodes and may miss character development, potentially resulting in a lower FEI assignment than warranted. Two synopses exceed 10,000 words, encompassing detailed multi-season descriptions. To partially tackle this problem I removed both extremely short and long synopses from the sample. Taken together, these sources of error do not point to a consistent direction of bias. The resulting combination of these errors is unlikely to be normally distributed, and thus attenuation bias is unlikely. Third, it is possible for a telenovela to receive a high FEI score (e.g., 5) without genuinely representing female empowerment, as the index does not capture the emotional or narrative context of empowerment. In Section 5.3 I tackle this specific concern and explore if the emotional attachment to characters influences the FEI effects.

3 Telenovelas, Gender Values, and Labor Force Participation

In this section, I present a framework to explain how the Female Empowerment Index (FEI) may influence labor market participation by shifting gender norms. I use census data from the Integrated Public Use Microdata Series (IPUMS) to document the relation between FEI exposure during adolescence and labor market outcomes in my Latin American sample. I then turn to the World Value Survey (WVS), which measures gender attitudes across Latin American countries, to show the link between FEI exposure and gender norms.

Conceptual Framework I focus on the cumulative FEI exposure during the teenage years (ages 13–18). The psychology and media literature suggest that long-term, cumulative exposure to media, rather than single instances, shapes individuals' perceptions and behaviors. Cultivation Theory argues that repeated exposure to particular media content can significantly shape an individual's worldview, complementing the influence of their personal experiences (Gerbner and Gross, 1976; Gerbner *et al.*, 2002). Thus, I focus on the sum of the FEI across telenovelas as a measure of this cumulative exposure to empowering narratives and characters depicted in telenovelas. A higher FEI sum indicates a media environment where empowered female characters are consistently visible, helping to normalize female agency. This repeated exposure can shape viewers' beliefs about what is possible and desirable for women in society. In addition to the frequency of exposure, the timing of exposure is also important. Research indicates that gender norms typically form between ages 13 and 18, with relatively little change occurring in later years (Ellemers, 2018). Exposure to empowering female characters during this formative period can have lasting effects, shaping individuals' career aspirations and choices well into adulthood when labor market decisions are made. Therefore, my exposure variable is TeenFEI, which is defined as the cumulative FEI exposure from all telenovelas aired during an individual's teenage years (ages 13–18) in their country.

TeenFEI and labor market participation To document the relation between TeenFEI and labor market outcomes at the country level, I start by making a simplifying assumption: all individuals within a country are potentially exposed to the same telenovelas. This assumption is quite strong, since some telenovelas may not be broadcast in some regions of the country.[14] However, it allows for a comparison of women across different countries and survey years. The results should therefore be

[14]Due to the limitations of broadcasting towers and the legal framework governing transmission rights, not all regions within a country are exposed to every telenovela. In fact, I exploit this geographic variation in Mexico to obtain causal estimates.

interpreted with caution and are intended to facilitate cross-country comparisons rather than provide causal estimates. At the end of this section, I will turn my attention to Mexico, where I can exploit an instrumental variable design to deal with potential confounding factors.

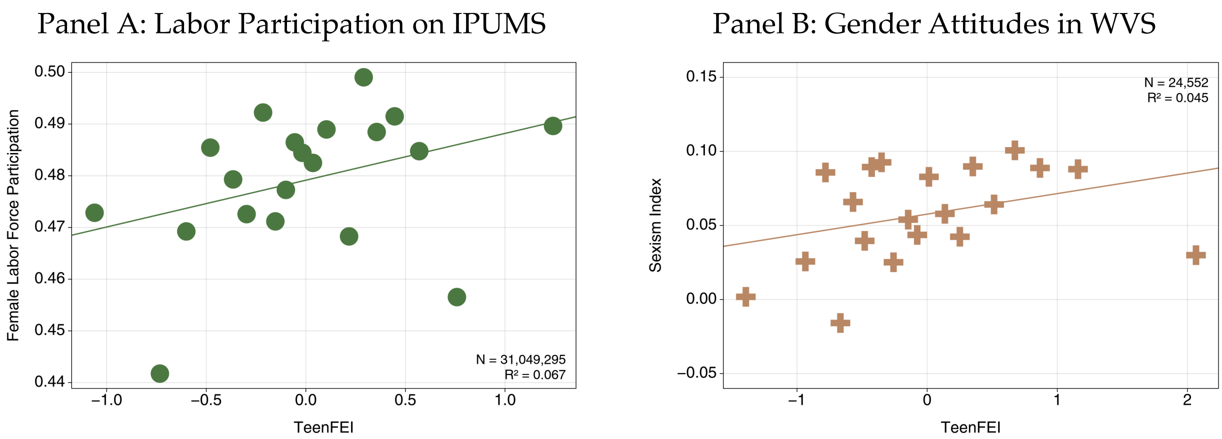
For labor participation data, I turn to census data downloaded from IPUMS for Mexico, Brazil, and Chile with individual-level data on labor participation.[15] These three countries have the highest GDP in the region and show a good cultural mix, with two Spanish-speaking countries at opposite ends of the region and a Portuguese-speaking country. In addition, these three countries have the highest quality telenovela data extracted from online sources. For each country, I use all available census waves and select all women in the sample. For each individual, I compute their TeenFEI as the total FEI exposure from all telenovelas aired during their teenage years (ages 13–18) in their country. I estimate the following model:

$$\mathrm{Y}_{i,c,w} = \alpha + \beta \mathrm{TeenFEI}_{i,c,w} + \delta_{c,w} + \delta_b + \varepsilon_{i,c,w}$$

where $\mathrm{Y}_{i,c,w}$ is an outcome variable for a working-age woman i living in country c surveyed in wave w. $\mathrm{TeenFEI}_{i,c,w}$ is a variable of the FEI sum of telenovelas that individual i was exposed to during their teenage years (ages 13–18) in country c and wave w. To make the interpretation of the coefficients easier, I calculate the TeenFEI z-scores, where a standard deviation increase in TeenFEI is equivalent to FEI exposure of 95. This shock is quite substantial since an average individual in this sample is exposed to an FEI of 184 over their teenage years, making it equivalent to around a 52% increase in FEI exposure. The shock is also equivalent to moving from the 25th percentile of the TeenFEI distribution to the 55th percentile. $\delta_{c,w}$ captures country×wave fixed effects, which would account for any unobserved factors that are constant within countries over time. I will include a birth decade cohort fixed effect δ_b to account for any cohort-specific trends in labor market participation and gender norms.

There is a positive correlation between TeenFEI and female labor market participation. Panel A of Figure 4 shows a positive relationship between the z-score of TeenFEI and female labor market participation. The estimated β is 0.01 (s.e. 0.003), indicating that a one standard deviation increase in TeenFEI is associated with a 1 percentage point increase in the probability of labor market partic-

Figure 4: Aggregate Relation of TeenFEI in LATAM



Notes: Both Panels show the binscatter plot of the individuals' outcomes using country × wave fixed effects and cohort fixed effects. TeenFEI is defined as the cumulative exposure to the FEI during the teenage years (13-18 years old) aired in their country for each individual. Panel A uses census data from IPUMS to measure labor force participation. The Sexism Index on Panel B is the average aggregation of WVS questions to measure gender attitudes. These questions include (i) should women be prioritized for jobs? (ii) Do women make better leaders than men? (iii) Are women better at running the economy? (iv) Is it important for women to have a university education? (v) Should women work outside of the house?

[15]The available censuses from IPUMS International are: Chile (1970, 1982, 1992, 2002, 2017), Mexico (1970, 1990, 2000, 2010, 2020), and Brazil (1970, 1980, 1991, 2000, 2010).

ipation. However, the effect size is modest, as a 50% increase in female empowered content during teenage years only translates to an average 2% increase in labor force participation. These results suggest that exposure to empowered female characters during the formative teenage years can encourage female labor market participation, even under the strong assumption that all individuals in a country are exposed to the same content.

TeenFEI and attitudes towards women Having shown the association between TeenFEI and female labor force participation, I next study whether this exposure is associated with changes in gender attitudes. Specifically, I use data from the World Values Survey (WVS), which covers countries worldwide and measures attitudes toward government, society, minorities, and related topics. As a proxy for individuals' attitudes toward working women, I use a sexism index available in the survey that takes the average response to the following questions: (i) Should women be prioritized for jobs? (ii) Do women make better leaders than men? (iii) Are women better at running the economy? (iv) Is it important for women to have a university education? (v) Should women work outside of the house? A higher value of the index indicates less sexist attitudes or increased approval for women in positions of power.

Estimating the model using the sexism index as the outcome, I find a strong positive relation between TeenFEI and this sexism. Panel B in Figure 4 shows the individual binscatter plot of TeenFEI on the sexism index. My estimates suggest that a one standard deviation increase in TeenFEI is associated with a 0.01 increase in the sexism index (s.e. 0.06), which translates to a substantial 10% increase in the index. Combined with the previous result, it presents suggestive evidence that exposure to telenovelas has a positive impact on both attitudes towards women and a potential passthrough to labor market participation.

Limitations There are two important caveats of this analysis. First, I do not observe what telenovelas each individual has watched. Therefore, the coefficient β captures an "intent-to-treat" effect. The coefficient of interest measures the impact of having access to telenovela content during adolescence on labor force participation, regardless of whether individuals actually watched the programs. It captures the direct effect of being exposed by watching or indirectly exposed through discussions within their location, network, or family at some point during their teenage years. Discussion of telenovelas or community groups watching telenovelas is a common phenomenon in Latin America, and thus this indirect channel is plausible (Lopez, 2002; Antezana *et al.*, 2023). Exposure may also occur indirectly, as peers who watch telenovelas can influence classmates and friends through changes in attitudes and behaviors, even if one does not personally watch the shows

Second, these results cannot be interpreted as the causal evidence, since the analysis suffer from both reverse causality and omitted variable bias. Even with country$\times$wave fixed effects, there is no clear way to disentangle the direction of the relationship between TeenFEI and the outcome. For example, if there is a positive relation between TeenFEI and labor force participation, it could suggest that either telenovela producers are more likely to create content that resonates with women who are already inclined to support gender equality and work, or that exposure to such content actively shapes viewers' attitudes. In addition, the omitted variable bias arises because the individuals' gender norms are unobserved. To address these issues, I will focus on Mexico, where I can use variation in physical signal loss from broadcasting towers as an instrumental variable for TeenFEI exposure. This strategy is discussed in detail in the following Section.

4 Instrumental Variable Strategy

I exploit geographic and temporal variation in television signal loss as an instrument for exposure to telenovelas in Mexico.[16] Television has been historically broadcast over free-to-access radio waves that propagate through the air. These radio waves weaken as they travel farther from the broadcasting tower, and the loss increases if obstacles, such as buildings or mountains, lie between the tower and the receiver. Individuals located at the same distance from a broadcasting tower may experience different signal strengths because of terrain and structural obstructions. This variation, driven by physical geography, generates plausibly exogenous differences in exposure to telenovelas.

Mexico provides a particularly advantageous setting for this analysis due to the availability of public records of government concessions to commercial channels, dating back to the 1940s. These historical archives contain detailed information about broadcasting tower locations and characteristics, making it possible to trace which television channels transmitted through specific broadcasting stations over time. Crucially, this unique historical data allows for the construction of a predicted signal loss that is plausibly exogenous to local labor market conditions.

4.1 TV Broadcasting in Mexico

The regulatory framework for Mexican television was established with the Ley Federal de Radio y Televisión (Federal Law of Radio and Television), enacted in January 1960. This foundational legislation declared radio and television as means of public interest and created a dual system of licenses: concessions and permits. Commercial stations, which were authorized to broadcast advertising, operated under concessions. In this context, a concession is a legal authorization granted by the government to a private entity, allowing it to broadcast content to the public at a certain frequency (i.e., TV channel) in a particular set of locations. In contrast, permits were granted to non-profit entities, explicitly prohibiting them from commercial use of frequencies. This system lets concessionaires jointly manage and operate frequencies, which, despite separate legal concessions, effectively creates a monopoly. For private broadcasters, this framework often provided advantages such as facilities for expansion, tax benefits, and discretionary allocation of concessions.

Specifically, commercial concessions for open television channels are granted exclusively to individuals of Mexican nationality and are typically awarded through public tenders where interested parties must meet various requirements, including an advance payment to the Federal Institute of Telecommunications (IFT). These concessions are valid for a period of up to 20 years and can be renewed for equal durations, with rare cases of closure or revocation. For instance, a concession granted to a Televisa subsidiary in 1994 for 62 channels had a validity of 16 years, with terms subject to review every five years. The three main broadcasting groups are Televisa, TV Azteca, and Grupo Imagen, which are financed by selling advertising time and accumulating profits while broadcasting to the general public.

Today, the Mexican open broadcast television industry is highly concentrated, with Televisa and Televisión Azteca dominating, collectively holding an estimated 95% of television concessions and commanding approximately 90% of the audience in Mexico City (Pareja Sánchez, 2010). Televisa's main channels include Las Estrellas, Canal 5, and NU9VE. Grupo Azteca operates Azteca Uno, Azteca 7, ADN 40, and a+. Grupo Imagen, with Imagen Televisión, has emerged as a third, albeit smaller, player. This concentration is underpinned by strong, historical political ties, exemplified by President Miguel Alemán's early involvement and Televisa's explicit loyalty to the ruling party, often providing favorable coverage in exchange for expansion facilities and tax advantages. In addition, television is massive, with over 90% of Mexicans owning a television set and regularly watching it since the 2000s (INEGI, 2022).

[16] Several studies have used predicted signal strength to study outcomes such as education (Kearney and Levine, 2019), violent conflict (Yanagizawa-Drott, 2014), and social capital (Olken, 2009).

4.1.1 Broadcasting Data

To construct the historical dataset of broadcasting stations in Mexico, I collected data from three sources. The Mexican National Institute of Statistics (INE) provides data on the coverage area of each television station for 2020–2025.[17] These data include basic details for each station, such as the geographic coordinates (latitude and longitude), a polygon representing the coverage area, and information on the coverage provided by auxiliary signal repeaters.

The Public Concession Registry (RPC) from the Federal Institute of Telecommunications (IFT) keeps public documents about concession rights for each station dating back to the 1940s. These documents state the rights owner, the frequency on which they can transmit (or channel number), the characteristics of the station at the time of the concession (e.g., power, antenna height, etc.), which populations they are targeting, and in some cases the degrees to which they are allowed to broadcast. The report *La televisión de la Nueva Generación* contains a table of the concessions historical records from 2004-2019. I use this table as a starting point and expand it with the public legal records to expand the coverage of the table to 1950-2025. The last step of the process is to match owners of stations/concessions with their channels. This information is contained in the Wikipedia list that lists every station and which channel names broadcast there. Then, I corroborated with a random selection of stations that the matches were correct by checking the owner groups with the concession records, channel numbers, station, and location of each station. Lastly, I aggregate the broadcasting data at the municipalidad level in Mexico, which is similar to a county in the USA. These counties are typically defined as high-density population centers and range from small urban centers to large rural regions, with significant variation in their areas.[18]

Figure 5 shows the expansion of television coverage in Mexico from 1960 to 1990. Panel A depicts the limited broadcast area in 1960, when the first major stations were constructed near the east of Mexico City and began expanding to the coast. Panel D shows the significantly broader coverage achieved by 1990. The expansion of television coverage in Mexico was driven primarily by the construction of new stations and the installation of auxiliary transmitters, rather than by upgrades to existing infrastructure. The first commercial television channels outside the east coast appeared in 1963, marked by the launch of Canal 5 and the establishment of new broadcasting stations in other regions. In preparation for hosting the 1968 Olympic Games and the FIFA World Cup, Grupo Televisa undertook a major expansion of its broadcasting network to support both domestic and international transmission.

4.2 Labor Data

The labor market data comes from Mexico's Central Bank EconLab's Local Labor Markets micro data (Aldeco *et al.*, 2024). These data combine both the National Survey of Occupation and Employment (ENOE)[19] and the Census surveys in 1990, 2000, 2010, 2015, and 2020 to generate a repeated cross-section.[20] It contains individual, household, and aggregated variables at the Local Labor Markets (LocalLM) level, as a group of counties that share labor market characteristics. Unlike the Metropolitan Statistical Areas (MSAs) defined by the US Census Bureau, these LocalLMs cover the entire geography of Mexico. An average LocalLM is a group of around 3 counties, and some of the larger LocalLM may encompass more than 10 counties. The individual data has detailed information on

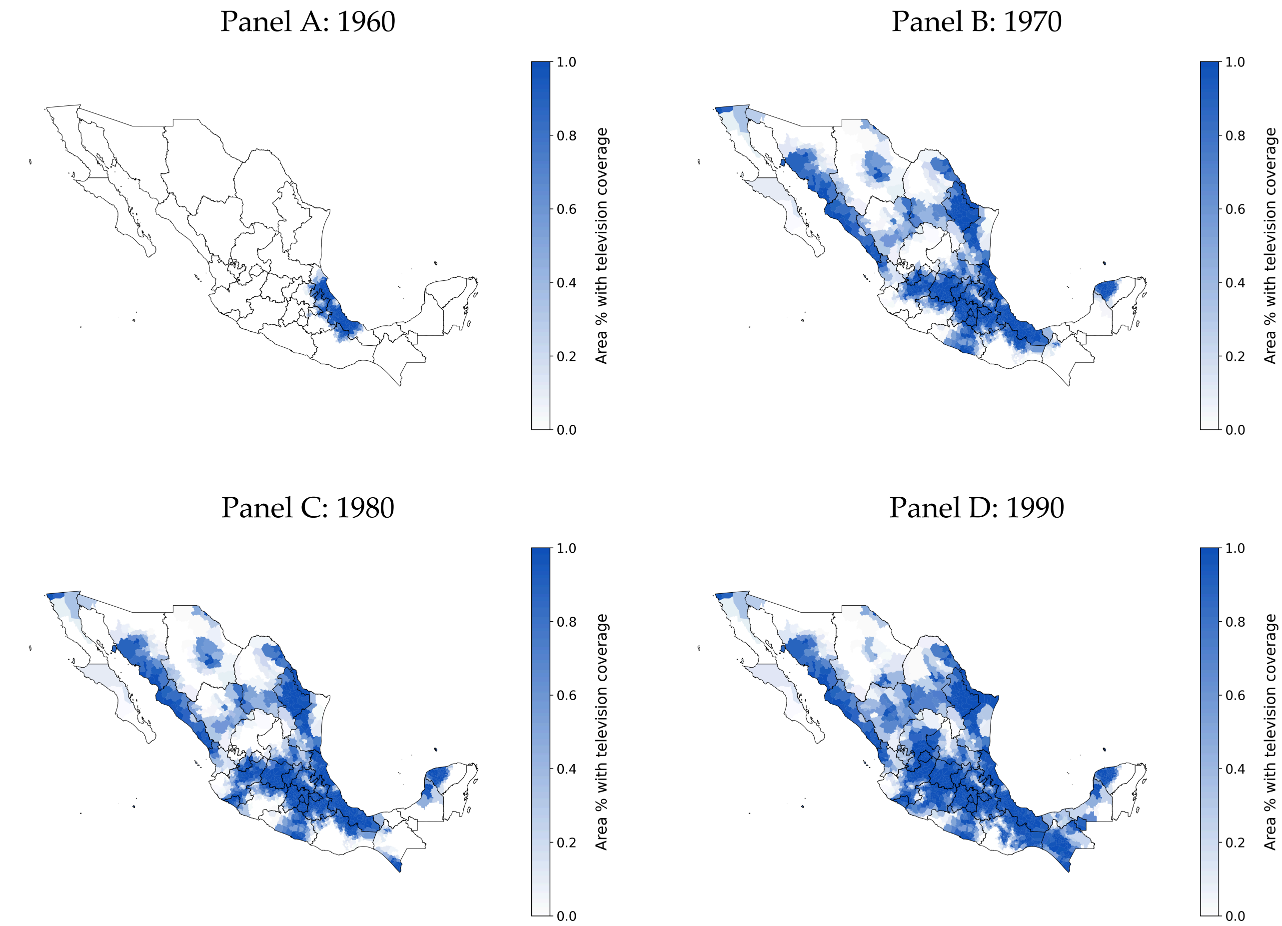
[17] The INE started this initiative around 2015, where versions of the same website are available on Internet Archive's Wayback Machine, but without access to the shapefiles.

[18] An average county in Baja California is around 23 times the size of an average county in Mexico City. At the same time, the state of Baja California is only divided into three counties with a larger total area, while Mexico City is divided into 6 counties.

[19] The ENOE labor data underrepresents Mexico's informal sector, where many women work (e.g., domestic services). This likely leads to an underestimation of the true female labor force participation rate.

[20] While some individuals may appear in multiple census years, the identification variable is only unique within each survey, so individuals cannot be tracked across waves.

Figure 5: Television Coverage Overtime



Notes: These maps show the percentage of the area covered by television broadcasting over time in Mexico. The boundaries of the states are shown in black lines.

location, demographics, education, labor market participation, occupation, industry, and commuting information. I use the 2010, 2015, and 2020 surveys to capture the current labor market decisions of Mexican women who were exposed to telenovelas during their teenage years. Having access to the most recent surveys allows me to capitalize on most of the cohort variation in exposure to telenovelas.

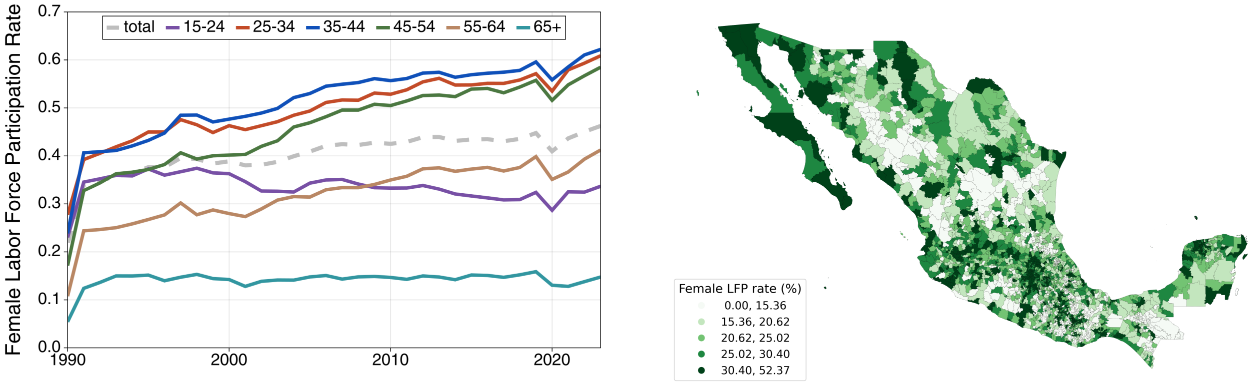
Panel A in Figure 6 displays the evolution of female labor force participation rates in Mexico over time, broken down by age group. Mexico's female labor force participation rate has gradually increased since the 1990s, stabilizing at around 41% from the 2000s onward. However, female labor force participation has seen a significant rise across all working age groups (25-64 years), with the younger cohorts reaching even close to 60% participation in 2023. In contrast, age groups 15-24 and 65+ have even experienced a decrease since their 1995 levels. Despite Mexico's significant progress in female labor force participation, it is still well below the average OECD country, with a female labor force participation rate of 41% in 2020 according to the ILO estimates. Panel B shows the spatial variation of female labor force participation in 2020 across counties in Mexico. It shows that the major concentrations of working women are near populated cities (e.g., Mexico City, Guadalajara, and Monterrey) and some parts of the north near the border with the USA in Baja California.

4.3 Identification

To test the effect of the FEI exposure during teenage years on women's labor force participation, one might run the following naive regression:

$$\mathrm{Y}_{i,c,w} = \alpha + \beta \mathrm{TeenFEI}_{i,c,w} + \varepsilon_{i,c,w}$$

Figure 6: Mexico's Labor Market



Panel A: Time series

Panel B: Across counties in 2020

Notes: Panel A shows the labor market participation rate in Mexico using ILOStats estimates for 1960-2024 for both men and women. Panel B shows the female labor market participation across counties in 2020 using the Bank of Mexico EconLab dataset for 2020.

$\mathrm{Y}_{i,c,w}$ is a labor market outcome of interest for individual i in county c surveyed on census wave w. $\mathrm{TeenFEI}_{i,c,w}$ measures the FEI of telenovelas that individual i was exposed to during their teenage years (ages 13–18). Where an individual i is exposed to a telenovela released in year t if she was between 13 and 18 years old in that year, and the TV channel covers more than 30% of the county given the broadcasting data. For example, consider a woman born in 1990. To calculate her TeenFEI, I first determine her teenage years, would be 2003–2008 for someone born in 1990. Next, I identify all TV channels that broadcast telenovelas and had coverage in her county during those years, using the broadcasting data. Suppose that only two channels reached her county between 2003 and 2008, and each channel aired a different set of telenovelas. For each year from 2003 to 2008, I list all telenovelas aired by these two channels that she could have been exposed to. Each telenovela has an associated FEI score, which reflects its level of female empowerment content. I then sum the FEI scores of all telenovelas broadcast by these two channels during her teenage years. This total represents her TeenFEI value, which captures potential exposure to empowerment content in telenovelas during adolescence, given the programming available in her county.

To improve the interpretation of the coefficients, I standardize the TeenFEI variable to its corresponding z-score. A standard deviation increase in TeenFEI is equivalent to an FEI of 116 during their teenage years. This is a substantial change since an average individual in this sample is exposed to and FEI of 203, making it equivalent to a 57% increase in FEI exposure. TeenFEI is designed to vary both between and within counties. The variation between counties is straightforward: an individual residing in an area served by multiple telenovela channels is likely exposed to more content than someone in an area with only one. However, the measure is more nuanced than a simple channel count, as it also accounts for the specific shows being broadcast. Since each channel has a unique programming slate, two counties with an equal number of channels can still have different TeenFEI values based on the distinct set of telenovelas available. This measure also captures changes over time within the same county, reflecting not only the evolving availability of TV channels but also the turnover of telenovelas as older shows end and new ones are introduced, exposing individuals from different cohorts to distinct media environments during their formative teenage years.

Similar to Section 3 specification, the coeficient β captures an intent-to-treat effect because I do not observe what telenovelas each individual watched in their teenage years.[21] The coefficient of

[21]Most media effects studies estimate an intent-to-treat (ITT) effect, as in La Ferrara, Chong and Duryea (2012) and Yanagizawa-Drott (2014), which use the introduction of channels or radio shows as proxies for exposure. In contrast, Schneider-Strawczynski and Valette (2025) estimates an average treatment effect by using data on individual TV preferences, ensuring that the analysis captures the specific shows each person was directly exposed to.

Table 3: FEI Patterns

Panel A: FEI Patterns Across Birth Cohorts

birth cohort	avg. age	pop. (%)	working age	labor part.	teenage decade	TV coverage	telenovela coverage	avg. FEI	std FEI
1930	84.85	0.01		0.03	1940	0.16	0.14	–	–
1940	75.21	0.03		0.10	1950	0.41	0.34	9.14	18.09
1950	65.30	0.07		0.23	1960	0.72	0.61	58.71	56.27
1960	55.41	0.10	✓	0.42	1970	0.75	0.64	96.44	82.74
1970	45.51	0.13	✓	0.54	1980	0.96	0.89	166.55	85.47
1980	35.57	0.15	✓	0.56	1990	0.99	0.98	294.59	70.46
1990	25.55	0.16	✓	0.49	2000	0.99	0.98	300.43	68.26
2000	15.42	0.17		0.12	2010	0.99	0.98	203.50	78.11
2010	5.59	0.16		0.00	2020	0.99	0.98	65.13	13.52

Panel B: FEI Patterns Across Age Groups and Waves

	25-60		35		45		55	
	labor part.	avg. FEI	labor part.	avg. FEI	labor part.	avg. FEI	labor part.	avg. FEI
2010 wave	0.46	159.84	0.48	172.68	0.49	101.44	0.35	59.69
2015 wave	0.45	191.69	0.48	278.36	0.48	106.08	0.35	87.71
2020 wave	0.52	219.10	0.57	292.40	0.55	172.79	0.43	103.30

Notes: The sample in Panel A contains all women in the 2020 wave. The working age column indicates whether most women in that cohort are of working age. The teenage decade corresponds to the decade when most women born in that cohort would have been teenagers. TV coverage corresponds to the average of people whose county has TV coverage in the decade they were teenagers. Telenovela coverage refers to the percentage of people who received a signal from a channel that broadcast telenovelas.

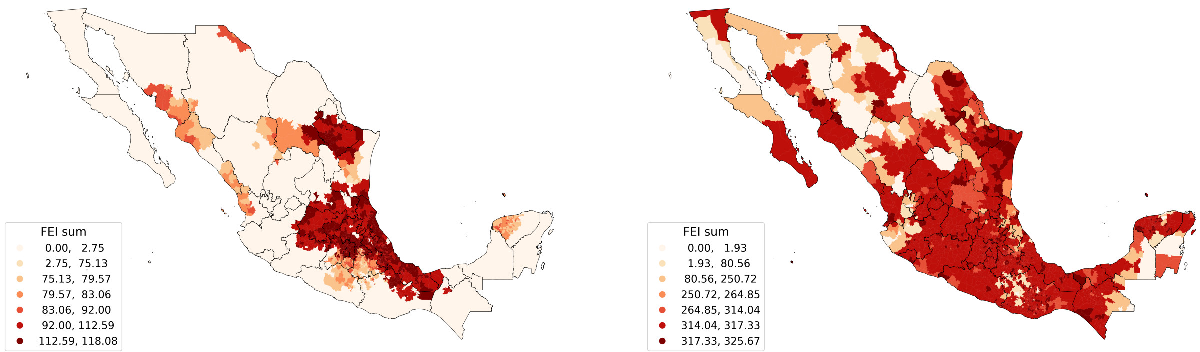
interest measures the impact of having access to telenovela content during adolescence on labor force participation, regardless of whether individuals actually watched the programs. It captures the direct effect by watching empowered women on telenovelas or indirectly through discussions within their location, network, or family at some point during their teenage years.[22]

Table 3 illustrates how does TeenFEI changes across cohorts and census waves. Panel A shows descriptive statistics for each woman's birth cohort in the 2020 wave. TV and telenovela penetration increases substantially over time. Women born in the 1930s had limited access to television during their teenage years, with an average of 16% of individuals living with TV coverage. In contrast, women born after 1970 had complete access to any TV channel or a telenovela channel by the time they were teenagers. Younger cohorts within the working-age population have experienced the highest levels of TeenFEI, as an average woman in the 1960 cohort is exposed to an FEI of 96, while an average woman from the 1990 cohort has a TeenFEI of 300. Labor market participation peaks for the 1970 and 1980 cohorts coincide with these cohorts also experiencing the highest levels of TeenFEI exposure. There is also substantial variation in TeenFEI within cohorts, with the standard deviation ranging from 30% to 90% of the cohort average. Panel B shows the average labor force participation rate and the average TeenFEI for different ages across multiple census waves. In general, more recent waves tend to have higher levels of TeenFEI, reflecting the overall increase in telenovela content over time. There is suggestive evidence of a positive correlation between TeenFEI and labor force participation, particularly among individuals aged 35 to 45.

[22]Discussion of telenovelas or community groups watching telenovelas is a common phenomenon in Latin America, and thus this indirect channel is plausible (Lopez, 2002; Antezana *et al.*, 2023).

Figure 7: FEI Spatial Distribution

Panel A: 1960-1969 Panel B: 1980-1989



Notes: Each map corresponds to the average FEI sum scores for a person living in a particular county in the decade 1960-1969 or 1980-1989, who would have been exposed to.

There is significant TeenFEI variation across counties and time periods. Figure 7 show the spatial distribution changes of the FEI across Mexico from 1960s to 1980s. Focusing on Panel A, the FEI is higher along the east coast of Mexico, particularly near the capital, Mexico City. There are significant differences in the FEI across regions, mainly due to the presence of different broadcasting channels. For example, Mexico City received signals from both Televisa and TV Azteca, while regions on the west coast received signals from only one of these networks. In Panel B, a similar pattern emerges, but with a broader coverage of the FEI across counties and generally higher FEI values. This is due to two main reasons. First, the expansion of broadcasting channels with government support in 1985 extended telenovela coverage to the west coast, parts of the north, and the Yucatan peninsula. Second, during this period, telenovelas began to feature more empowered female characters compared to those in the 1960s.

However this naive approach has several potential issues.

Internal migration First, internal migration of individuals could lead to selection bias in my analysis. I do not observe if individuals have remained in their locality since birth, which could bias the estimates if individuals who moved to a different locality had different labor market outcomes than those who stayed. Thus, the naive regression might be capturing individuals who moved from a low-coverage area to a high-coverage area, potentially inflating the estimated effects of telenovela exposure. I can partially address this concern by examining available migration indicators. Specifically, I test whether individuals currently live in the same state where they were born, and whether they have changed municipalities in the last five years. Table A.1 presents evidence from Mexico's population census. Approximately 82% of Mexicans remain in the state where they were born across all survey waves, and 83% have not moved to another county in the last five years. These patterns suggest that, although some internal migration occurs, most individuals remain in the same general location, reducing concerns about selection bias.[23]

Spatial and Cohort Differences Second, counties in Mexico exhibit substantial heterogeneity in their economic development, geography, cultural norms, and size. For instance, counties in Mexico City are small, urban, and offer abundant economic opportunities, whereas counties in Baja California tend to be much larger, more rural, and may adhere to more traditional social values. These underlying differences can influence both historical TV access and labor market outcomes, potentially confounding the analysis. To address these regional disparities, I include Local Labor Market (LocalLM) fixed effects in the regression. LocalLMs are geographic units defined by Aldeco

[23]Table A.2 shows that the results are robust (and even stronger) to restricting the sample to Mexicans, who have stayed in the same state they were born and have not changed municipalities in the last 5 years.

et al. (2024) that group together counties with similar economic characteristics, such as commuting patterns and job opportunities. As a result, counties within the same LocalLM share comparable labor market conditions, while different LocalLMs capture broader regional variation.

Beyond geography, an individual's cohort can strongly influence their career outcomes within any given survey wave. For instance, at a single point in time, older individuals might face different challenges or opportunities in the labor market compared to younger individuals, simply due to life-cycle factors. At the same time, different birth cohorts also experienced varying levels of telenovela exposure as television expanded across Mexico. In fact, Table 3 shows that different birth cohorts experienced varying levels of telenovela coverage, which could influence their labor market outcomes. To address this, I will include decade-of-birth cohort fixed effects, which isolate the impact of telenovela exposure from both life-cycle effects and other broad trends that change from one cohort to the next.

Endogeneity The third concern is the potential reverse causality and omitted variable bias of TeenFEI. Even controlling for LocalLM and cohort differences, TV channels may have chosen to broadcast in locations with particular gender norms, socioeconomic characteristics, or other unobserved factors correlated with labor market outcomes, leading to reverse causality. In addition, omitted variable bias may arise from unobserved individual gender norms. While these risks are mitigated with the focus on a historical exposure, I address them by employing an instrumental variable approach. Specifically, I use spatial and cohort variation in television signal loss as an instrument for TeenFEI exposure.

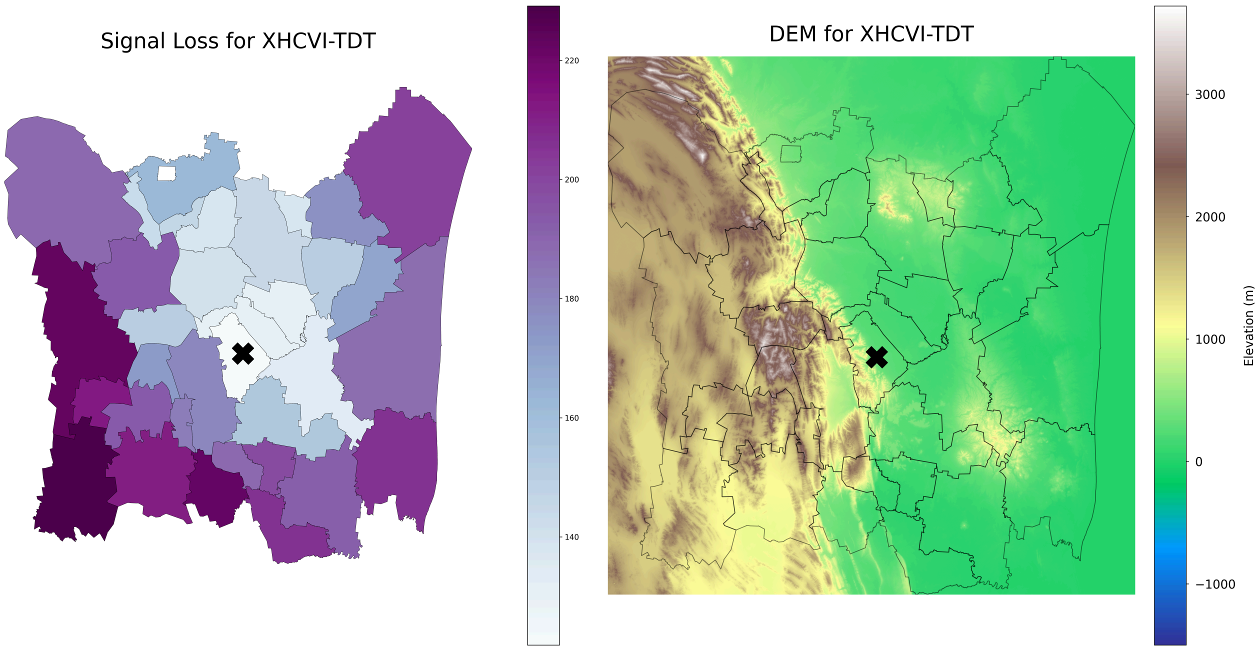
The Irregular Terrain Model (ITM) estimates the signal loss of a radio wave as it travels through the terrain, taking into account the frequency, power of the antenna, the distance between the transmitter and receiver, and the terrain profile between them (Oughton *et al.*, 2020). It's based on the physical principle that radio waves lose potency when they are obstructed or reflected by physical obstacles. The best possible signal is when the receiver has a clear line-of-sight to the transmitter. Therefore, the broadcasting signal should only depend on how many obstacles (e.i., topography) are in the way between the receiver and transmitter.

To estimate signal loss at the county level, I sample 30 random points within each county and identify all broadcasting towers within a 100km radius of each point. Then, I calculate the average propagation signal loss in decibels (dB) with reliability 99 and confidence 90 for that specific station using the Irregular Terrain Model (Oughton *et al.*, 2020).[24] Figure 8 illustrates the signal loss mapping for the `XHCVI-TDT` station. Panel A shows the signal loss in decibels (dB) across counties near the station, with darker areas indicating higher signal loss. Panel B shows the elevation map near that station. Places to the left of the station have higher loss due to the mountains, while places to the right have lower loss due to the flat terrain. Since counties may receive broadcasts from multiple stations in a given year, I then take the minimum predicted signal loss. This ensures that each county is assigned the best possible signal quality it could receive at any point in time, reflecting the optimal broadcasting conditions available to residents. However, it will restrict the spatial variation of signal structure since it only captures the best-case scenario for each county.

The instrument $\text{SignalLoss}_{i,w}$ represents the minimum signal loss in decibels (dB) that an individual i surveyed in wave w received in their county c during their teenage years, and varies across counties and cohorts. The minimum is used to capture the best possible signal reception for that individual. The introduction of a closer or less obstructed broadcasting tower over cohorts can lead to significant changes in signal loss for a given county. While it does not directly measure the content of telenovelas, the instrument strongly predicts the quantity of telenovelas. Figure 9 illustrates the

[24]The confidence level is the probability that the predicted signal strength will be met at a given location over time, while the reliability level is the probability that the prediction holds across different locations or scenarios.

Figure 8: Propagation Loss Map for XHCVI-TDT

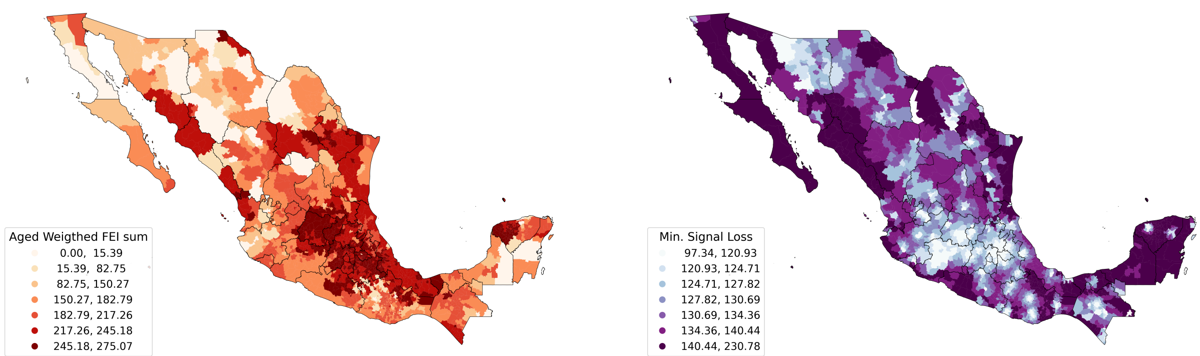


Notes: The left panel shows the coverage polygon of station XHCVI and the estimated propagation signal loss in the nearby counties. Darker colored polygons show higher losses of signal, while white shows small signal loss from the XHCVI stations. The right panel superimposes the Digital Elevation Model (DEM) of the area around the station XHCVI.

spatial relationship between minimum signal loss and age-weighted FEI across Mexican counties, providing suggestive evidence of a strong first-stage association between signal quality and telenovela exposure.[25] There is a clear pattern where places with a high minimum signal loss tend to have a lower FEI sum, like in Baja California, while places with a low minimum signal loss tend to have a higher FEI sum, like in Mexico City.

Several studies have used predicted signal strength as an instrument to study outcomes such as education (Kearney and Levine, 2019), violent conflict (Yanagizawa-Drott, 2014), and social capital (Olken, 2009). These papers rely on the exclusion restriction that geographic terrain, affecting signal loss between transmitters and receivers, is essentially random and thus uncorrelated with other determinants of the outcome. To test this assumption, most studies examine whether signal strength

Figure 9: FEI and Min. Signal Loss



Panel A: Age Weighted FEI

Panel B: Min. Propagation Signal Loss

Notes: Panel A displays the average Female Empowerment Index (FEI) weighted average for each county, based on age composition. Panel B displays the minimum propagation signal loss (in dB) in the county since 1960.

[25] The age-weighted FEI summarizes the spatial variation in treatment exposure, while also accounting for differences in cohort composition within each county.

is correlated with other outcomes after controlling for geographic and propagation characteristics. For example, Yanagizawa-Drott (2014) shows that residual signal loss is uncorrelated with other determinants of conflict after controlling for latitude, longitude, polynomials of mean and standard deviation of elevation, and distance to the nearest broadcasting tower. Following this literature, I include similar propagation controls in my regressions.

Unlike the existing literature, this paper exploits both the spatial and cohort variation in predicted signal loss. This introduces a new challenge since the construction of new broadcasting antennas could be influenced by local development trends, potentially biasing the instrument. Take, for instance, improvements in electricity in rural areas during the individual's teenage years. These local investments in infrastructure can coincide with investments in new broadcasting antennas near that region. In turn, these new development trends could persist and influence the labor market outcomes. Thus, the instrument will not be capturing television exposure but overall development in the region. To account for these development differences, I interact the spatial LocalLM fixed effect with the waves available in ECONLAB's. These should account for any LocalLM specific developmental trends, absorbing any trend differences in infrastructure investment across regions.

To further ensure the instrument does not capture development trends, it should not be correlated with changes in development. In particular, I check whether SignalLoss is correlated with changes in housing infrastructure during individuals' teenage years and subsequent years. First, I aggregate the SignalLoss to the county, cohort, wave level, resulting in $\overline{\text{SignalLoss}}_{c,t}$, which represents the average signal loss in county c during year t when individuals were teenagers. The logic is as follows: if the instrument is capturing development trends, then places with lower signal loss should also experience higher growth in infrastructure, using them as a proxy for this development. In particular, I estimate the following regression:

$$H_{c,t+x} = \alpha + \beta\overline{\text{SignalLoss}}_{c,t} + \Gamma X_{c,t} + \gamma H_{c,t-10} + \delta_{l,w} + \delta_b + \varepsilon_{c,t}$$

Where $H_{c,t+x}$ is the percentage of houses with access to a basic service in county c at year $t + x$, with x being after year t corresponding to individuals' teenage exposure decade. More explicitly, if $t = 1960$ this will correspond to the average access to housing infrastructure for individuals born 1950s cohort ($c = 1950$), who were teenagers in the 1960s. $\overline{\text{SignalLoss}}_{c,t}$ will then correspond to the average signal loss experienced by individuals in county c during their teenage years t. $X_{c,t}$ are the propagation controls described before, which include the county's latitude and longitude, polynomials of mean elevation, standard deviations, and minimum distance to nearest broadcast tower at time t. I test this by regressing county-level infrastructure growth on my predicted signal loss instrument, partialing out LocalLM fixed effects $\times$ wave and cohort fixed effects.

Table 4 shows the estimated coefficients of such exercise. Household infrastructure refers to access to electricity, water, and sewage systems available in every Mexican Census from 1960 to 2020. Each cell represents the estimated coefficient from the regression explained above. The first row shows the relationship between $\overline{\text{SignalLoss}}$ and contemporary basic household infrastructure. The instrument is generally negatively correlated with current infrastructure access, as places with lower signal loss (better signal) tend to have higher access to electricity, water, and sewage systems. However, these coefficients become statistically insignificant once I include the propagation controls. This highlights the importance of using the residual signal loss after controlling for geographic and propagation characteristics in each county. The second, third, and last rows show the relationship between $\overline{\text{SignalLoss}}$ and changes in each household infrastructure over a 10-year period (from year t to $t + 10$), 20-year period, and from t to the current survey wave as proxies for development, respectively. Across electricity, water, and sewage access, the coefficients are statistically insignificant, suggesting that the instrument is not capturing long-term development trends.

Table 4: SignalLoss Exogeneity Test on Housing Infrastructure Growth

	Infrastructure Development					
	Electricy		Water		Sewage	
SignalLoss effect timing						
At teenage exposure	−0.002	−0.001	0.006	0.006	−0.009*	0.000
	(0.003)	(0.004)	(0.005)	(0.006)	(0.005)	(0.005)
At teenage exposure +10 years	−0.003	0.001	0.005	0.007	−0.006	0.001
	(0.002)	(0.003)	(0.005)	(0.007)	(0.007)	(0.009)
At teenage exposure +20 years	−0.008***	−0.003	−0.003	0.007	−0.006	0.006
	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
At current survey wave	−0.003**	0.002	−0.002	0.004	−0.014***	−0.000
	(0.001)	(0.002)	(0.006)	(0.007)	(0.005)	(0.005)
Propagation Controls		✓		✓		✓

Notes: Each cell represents the coefficient of the estimated regression Y. Each column represents the percentage of households that have access to electricity, water, or a sewage system as proxies for housing infrastructure development growth. Rows show the outcome is contemporary or changes over time. Propagation controls include the county's latitude and longitude, polynomials of mean elevation, standard deviations, and minimum distance to the nearest broadcast tower at time t. Clustered standard errors at the county level are displayed in parentheses.

IV Regression Taking all the previous issues into account, the main regression specification is as follows:

$$\widehat{\text{TeenFEI}}_{i,c,w} = \gamma_1 + \gamma_2 \text{SignalLoss}_{i,c,w} + \delta_{l,w} + \delta_b + \varepsilon^1_{i,c,w}$$

$$\text{Y}_{i,c,w} = \alpha + \beta \widehat{\text{TeenFEI}}_{i,c,w} + \delta_{l,w} + \delta_b + \varepsilon^2_{i,c,w}$$

The sample is restricted to all working-age Mexican women i living in county c surveyed in wave w. $\text{SignalLoss}_{i,c,w}$ captures the minimum signal loss for individual i living in county c during their teenage years, measured in decibels (dB). $\text{TeenFEI}_{i,c,w}$ measures the z-score of the cumulative FEI score of telenovelas that individual i was exposed to during their teenage years (ages 13–18). $\text{Y}_{i,c,w}$ is the labor market outcome of interest for individual i living in county c during wave w. The LocalLM fixed effects and the LocalLM fixed effect × wave trends are represented by $\delta_{l,w}$, which should capture LocalLM specific differences and differential development trends across space. The cohort specific trends reflecting different gender norms are captured by δ_b.

5 Results

Table 5 presents the main estimation results, where I progressively add fixed effects across columns. For each specification, both OLS and IV estimates are reported. The table demonstrates that exposure to a standard deviation of TeenFEI increases women's likelihood of participating in the formal labor market across all specifications. Column 1 shows the effect without any controls or fixed effects, and an additional standard deviation of TeenFEI is associated with a 3 percentage point increase in labor market participation using OLS and a 18 p.p. increase using IV. Column 2 adds LocalLM fixed effects, which reduces the IV estimates to a 8 p.p. increase. After accounting for all spatial heterogeneity, the instrument remains strong, with a first-stage F-statistic of 521. This suggests that the instrument mostly captures differences in signal loss across cohorts. Column 3 includes census wave fixed effects that account for any temporal trends across cross-section periods, yielding an IV estimate similar to that in the specification without fixed effects. Column 4 includes cohort fixed effects, with an estimated impact of 30 p.p. on the likelihood of labor participation. The first-stage F-statistic in this specification is 48, significantly lower than in Column 2. This suggests that most of the variation in TeenFEI captured by the instrument comes from differences in exposure between

Table 5: TeenFEI Effects on Female Labor Market Participation

	Labor Market Participation						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: OLS							
TeenFEI	0.032***	0.032***	0.029***	0.023***	0.027***	0.009***	0.009***
	(0.003)	(0.001)	(0.003)	(0.006)	(0.001)	(0.003)	(0.003)
N				10,908,351			
R^2	0.004	0.058	0.008	0.008	0.065	0.068	0.068
Panel B: IV							
TeenFEI	0.179***	0.082***	0.179***	0.311***	0.077***	0.157***	0.128***
	(0.024)	(0.006)	(0.024)	(0.055)	(0.006)	(0.023)	(0.018)
N				9,572,244			
First-stage F	119.3	512.5	118.0	48.0	512.2	123.3	41.6
LocalLM FE		✓			✓	✓	✓
Wave FE			✓		✓	✓	✓
Cohort FE				✓		✓	✓
LocalLM FE × Wave					✓	✓	✓
Propagation Controls							✓

Notes: The sample for all columns consists of Mexican women between 25 and 60. The outcome variable is whether the individual is participating in the formal labor market. LocalLM refers to the geographic unit of analysis of Local Labor Markets defined by (Aldeco *et al.*, 2024). Propagation controls include the county's latitude and longitude, polynomials of mean elevation, standard deviations, and minimum distance to nearest broadcast tower. Clustered standard errors at the county level are displayed in parentheses.

cohorts, with a smaller proportion attributable to differences within cohorts across space. Column 5 combines LocalLM fixed effect with a wave trend, with an estimated impact of 8 p.p. on the likelihood of labor participation. Column 6 adds both the LocalLM fixed effect with a census wave trend and cohort fixed effects. Column 7 includes signal propagation controls. Using the last specification, I find that a standard deviation increase in FEI during teenage years increases the probability of labor participation by 13 p.p., suggesting that constant exposure to telenovelas with female empowerment content has a long-run effect on women's labor market participation.

The substantial differences between OLS and IV estimates in Table 5 can be attributed to several factors. First, the instrumental variable addresses the measurement error inherent in the FEI index. As discussed in Section 2.1.3, the generative AI models used to construct the FEI are not perfectly accurate at predicting female empowerment content, and the measure may not capture all aspects of empowered female characters that could serve as role models for teenagers. The determinants of signal loss (such as terrain elevation and distance to broadcast towers) are unrelated to the generative AI models used to construct the FEI, ensuring that the instrument does not directly affect the measurement process or content classification. Second, the 2SLS specification yields a Local Average Treatment Effect (LATE) that represents a weighted average across different complier groups (Angrist and Pischke, 2009). The IV estimates are therefore higher because they assign greater weight to women in specific regions and cohorts who are most responsive to telenovela exposure. Following this causal inference framework, the estimated compliance rate is around 7%, suggesting that around 670,000 Mexican women are most susceptible to empowered female characters.[26]

[26] The compliance rate is the intent to treat captured by the OLS regression (0.009), divided by the estimate from the IV regression (0.128) equal to 0.07 (Angrist and Pischke, 2009).

The estimated effects are sizable for the complier group, so it is important to interpret what a one standard deviation change in TeenFEI exposure represents. For the average individual in the sample (mean exposure of 203), a one standard deviation increase corresponds to a 116 change of FEI. This change is equivalent to a 57% increase in exposure, roughly equivalent to moving from the 25th to the 60th percentile of the distribution. Changes of this magnitude are common across cohorts: the difference in TeenFEI between the 1960 and 1970 cohorts is approximately 70 units (roughly two-thirds of a standard deviation), while the gap between the 1960 and 1980 cohorts reaches nearly 200 units, or 1.7 standard deviations. Given that these effects are large, it is useful to interpret them in terms of marginal changes in exposure. Overall, a 10% increase in TeenFEI exposure (roughly 20 units) translates into a 4% rise in the likelihood of labor market participation, or about two percentage points from a 47% average female labor force participation rate.

These results are quantitatively similar to findings from other studies examining other cultural shocks. Fernández, Fogli and Olivetti (2004) find that wives whose husband's mother worked are 9 to 24 percentage points more likely to join the labor force, a magnitude comparable to the effect of a standard deviation increase in TeenFEI in my estimates. This highlights that cultural effects can be substantially larger for specific subgroups of the population. Within more minor marginal effects, the results are also comparable to the impact of women who move to less sexist U.S. states, experiencing an approximately 3 p.p. increase in labor market participation probability (Charles, Guryan and Pan, 2022).[27] To achieve a similar 3 p.p. effect, a susceptible Mexican woman would need to increase her exposure to FEI by 30 points. Lastly, Fernández and Fogli (2009) shows that increases in 1950s labor force participation by country of origin, as a proxy for culture, increase the labor force participation of second-generation Americans by 5 p.p., comparable to a TeenFEI increase of 50.

What fraction of the overall labor force participation rate increase from 2010 to 2020 can be attributed to changes in exposure to empowered female characters? I use the estimated intent to treat effect derived from the OLS estimates found in Column 7 of Table 5, since the IV estimates reflect local effects for a specific complier group and should not be generalized to the entire population. The data, detailed in Panel B of Table 3, provides the average TeenFEI exposure and labor force participation rate from the 2010 and 2020 census waves for various age groups. The analysis hinges on the estimate that an increase of TeenFEI of 116 is associated with a 1 percentage point increase in labor market participation. For an average working-age Mexican woman the labor force participation rate increased from 46% to 52% from 2010 to 2020, with a gap of 6 percentage points. Over the same time period, the average TeenFEI increased from 160 to 219, a change of 59 units. A 59 increase in TeenFEI is thus associated to a 0.5 p.p. increase in labor market participation, implying that the increase in TeenFEI can explain 8.4% of the total increase in labor force participation.[28] Applying the same calculation to specific ages, I find that changes in TeenFEI can explain 11% of the increase in the labor force participation rate for 35-year-olds, 10% for 45-year-olds, and 5% for 55-year-olds.

5.1 Heterogeneous Effects

I next examine how the effects of telenovela exposure differ across various demographic characteristics and life circumstances. I start by exploring the effects of TeenFEI on different age groups. Table 6 shows the results of estimating the equation across different age group samples in columns 1 to 3. The results still indicate that across all age groups, a standard deviation increase of TeenFEI has a positive effect on the likelihood of women participating in the labor market. The effect is most pronounced in older women, with an estimated effect of 30 p.p. for those aged 51-65, compared to a 13 p.p. increase

[27] An important caveat is that these studies examine different cultural contexts. The media landscape and societal norms in the U.S. may differ significantly from those in Mexico. For instance, World Values Survey data shows that 24% of Mexicans agree that men make better business leaders than women, compared to 17% of U.S. respondents.

[28] This proportion is somewhat higher than the 7% of Brazil's fertility decline attributed to telenovela exposure by La Ferrara, Chong and Duryea (2012).

Table 6: TeenFEI Effects on Labor Market Participation by Age and Education Group

	Labor Market Participation					
	Age Groups			Educ Attainment Group		
	25-35	35-50	50-60	elem.	high.	uni.
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
TeenFEI	0.02**	0.01***	0.04***	0.01***	0.01**	0.00
	(0.007)	(0.003)	(0.007)	(0.003)	(0.003)	(0.002)
N	4,147,275	4,644,787	2,116,289	9,841,859	3,034,828	1,142,323
R^2	0.069	0.068	0.051	0.059	0.024	0.022
Panel B: IV						
TeenFEI	0.14***	0.13***	0.31***	0.13***	0.07***	0.01
	(0.030)	(0.015)	(0.089)	(0.017)	(0.013)	(0.012)
N	4,144,058	4,069,156	1,359,030	8,837,165	2,861,092	1,084,173
First-stage F	13.7	75.0	3.9	41.0	24.0	19.0
LocalLM FE	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓
Wave FE	✓	✓	✓	✓	✓	✓
LocalLM FE $\times$ Wave	✓	✓	✓	✓	✓	✓
Propagation Controls	✓	✓	✓	✓	✓	✓
Outcome mean	0.39	0.40	0.30	0.39	0.59	0.73

Notes: The sample consists of Mexican women aged 25–60. Each column corresponds to a specific subgroup: columns 1–3 display results by age group (25–35, 36–50, 51–65), and columns 4–6 display results by education level (elementary, high school, college). The outcome is formal labor market participation. LocalLM refers to the geographic unit of analysis of Local Labor Markets defined by (Aldeco *et al.*, 2024). Propagation controls include the county's latitude and longitude, polynomials of mean elevation, standard deviations, and minimum distance to nearest broadcast tower. Clustered standard errors at the county level are shown in parentheses.

for those aged 25-50. However, the first-stage relationship is weaker for this older group because there is less variation in the residual signal loss. With a first-stage F-statistic of 3.9, the estimates for this group should be interpreted with caution.

There are also significant differences in the effect of TeenFEI across different education attainment groups. Table 6 shows the results of estimating the equation using different educational attainment samples in columns 4 to 6. A standard deviation of TeenFEI increases the likelihood that women with at most an elementary or high school diploma participate in the labor market by around 13 and 7 p.p., respectively. However, those with a college degree do not experience a significant effect. This is expected since women with sights to get a college degree would be a subgroup that is less likely to be influenced by telenovelas. Women who aspire to higher education may be less susceptible to the influence of telenovelas because they face lower uncertainty about their future career paths and may rely less heavily on media representations. In other words, these women are likely to have more favorable gender norms and, therefore, are less influenced by what they watch on television.

5.2 Other Outcomes

Beyond labor market participation, I also examine how TeenFEI exposure affects other labor market outcomes. Table 7 presents the estimated effects of FEI exposure on a range of labor market indicators and personal decisions. Conditional on participating in the labor market, women exposed to higher FEI seem to be less likely to work in male-dominated industries and occupations; however, this is not statistically significant for male-dominated industries when considering the IV estimates. In particular, a standard deviation increase in TeenFEI decreases by 2.5 p.p. the likelihood of working in male-dominated industries and decreases by 3.8 p.p. in male-dominated occupations. Male-dominated industries and occupations are typically in STEM fields. Another outcome focused on in the literature is the likelihood of a managerial position. Column 3 shows that the TeenFEI effect is positive and statistically significant in determining her likelihood of occupying a managerial role. Although the point estimate seems small (1.5 p.p. increase), this corresponds to a 60% increase, since there are only around 2.5% of working women who occupy a managerial role. Lastly, FEI exposure is associated with higher monthly wages for women, suggesting that women either chose higher paying occupations or were able to negotiate better salaries. Overall, these results indicate that FEI exposure primarily influences women's decisions to participate in the labor market and, once employed, to pursue better-paying and higher-ranking positions. However, it does not necessarily lead women to enter male-dominated fields.

Table 7 also presents the results of teenage exposure to empowered female characters on educational choices. A standard deviation increase in TeenFEI increases the number of years in education by 2.2 years. Additionally, this exposure increases the likelihood that an individual will obtain an elementary, high school, or university degree. A standard deviation increase in TeenFEI increases the likelihood of obtaining a university degree by 0.04 percentage points, which translates to a 3.6% increase from an average Mexican woman's likelihood of obtaining a university degree. This pattern

Table 7: TeenFEI Effects on Other Labor Market Outcomes and Educational Attainment

	Labor Market				Educational Choices			
	Male Dominated		Manager	Wages	Educ	Highest Educ. Level		
	industry	occup.		MX$	years	elem.	high.	uni.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: OLS								
TeenFEI	−0.000	0.001	−0.000	0.226	0.248***	0.017***	0.007**	0.006***
	(0.002)	(0.002)	(0.000)	(0.786)	(0.038)	(0.003)	(0.003)	(0.002)
N	4,002,120				10,846,398	10,871,312	10,871,312	10,871,312
R^2	0.039	0.066	0.007	0.003	0.250	0.155	0.121	0.057
Panel B: IV								
TeenFEI	−0.025**	−0.038***	0.015***	13.324*	2.204***	0.079***	0.156***	0.088***
	(0.012)	(0.011)	(0.003)	(6.988)	(0.295)	(0.016)	(0.022)	(0.013)
N	3,634,601				9,517,860	9,540,069	9,540,069	9,540,069
First-stage F	33.9	33.8	33.9	33.8	41.4	41.4	41.4	41.4
LocalLM FE	✓	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓
Wave FE	✓	✓	✓	✓	✓	✓	✓	✓
LocalLM FE × Wave	✓	✓	✓	✓	✓	✓	✓	✓
Propagation Controls	✓	✓	✓	✓	✓	✓	✓	✓
Outcome mean	0.435	0.359	0.025	10.872	7.918	0.905	0.279	0.105

Notes: The sample for all columns consists of Mexican women aged 25-60. Each column represents a different outcome variable: columns 1-4 cover labor market outcomes, while columns 5-8 focus on educational choices. The industry refers to broad sectors, such as transportation and tourism, whereas occupation refers to more specific job roles, like a truck driver. An industry or occupation is considered male-dominated if more men than women work in that category. Manager indicates the likelihood of a woman holding a managerial position. Wages are measured as monthly earnings in 1,000 $MX. The highest education level refers to the highest degree that an individual has earned. The outcome mean is the average value for that outcome in the sample. LocalLM refers to the geographic unit of analysis of Local Labor Markets defined by (Aldeco *et al.*, 2024). Propagation controls include the county's latitude and longitude, polynomials of mean elevation, standard deviations, and minimum distance to nearest broadcast tower. Clustered standard errors at the county level are shown in parentheses.

suggests that exposure to empowering telenovela content during adolescence may encourage young women to pursue higher education, aligning with their aspirations for future formal labor market participation.

Empowered female characters also influence women's family decisions. Table 8 reports estimates for cohabitation, marriage, motherhood, and the number of children using the main specification. A standard deviation increase in TeenFEI is associated with small declines in the probabilities of cohabitation, marriage, and motherhood (each around a 2 percentage point decrease), while it is associated with a reduction of about one children per woman. These findings are consistent with previous work. For example, La Ferrara, Chong and Duryea (2012) finds that the introduction of a new telenovela channel in Brazil reduced fertility and changed naming patterns. Overall, the evidence suggests that modestly empowering telenovela content shifts family behavior, resulting in lower fertility, while its effects on marriage and cohabitation are negligible. This reinforces the notion that the most significant impacts of exposure are observed in labor market participation, education, and fertility.

5.3 Potential mechanisms

5.3.1 Emotions

Psychology research indicates that emotionally charged messages can influence individuals differently, a phenomenon known as attribute framing (Tversky and Kahneman, 1981; Piñon and Gambara, 2005). In this context, the magnitude and direction of TeenFEI effects depend on the emotions associated with the empowered characters. For example, in our example telenovela, *La Patrona*, the main character is mostly associated with revenge, as her entire story revolves around

Table 8: TeenFEI Effects on Family Choices

	Co-habitation	Married	Mother	Fertility
	(1)	(2)	(3)	(4)
Panel A: OLS				
TeenFEI	0.0056***	−0.0070***	−0.0000	−0.1716***
	(0.00132)	(0.00129)	(0.00138)	(0.01800)
N	10,908,351	10,883,884	10,908,351	10,908,351
R^2	0.014	0.182	0.046	0.166
Panel B: IV				
TeenFEI	−0.0391***	−0.0117*	−0.0155**	−1.0346***
	(0.00781)	(0.00644)	(0.00671)	(0.11747)
N	9,572,244	9,551,057	9,572,244	9,572,244
First-stage F	41.6	41.3	41.6	41.6
LocalLM FE	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓
Wave FE	✓	✓	✓	✓
LocalLM × Wave FE	✓	✓	✓	✓
Propagation Controls	✓	✓	✓	✓
Outcome mean	0.691	0.552	0.866	2.917

Notes: The sample for all columns consists of Mexican women aged 25-65. TODO The outcome mean is the average value for that outcome in the sample. LocalLM refers to the geographic unit of analysis of Local Labor Markets defined by (Aldeco *et al.*, 2024). Propagation controls include the county's latitude and longitude, polynomials of mean elevation, standard deviations, and minimum distance to nearest broadcast tower. Clustered standard errors at the county level are shown in parentheses.

payback and her anger towards the villain. In these cases, the audience might interpret this as a negative message, suggesting that empowered women must face hardships or be disliked by those they fight against. In contrast, telenovelas that depict empowered women with emotions such as joy might create a positive association with empowerment.

To explore this, I estimate the emotion attached to empowered characters in the synopsis using `pysentimiento`. This is a machine learning model that predicts the emotions of a piece of text (Pérez *et al.*, 2024).[29] The model outputs six possible emotions: anger, sadness, fear, surprise, love, and joy, with a corresponding probability. For example, the model predicts the main emotion associated with *La Patrona* is anger. Other Telenovelas provide other sentiments. For example, the telenovela *Las Bandidas* (The Female Bandits), a comedy about the lives and loves of the Montoya sisters with different social classes who begin to hunt down the killers of their loved ones as inexperienced bandits, is categorized as joy. This makes sense, since, although their dark journey is filled with challenges, the overall tone remains light-hearted and humorous. A classic example of a telenovela associated with love is *Enamorandome de Ramon* (Falling in Love with Ramon), where two sisters navigate familial greed and social prejudices while discovering love amidst trials with a mechanic.

To assess whether attribute framing plays a role in telenovela content, I estimate the effect of TeenFEI exposure linked to each of the six predicted emotions on labor market participation using the following OLS model:

$$\mathrm{Y}_{i,c,w} = \alpha + \sum^{E} \beta_e \mathrm{TeenFEI}_{i,c,w,e} + X_i'\omega + \delta_{c,w} + \delta_b + \varepsilon_{i,c,w}$$

where $E = \{\text{anger, sadness, fear, surprise, love, and joy}\}$. The coefficients β_e represent the effect of the TeenFEI exposure associated with emotion e on labor market participation, with the sum TeenFEI of other telenovelas depicting the other 5 emotions as constant. As an illustration, β_e for anger is the equivalent of adding several telenovelas that primarily evoke anger to the viewing experience. Since I have only one instrument for TeenFEI, I cannot separately instrument each emotion-specific TeenFEI variable, which makes it difficult to fully address potential endogeneity concerns for these coefficients. Therefore, these results should be interpreted cautiously, and the magnitudes may be attenuated compared to the causal effects identified in the main analysis.

Figure 10 shows the results from the emotion model compared to the benchmark case.[30] The positive emotions of joy, love, and surprise are associated with a positive effect on labor market participation. This is consistent with the idea that female empowered content associated with a positive emotion should incentivize women to work. Emotions like fear and sadness have a negative effect on labor market participation, confirming that somewhat negative emotions can discourage participation. The emotion of anger has a strong positive effect on labor market participation. Although counterintuitive, this can also be explained by attribute framing. Anger becomes a motivational force that empowers women to challenge existing constraints and pursue economic independence. The narrative of righteous anger against injustice may resonate particularly strongly with viewers who face similar barriers in their own lives, encouraging them to take action rather than accept the status quo.

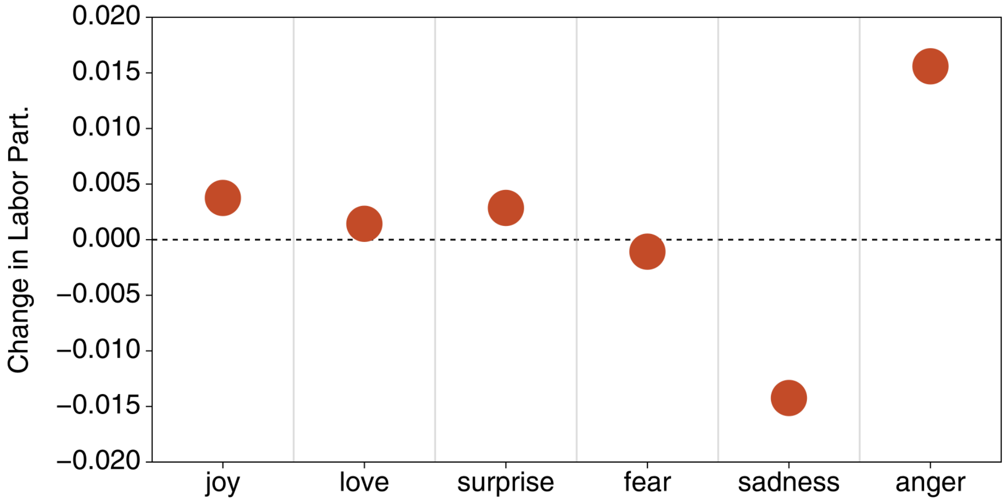
5.3.2 Traditional versus Non-Traditional Female Jobs

The type of job that the female characters have in telenovelas can also have differential effects on female labor market participation. If female characters consistently have traditional female jobs (e.g., nurse, teacher, or homemaker), this may reinforce existing gender norms and limit the perceived

[29]This model can also process text and predict emotions for English, Portuguese, and Italian.

[30]Note that $\sum^{E} \mathrm{TeenFEI}_{i,l,e} = \mathrm{TeenFEI}_{i,l}$ should be the same as the benchmark case since these emotions are exhaustive.

Figure 10: Effects of TeenFEI by Emotional Content

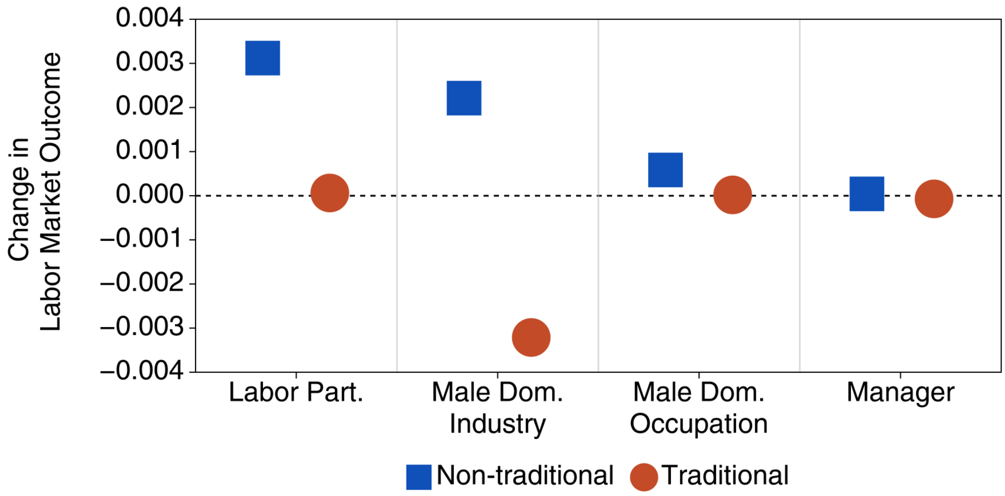


Notes: Each point shows the estimated effect of a standard deviation of TeenFEI of telenovelas featuring different emotions. The 95% confidence intervals are shown in grey error bars using clustered standard errors at the county level.

range of career options for viewers. Conversely, if female characters occupy non-traditional roles (e.g., businesswoman, doctor, or politician), this could inspire viewers to pursue similar paths and break the gender norms.

To examine this, I categorize my telenovela sample based on whether the female protagonist or other secondary female characters hold traditional or non-traditional female jobs. I then estimate the effect of TeenFEI exposure to those telenovelas on labor force participation, probability of working in a male-dominated industry, male-dominated occupation, and holding a managerial role. Figure 11 shows the coefficients of a single regression, including the traditional TeenFEI and non-traditional TeenFEI. Exposure to non-traditional female jobs is positively associated with higher labor market participation, while traditional female jobs have a negligible and statistically insignificant effect. In terms of working in a male-dominated industry, I find that the depiction of jobs associated with traditional gender norms reduces the likelihood of working in such an industry. Non-traditional jobs are positively associated with a higher likelihood of working in such industries. However, the effects of both types of jobs represented in telenovelas have small and non-significant effects on the

Figure 11: Effects of TeenFEI by Traditional vs Non-Traditional Female Jobs



Notes: Each point shows the estimated effect of a standard deviation of TeenFEI of telenovelas featuring traditional or not traditional female jobs. The 95% confidence intervals are shown in grey error bars using clustered standard errors at the county level.

likelihood of working in male-dominated occupations or holding a managerial role. Overall, these results suggest that the type of work female characters portray in telenovelas can influence viewers' perceptions of gender roles, albeit with limited effects on their occupational choices.

6 Conclusion

This study examines the causal impact of exposure to female empowerment narratives in Latin American telenovelas on women's labor force participation. I developed a novel Female Empowerment Index (FEI) for approximately 2,000 telenovelas (1960-2024) using generative AI. I find that adolescent exposure to FEI is linked to higher labor force participation and more progressive gender attitudes in Latin America. To establish causality, I employed an instrumental variable strategy in Mexico, leveraging exogenous geographic and temporal variation in broadcast television signal loss. The central finding is that greater exposure to empowered female characters significantly increases a woman's likelihood of participating in the labor force as an adult. Quantitatively, a 10% increase in FEI exposure is associated with a 4% rise in the probability of labor market participation, accounting for nearly 10% of the observed changes in labor force participation between 2010 and 2020. This effect is most pronounced among women with lower educational attainment.

The research also identified several positive impacts on other life outcomes. Increased exposure to FEI during adolescence was associated with higher educational attainment, including a greater likelihood of completing a university degree. It was linked to higher wages and an increased probability of becoming a manager. Additionally, exposure significantly decreased fertility rates. However, the impact of content proved sensitive to the specifics of portrayal. Increases in female labor force participation were primarily driven by portrayals of women in non-traditional careers (e.g., doctor or lawyer). In contrast, portrayals of traditional female jobs had negligible or even negative effects. Furthermore, empowerment narratives associated with joy or anger were strong motivators, while those linked to sadness or fear appeared to function as cautionary tales. Overall, the findings demonstrate that entertainment media is an active force in cultural formation, translating specific gender narratives into tangible changes in women's economic decisions.

Bibliography

Aldeco, L. *et al.* (2024) *Local Labor Markets in Mexico: definition, data bases, and descriptive analysis 1990-2020*. Available at: https://www.banxico.org.mx/apps/datasets/merclab/%7B2BB5ECD9-4503-30DD-93A3-B70FFE6A52FF%7D.pdf.

Alesina, A., Giuliano, P. and Nunn, N. (2013) "On the Origins of Gender Roles: Women and the Plough," *The Quarterly Journal of Economics*, 128(2), pp. 469–530. Available at: https://doi.org/10.1093/qje/qjt005.

Angrist, J.D. and Pischke, J.-S. (2009) *Mostly Harmless Econometrics | Princeton University Press*. Princeton University Press. Available at: https://press.princeton.edu/books/paperback/9780691120355/mostly-harmless-econometrics (Accessed: September 2, 2025).

Antezana, L. *et al.* (2022) *Telenovelas Chilenas*. FONDECYT. Available at: https://uchile.cl/dam/jcr:0fac0e48-2c1e-4ce2-a1e4-72211c5577a7/Boletin1fondecyt%20(1)%20(1).pdf.

Antezana, L. *et al.* (2023) *Audiencias Juveniles*. FONDECYT. Available at: https://uchile.cl/dam/jcr:c764eb9e-70aa-44c9-964c-00f59b6b30bd/Boletin2fondecyt%20%5B06-10-2023%5D%20(1)%20(1).pdf.

Apel, M., Blix Grimaldi, M. and Hull, I. (2022) "How Much Information Do Monetary Policy Committees Disclose? Evidence from the FOMC's Minutes and Transcripts," *Journal of Money, Credit and Banking*, 54(5), pp. 1459–1490. Available at: https://doi.org/10.1111/jmcb.12885.

Appel, M. and Gnambs, T. (2023) "Women in fiction: Bechdel-Wallace Test results for the highest-grossing movies of the last four decades," *Psychology of Popular Media*, 12(4), pp. 499–504. Available at: https://doi.org/10.1037/ppm0000436.

Arceneaux, K. and Johnson, M. (2013) *Changing Minds or Changing Channels?: Partisan News in an Age of Choice*. Chicago, IL: University of Chicago Press (Chicago Studies in American Politics). Available at: https://press.uchicago.edu/ucp/books/book/chicago/C/bo15731464.html (Accessed: December 31, 2024).

Ash, E. and Hansen, S. (2023) "Text Algorithms in Economics," *Annual Review of Economics*, 15(Volume15, 2023), pp. 659–688. Available at: https://doi.org/10.1146/annurev-economics-082222-074352.

Banerjee, A., Ferrara, E.L. and Orozco, V. (2019) "Entertainment, Education, and Attitudes Toward Domestic Violence," *AEA Papers and Proceedings*, 109, pp. 133–137. Available at: https://doi.org/10.1257/pandp.20191073.

Banerjee, A., La Ferrara, E. and Orozco-Olvera, V.H. (2019) "The Entertaining Way to Behavioral Change: Fighting HV with MTV." *Working Paper Series*.

Bertrand, M. (2020) "Gender in the Twenty-First Century," *AEA Papers and Proceedings*, 110, pp. 1–24. Available at: https://doi.org/10.1257/pandp.20201126.

Besley, T. and Burgess, R. (2002) "Political Economy of Government Responsiveness: Theory and Evidence from India," *The Quarterly Journal of Economics*, 117(4), pp. 1415–1451. Available at: https://doi.org/https://doi.org/10.1162/003355302320935061.

Chann, S. (2023) *Non-determinism in GPT-4 is caused by Sparse MoE*. Available at: https://152334h.github.io/blog/non-determinism-in-gpt-4/ (Accessed: December 11, 2024).

Charles, K.K., Guryan, J. and Pan, J. (2022) "The Effects of Sexism on American Women: The Role of Norms vs. Discrimination," *Journal of Human Resources* [Preprint]. Available at: https://doi.org/10.3368/jhr.0920-11209R3.

Chong, A. and La Ferrara, E. (2009) "Television and Divorce: Evidence from Brazilian "Novelas"," *Journal of the European Economic Association*, 7(2/3), pp. 458–468. Available at: https://www.jstor.org/stable/40282763 (Accessed: November 27, 2023).

Clayton, C. *et al.* (2025) *Geoeconomic Pressure*. Rochester, NY: Social Science Research Network. Available at: https://papers.ssrn.com/abstract=5350350 (Accessed: July 22, 2025).

Desmet, K., Ortuño-Ortín, I. and Wacziarg, R. (2017) "Culture, Ethnicity, and Diversity," *American Economic Review*, 107(9), pp. 2479–2513. Available at: https://doi.org/10.1257/aer.20150243.

Drago, F., Nannicini, T. and Sobbrio, F. (2014) "Meet the Press: How Voters and Politicians Respond to Newspaper Entry and Exit," *American Economic Journal: Applied Economics*, 6(3), pp. 159–188. Available at: https://doi.org/10.1257/app.6.3.159.

Duclos, J.-Y., Esteban, J. and Ray, D. (2004) "Polarization: Concepts, Measurement, Estimation," *Econometrica*, 72(6), pp. 1737–1772. Available at: https://doi.org/10.1111/j.1468-0262.2004.00552.x.

Ellemers, N. (2018) "Gender Stereotypes," *Annual Review of Psychology*, 69(Volume69, 2018), pp. 275–298. Available at: https://doi.org/10.1146/annurev-psych-122216-011719.

Esteban, J.-M. and Ray, D. (1994) "On the Measurement of Polarization," *Econometrica*, 62(4), pp. 819–851. Available at: https://doi.org/10.2307/2951734.

Etchegaray, N., Scherman, A. and Valenzuela, S. (2019) "Testing the Hypothesis of "Impressionable Years" With Willingness to Self-Censor in Chile," *International Journal of Public Opinion Research*, 31(2), pp. 331–348. Available at: https://doi.org/10.1093/ijpor/edy012.

Fernández, R. (2013) "Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century," *American Economic Review*, 103(1), pp. 472–500. Available at: https://doi.org/10.1257/aer.103.1.472.

Fernández, R. and Fogli, A. (2009) "Culture: An Empirical Investigation of Beliefs, Work, and Fertility," *American Economic Journal: Macroeconomics*, 1(1), pp. 146–177. Available at: https://doi.org/10.1257/mac.1.1.146.

Fernández, R., Fogli, A. and Olivetti, C. (2004) "Mothers and Sons: Preference Formation and Female Labor Force Dynamics*," *The Quarterly Journal of Economics*, 119(4), pp. 1249–1299. Available at: https://doi.org/10.1162/0033553042476224.

Ferrara, E.L. (2016) "Mass Media and Social Change: Can We Use Television to Fight Poverty?," *Journal of the European Economic Association*, 14(4), pp. 791–827. Available at: https://ideas.repec.org//a/oup/jeurec/v14y2016i4p791-827..html (Accessed: June 12, 2025).

La Ferrara, E., Chong, A. and Duryea, S. (2012) "Soap Operas and Fertility: Evidence from Brazil," *American Economic Journal: Applied Economics*, 4(4), pp. 1–31. Available at: https://doi.org/10.1257/app.4.4.1.

Galor, O., Özak, Ö. and Sarid, A. (2018) *Geographical Roots of the Coevolution of Cultural and Linguistic Traits*. Rochester, NY: Social Science Research Network. Available at: https://doi.org/10.2139/ssrn.3284239.

Gentzkow, M. (2006) "Television and Voter Turnout," *The Quarterly Journal of Economics*, 121(3), pp. 931–972. Available at: https://www.jstor.org/stable/25098813 (Accessed: April 22, 2024).

Gentzkow, M. and Shapiro, J.M. (2008) "Preschool Television Viewing and Adolescent Test Scores: Historical Evidence from the Coleman Study," *The Quarterly Journal of Economics*, 123(1), pp. 279–323. Available at: https://doi.org/10.1162/qjec.2008.123.1.279.

Gentzkow, M., Shapiro, J.M. and Stone, D.F. (2015) “Media Bias in the Marketplace: Theory,” *Handbook of Media Economics*. Edited by S.P. Anderson, J. Waldfogel, and D. Strömberg. North-Holland (Handbook of Media Economics). Available at: https://doi.org/10.1016/B978-0-444-63685-0.00014-0.

Gerbner, G. and Gross, L. (1976) “Living with television: The violence profile,” *Journal of Communication*, 26(2), pp. 172–199. Available at: https://doi.org/10.1111/j.1460-2466.1976.tb01397.x.

Gerbner, G. *et al.* (2002) “Growing up with television: Cultivation processes,” *Media effects: Advances in theory and research, 2nd ed*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers (LEA's communication series).

Gonzalez, J. (2025) *Drama and Redistribution: The Impact of Telenovelas on Preferences for Redistribution in Latin America*. Available at: https://jjgecon.github.io/assets/papers/Gonzalez_INE_Telenovelas.pdf.

Green, M.C., Brock, T.C. and Kaufman, G.F. (2004) “Understanding Media Enjoyment: The Role of Transportation Into Narrative Worlds,” *Communication Theory*, 14(4), pp. 311–327. Available at: https://doi.org/10.1111/j.1468-2885.2004.tb00317.x.

Guiso, L., Sapienza, P. and Zingales, L. (2006) “Does Culture Affect Economic Outcomes?,” *Journal of Economic Perspectives*, 20(2), pp. 23–48. Available at: https://doi.org/10.1257/jep.20.2.23.

Gulesci, S., Lombardi, M. and Ramos, A. (2024) “Telenovelas and attitudes toward the LGBTQ+ community in Latin America,” *Labour Economics*, 87, p. 102488. Available at: https://doi.org/10.1016/j.labeco.2023.102488.

Haese, J. (2025) “Gender in Movies.”

Hidalgo-Ternero, C.M. (2020) “Google Translate vs. DeepL: analysing neural machine translation performance under the challenge of phraseological variation,” *MonTI. Monographs in translation and interpreting*, pp. 154–177. Available at: https://doi.org/10.6035/MonTI.2020.ne6.5.

Holbrook, R.A. and Hill, T.G. (2005) “Agenda-Setting and Priming in Prime Time Television: Crime Dramas as Political Cues,” *Political Communication*, 22(3), pp. 277–295. Available at: https://doi.org/10.1080/10584600591006519.

INEGI (2022) *Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares (ENDUTIH)*. Available at: https://en.www.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/889463911180.pdf.

Jensen, R. and Oster, E. (2009) “The Power of TV: Cable Television and Women's Status in India,” *The Quarterly Journal of Economics*, 124(3), pp. 1057–1094. Available at: https://www.jstor.org/stable/40506252 (Accessed: April 22, 2024).

Jurafsky, D. and Martin, J.H. (2024) *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models*. 3rd ed. Available at: https://web.stanford.edu/~jurafsky/slp3 (Accessed: December 11, 2024).

Kearney, M.S. and Levine, P.B. (2015) “Media Influences on Social Outcomes: The Impact of MTV's 16 and Pregnant on Teen Childbearing,” *American Economic Review*, 105(12), pp. 3597–3632. Available at: https://doi.org/10.1257/aer.20140012.

Kearney, M.S. and Levine, P.B. (2019) “Early Childhood Education by Television: Lessons from Sesame Street,” *American Economic Journal: Applied Economics*, 11(1), pp. 318–350. Available at: https://doi.org/10.1257/app.20170300.

Krosnick, J.A. and Alwin, D.F. (1989) “Aging and susceptibility to attitude change,” *Journal of Personality and Social Psychology*, 57(3), pp. 416–425. Available at: https://doi.org/10.1037//0022-3514.57.3.416.

Lippmann, Q. and Montalbo, A. (2025) “Power, Media and Gender.”

Lopez, A. (2002) *Our welcomed guests: Telenovelas in Latin America*. Routledge.

Malmendier, U. and Nagel, S. (2011) “Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?*,” *The Quarterly Journal of Economics*, 126(1), pp. 373–416. Available at: https://doi.org/10.1093/qje/qjq004.

Martin, G.J. and Yurukoglu, A. (2017) “Bias in Cable News: Persuasion and Polarization,” *American Economic Review*, 107(9), pp. 2565–2599. Available at: https://doi.org/10.1257/aer.20160812.

Michalopoulos, S. (2012) “The Origins of Ethnolinguistic Diversity,” *American Economic Review*, 102(4), pp. 1508–1539. Available at: https://doi.org/10.1257/aer.102.4.1508.

Michalopoulos, S. and Rauh, C. (2024) *Movies*. *Working Paper Series*, National Bureau of Economic Research. Available at: https://doi.org/10.3386/w32220.

Michalopoulos, S. and Xue, M.M. (2021) “Folklore,” *The Quarterly Journal of Economics*, 136(4), pp. 1993–2046. Available at: https://doi.org/10.1093/qje/qjab003.

Morgan, M. and Shanahan, J. (2010) “The State of Cultivation,” *Journal of Broadcasting & Electronic Media*, 54(2), pp. 337–355. Available at: https://doi.org/10.1080/08838151003735018.

Olivetti, C., Pan, J. and Petrongolo, B. (2024) *The Evolution of Gender in the Labor Market*. *Working Paper Series*, National Bureau of Economic Research. Available at: https://doi.org/10.3386/w33153.

Olken, B.A. (2009) “Do Television and Radio Destroy Social Capital? Evidence from Indonesian Villages,” *American Economic Journal: Applied Economics*, 1(4), pp. 1–33. Available at: https://doi.org/10.1257/app.1.4.1.

Oughton, E.J. *et al.* (2020) “itmlogic: The Irregular Terrain Model by Longley and Rice,” *Journal of Open Source Software*, 5(51), p. 2266. Available at: https://doi.org/10.21105/joss.02266.

Pareja Sánchez, N. (2010) “Televisión y democracia: La televisión abierta y su oferta en la Ciudad de México,” *Andamios*, 7(14), pp. 101–135. Available at: http://www.scielo.org.mx/scielo.php?script=sci_abstract&pid=S1870-00632010000300005&lng=es&nrm=iso&tlng=es (Accessed: July 24, 2025).

Pastina, A.C.L., Rego, C.M. and Straubhaar, J.D. (2003) “The centrality of Telenovelas in Latin America's Everyday Life:Past Tendencies, Current Knowledge, and Future Research,” *Global Media Journal*, 2(2). Available at: https://www.globalmediajournal.com/open-access/the-centrality-of-telenovelas-in-latin-americas-everyday-lifepast-tendencies-current-knowledge-and-future-research.php?aid=35095 (Accessed: June 9, 2025).

Piñon, A. and Gambara, H. (2005) “A meta-analytic review of framming effect: risky, attribute and goal framing,” *Psicothema*, 17(2), pp. 325–331. Available at: https://www.psicothema.com/pdf/3107.pdf.

Porter, C. and Serra, D. (2020) “Gender Differences in the Choice of Major: The Importance of Female Role Models,” *American Economic Journal: Applied Economics*, 12(3), pp. 226–254. Available at: https://doi.org/10.1257/app.20180426.

Pérez, J.M. *et al.* (2024) *pysentimiento: A Python Toolkit for Opinion Mining and Social NLP tasks*. arXiv. Available at: https://doi.org/10.48550/arXiv.2106.09462.

Rathje, S. *et al.* (2024) "GPT is an effective tool for multilingual psychological text analysis," *Proceedings of the National Academy of Sciences*, 121(34), p. e2308950121. Available at: https://doi.org/10.1073/pnas.2308950121.

Sanz-Valdivieso, L. and López-Arroyo, B. (2023) "Google Translate vs. ChatGPT: Can non-language professionals trust them for specialized translation?," *Proceedings of the International Conference HiT-IT 2023*, pp. 97–107. Available at: https://unora.unior.it/bitstream/11574/219580/1/HiT-IT-2023-proceedings_compressed.pdf#page=113.

Schneider-Strawczynski, S. and Valette, J. (2025) "Media Coverage of Immigration and the Polarization of Attitudes," *American Economic Journal: Applied Economics*, 17(1), pp. 337–368. Available at: https://doi.org/10.1257/app.20230300.

Soroka, S.N. (2014) *Negativity in Democratic Politics: Causes and Consequences*. Cambridge: Cambridge University Press (Cambridge Studies in Public Opinion and Political Psychology). Available at: https://doi.org/10.1017/CBO9781107477971.

Strömberg, D. (2004) "Radio's Impact on Public Spending*," *The Quarterly Journal of Economics*, 119(1), pp. 189–221. Available at: https://doi.org/10.1162/003355304772839560.

Tversky, A. and Kahneman, D. (1981) "The Framing of Decisions and the Psychology of Choice," *Science*, 211(4481), pp. 453–458. Available at: https://doi.org/10.1126/science.7455683.

Törnberg, P. (2023) *ChatGPT-4 Outperforms Experts and Crowd Workers in Annotating Political Twitter Messages with Zero-Shot Learning*. arXiv. Available at: https://doi.org/10.48550/arXiv.2304.06588.

Wang, Z., Pang, Y. and Lin, Y. (2023) *Large Language Models Are Zero-Shot Text Classifiers*. arXiv. Available at: https://doi.org/10.48550/arXiv.2312.01044.

Yanagizawa-Drott, D. (2014) "Propaganda and Conflict: Evidence from the Rwandan Genocide *," *The Quarterly Journal of Economics*, 129(4), pp. 1947–1994. Available at: https://doi.org/10.1093/qje/qju020.

7 Appendix

A. Additional Tables and Figures

Table A.1: Internal Migration

Year	Same birth state	Same muni. (t-5)	Mexican Sample
1990	0.82	–	8.1m
2000	0.81	0.81	10.1m
2010	0.81	0.83	11.9m
2015	0.82	0.84	22.7m
2020	0.82	0.86	15.0m

Notes: The same birth state corresponds to the percentage of Mexicans who still live in the state where they were born. The same county corresponds to the percentage of people who have not moved to another county in the last 5 years. Source: Mexican ENOE's population census.

Table A.2: Results with Internal Migration

	Labor Market Participation				
	(1)	(2)	(3)	(4)	(5)
TeenFEI	0.128***	0.132***	0.120***	0.123***	0.122***
	(0.018)	(0.019)	(0.017)	(0.018)	(0.018)
LocalLM FE	✓	✓	✓	✓	✓
Wave FE	✓	✓	✓	✓	✓
LocalLM × Wave FE	✓	✓	✓	✓	✓
Propagation Controls	✓	✓	✓	✓	✓
Mexican		✓			✓
Remained State of Birth			✓		✓
Remained County past 5y				✓	✓
N	9,572,244	7,492,216	8,176,114	8,947,131	6,174,875
First-stage F	41.6	38.6	42.1	41.4	39.0

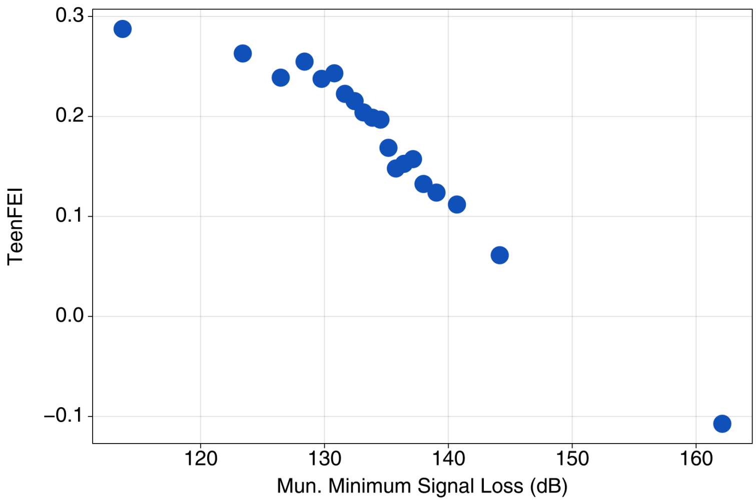
Notes: The table corresponds to the 2SLS coefficient estimates of a standard deviation of FEI on labor market participation. Mexican corresponds to the subset of individuals in the sample that are born in Mexico. The Remained State of Birth corresponds to the subset of individuals who still live in the state where they were born. Remained County past 5y corresponds to the subset of individuals who have not moved counties in the last 5 years. Clustered standard errors at the county level are reported in parentheses. Source: Mexican ENOE's population census.

Table A.3: Models performance

		Inequality		FEI	
Gen. AI model	Commercial	accuracy	precision	accuracy	precision
`gpt-4o`	✓	0.65	0.87	0.81	0.83
`gemini-2.0-flash`	✓	0.74	0.94	0.81	0.83
`claude-3-5`	✓	0.77	0.85	–	–
`phi-4`	×	0.68	0.81	0.82	0.84
`deepseek-8b`	×	0.66	0.55	0.83	0.84

Notes: The table shows the performance of different generative AI models in capturing inequality or the Female Empowerment Index (FEI) in telenovela synopses, using traditional metrics from the machine learning literature. Commercial refers to whether the model is a commercial product or open source. Higher numbers in accuracy and precision are better.

Figure A.1: First Stage Binscatter



Notes: The figure shows the first-stage relationship between the TeenFEI variable and the instrument variable, representing the minimum signal loss in a binscatter.