Drama and Redistribution:

The Impact of Telenovelas on Preferences for Redistribution in Latin America*

Javier Gonzalez

Southern Methodist University
gonzalezjj@smu.edu

September 2025

Latest Available Version

Abstract

This paper investigates the impact of telenovelas that depict inequality on support for redistribution in Latin America. Using a novel dataset of telenovelas aired from 1960 to 2024 in the region, I employed three Natural Language Processing methods to identify inequality telenovelas. My analysis focuses on instances where a new inequality telenovela is introduced during the LAPOP's fieldwork, ensuring that no other inequality telenovelas are simultaneously airing. Consistently, I observe a 4% reduction in support for redistribution following exposure to such telenovelas. These narratives seem to misrepresent the problems of inequality, portraying fewer issues than are present in reality, making individuals adjust their policy priorities. This study provides evidence of how entertainment media can shape the perceptions of inequality, thereby influencing policy preferences.

JEL Clasification: H11, H23, N16, P16, L82
Keywords: Redistribution, Media, Latin America, Inequality, Perceptions, Telenovelas

* Thanks to Ömer Özak, Klaus Desmet, Wookun Kim, Ivan de las Heras Tranche, Esteban Aleman, and Rouzhi Liang for unvaluable advise. I would also like to thank the Dedman College Interdisciplinary Institute at SMU for their financial support.

1 Introduction

Inequality perceptions are central to understanding the demand for redistributive policies (Alesina and Giuliano, 2011; Stantcheva, 2021). These perceptions typically arise from individuals' direct exposure to inequality within their country or neighborhood (Alesina, Stantcheva and Teso, 2018; Alesina, Miano and Stantcheva, 2023; Domènech-Arumí, 2023). The portrayal of inequality in traditional media can skew individuals' perceptions of income inequality within their country, impacting their trust in others (Diermeier *et al.*, 2017), and shaping their beliefs about upward mobility (Kim, 2023). Consequently, these (mis)perceptions of inequality subsequently impact individuals' likelihood of approving redistributive policies. Despite this, there is limited understanding of how entertainment media, with its potential to present a fictional and potentially inaccurate representation of inequality, affects perceptions of or attitudes towards redistribution.

This study explores the effects of exposure to telenovelas that depict inequality on support for redistribution in Latin America. I focus on telenovelas, commonly known as soap operas, a publicly and freely available TV show usually aired during prime time and central to Latin American culture. These shows appeal to a mass audience and usually display social norms, cultural identity, and the historical roots of each country (Lopez, 2002; Pastina, Rego and Straubhaar, 2003; Green, Brock and Kaufman, 2004; Antezana *et al.*, 2022; 2023). Their cultural impact is so strong that many newborns are named after telenovela characters (La Ferrara, Chong and Duryea, 2012). Beyond their entertainment value, telenovelas serve as powerful instruments for influencing individuals' behaviors and attitudes on various social issues throughout Latin America (Chong and La Ferrara, 2009; Bertrand, 2020; Gulesci, Lombardi and Ramos, 2024).

To estimate the effect of inequality telenovelas on support for redistribution, I use a novel dataset comprising telenovelas aired in the region from 1960 to 2024. This dataset contains information on each show's airtime, start and end dates, broadcasting channel, genre, and a plot synopsis. To identify telenovelas that specifically address issues of inequality, I employed Natural Language Processing (NLP) techniques on the synopses. This process involved implementing three distinct methods: (i) a dictionary-based approach that utilizes related terms extracted from ConceptNet, (ii) a vector embedding approach using the Universal Sentence Encoder (USE), (iii) a generative AI model capable of classifying content with no examples. Each of these methods has its strengths and limitations. The ConceptNet approach is the most intuitive but provides the simplest way to capture inequality themes, while the generative AI models are less transparent yet show the best performance in classification precision.

My empirical strategy centers on instances where a new inequality telenovela is introduced during the Latin American Public Opinion Project (LAPOP) fieldwork. I compare individuals currently exposed to inequality telenovelas (those interviewed after the telenovela's debut) with those who are not exposed (interviewed beforehand). The identification strategy encounters two key challenges. First, there is the concern that the groups interviewed before and after the telenovela's premiere differ significantly. To address this, I provide a balance table that reveals no significant differences among the samples. Second, the validity of the results depends on the accurate identification of inequality themes by the NLP methods. To validate this, I show that the introduction of an inequality telenovela is not associated with other political and social issues like the support for minorities. Moreover, these methods prove robust to variations in the threshold and indicate that telenovelas identified as lacking inequality content show no effect of exposure.

Exposure to an inequality telenovela reduces support for redistribution, with the estimated effects ranging from 3% to 5% reduction depending on the categorization method. Taking an ensemble of all three methods, I find a statistically significant 3% reduction in support for redistribution after being exposed to an inequality telenovela. This effect appears to reverse in cases where an inequality telenovela begins airing while another is already on-air. Under these conditions, exposure is associated with a significant increase in support for redistribution. This finding, however, is not robust to changes in the classification threshold, unlike the main finding.

The results suggest that these telenovelas misrepresent the inequality issues in their respective countries, leading individuals to alter their policy preferences based on this distorted depiction. For instance, when these telenovelas portray the disparity between the rich and the poor in a comedic light, they may diminish the perceived urgency of addressing inequality-related challenges. Additional evidence suggests a shift in policy priorities; exposed individuals report placing less importance on inequality and poverty, while focusing more on issues like the economy, violence, or government performance. Individuals with a keen interest in politics, and potentially more informed about economic conditions, experience no significant impact from inequality telenovelas. In contrast, those who identify as lower class demonstrate the most considerable reductions in support, indicating that these portrayals may lead them to perceive systemic inequality as less severe than it is.

Inequality telenovelas are also associated with tangible effects on voting behavior. I merge my telenovela dataset with presidential election data across Latin America and test whether the premier of an inequality telenovela in the months leading to presidential elections affects voter behavior. I find that an additional inequality-focused telenovela is associated with a 0.02 percentage point decrease in voter turnout. Furthermore, this exposure is also associated with a narrower victory margin for the winning candidate and a lower probability of incumbent turnover. These findings suggest that entertainment media can subtly influence political landscapes, potentially by dissuading voters to participate in their civic duties or generating a bias towards the status quo.

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) are the main drivers of 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).

Entertainment media offers individuals a fictional or idealized economic reality and has been shown to shape various political attitudes by "transporting" the viewer into the fictional world and generating an emotional connection (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), HIV attitudes (Banerjee, La Ferrara and Orozco-Olvera, 2019), approval of domestic violence (Banerjee, Ferrara and Orozco, 2019), and early educational outcomes (Kearney and Levine, 2019). Within entertainment media, 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). Outside of economics, the literature has focused on the contents of telenovelas and their relation to inequality (Tufte, 2000; Mayer, 2003; Coppini, Alvarez and Rojas, 2018). In economics, researchers have focused on the effects of telenovelas on demographic changes or 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 support for said community in Latin American countries. The latter study highlights media representation's complex and potentially contradictory effects on attitudes toward specific communities or topics. This paper contributes to the literature on telenovelas and culture by focusing on the subject of inequality and considering both the content and the effects of telenovelas.

This paper also extends to the broader literature seeking to understand how perceptions of inequality influence the demand for redistributive policies (Alesina and Giuliano, 2011; Stantcheva, 2021). Alesina, Stantcheva and Teso (2018) explores the role of upward mobility in shaping perceptions of inequality, shedding light on how individuals' aspirations and opportunities for advancement influence their attitudes toward redistributive policies. Alesina, Miano and Stantcheva (2023) shows evidence that the (mis)perceptions of inequality about the intensity of immigration decrease support for redistributive policies. Domènech-Arumí (2023) argues that exposure to inequality within one's neighborhood can improve the approval of redistribution policies, adding to our understanding of the complex nature of individual experiences and redistributive preferences. Note that this part of the literature focuses on experimental designs. This study highlights the role of entertainment media in shaping individuals' perceptions of their own social class, uses a survey approach rather than experiments, and provides evidence outside of online experiments that (mis)perceptions of inequality are the main drivers of support for redistributive policies.

This paper contributes to the existing literature by shedding light on the role of telenovelas in shaping viewers' preferences for redistributive policies. Latin America has a long history of significant income inequality (Fergusson, Robinson and Torres, 2023) alongside lower levels of redistribution compared to developed economies (Ocampo and Gómez-Arteaga, 2018). Given the ongoing struggles with income distribution in the region, it's vital to grasp the impact of entertainment media on the survival of various redistributive policies through political approval.

A small part of this literature has studied the portrayal of inequality or upward mobility in the media and its impact on perceptions. Diermeier *et al.* (2017) shows that intense coverage of inequality in newspapers worsens individuals' perceptions of social fairness in Germany. Kim (2023) shows that exposure to programs that promote "rags-to-riches" narratives (e.g., American Idol or Shark Tank) reinforces Americans' belief in upward mobility through laboratory and online experiments. In contrast, our study focuses on preferences for redistribution instead of fairness or upward mobility, allowing for a more direct examination of individuals' support for policies addressing inequality.

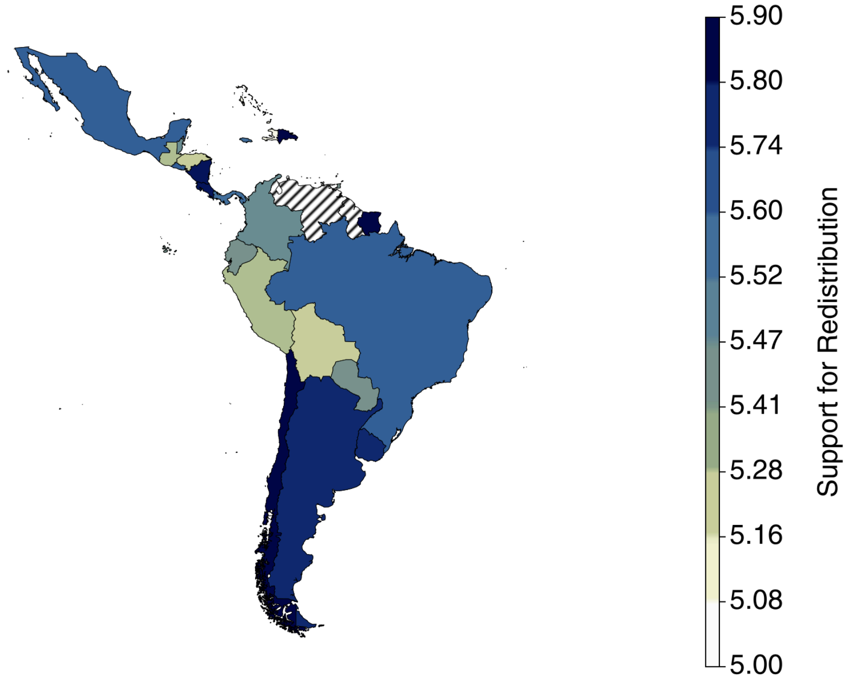
2 Data

2.1 Preference for Redistribution Surveys

Data on preferences for redistribution was sourced from the Latin American Public Opinion Project (LAPOP). The LAPOP has conducted numerous surveys across Latin America, collecting answers to a wide range of socio-political questions. To measure support for redistributive policies, individuals were asked to rank their agreement to the following statement from 1 to 7. *"The (Country) government should implement strong policies to reduce income inequality between the rich and the poor"*.[1] An answer of 1 indicates strong disagreement with the statement and 7 a strong agreement with the statement. This question is available across all survey waves in the sample, spanning from 2006 to 2023. Figure 1 shows the mean value of the outcome question across the survey waves in each country. There is a

[1]Some studies elicit preferences for redistribution using similar a question (Almås *et al.*, 2025). Others measuresupport for redistribution using alternative questions centered on the role of government to reduce inequality. Alesina and Giuliano (2011) uses as their main measure for support for redistribution the following question from the World Value Survey *"Some people think that the government in Washington should do everything to improve the standard of living of all poor Americans (they are at point 1 on this card). Other people think it is not the government's responsibility, and that each person should take care of himself (they are at point 5). Where are you placing yourself in this scale?"* Alesina, Miano and Stantcheva (2023) elicited support for redistribution by asking how much of the government's budget should be destined to social causes like education, and how much should the government tax the top 1% and the bottom 90%.

Figure 1: Average Support for Redistribution in Latin America



Notes: This figure displays the average responses to the support for redistribution question in LAPOP surveys. A score of 7 indicates strong agreement, while a score of 1 represents strong disagreement. For each country, I calculate the average response across the surveyed years, applying the provided population weights.

high level of support for redistribution in Latin America with an overall mean of 5.5 out of 7, with most of the individuals in each country concentrating on high values of support. The countries with the highest approval for redistribution are Chile, Nicaragua, and Costa Rica, while Haiti and Bolivia exhibit the lowest support.

2.2 Telenovelas

Data on telenovelas was collected using various online sources tailored to each country of interest. From each source, we gathered a written synopsis, the air time, the duration (air and end date), the number of episodes, the broadcasting channel, and the genre of the show (e.g. comedy, drama, or melodrama). Additionally, we extracted ratings for each telenovela from the IMDb website, indicating of how well-received each telenovela was by its audience. Telenovelas aired in Chile were gathered from the chilenovelas, a Wiki-like page, which provides comprehensive lists of telenovelas along the genre of telenovelas, ratings, duration, airing times, cast members, and synopses. Similar information for Brazilian telenovelas was sourced from teledramaturgia. For other Latin American countries, lists of telenovelas are available on Wikipedia and user's IMDb lists. Each synopsis was translated into English using OpenAI's gpt-4o-mini model, which provides a good quality translation from Spanish to English and performs reasonably well translating from Portuguese to English (Sanz-Valdivieso and López-Arroyo, 2023; Törnberg, 2023).[2]

A total of 1852 telenovelas were gathered from Chile, Brazil, Mexico, Colombia, Panama, and Argentina, with around 90% concentrated in Chile, Brazil, and Mexico. Table 1 shows some descriptive statistics for all the telenovelas in my sample. Around 34% of telenovelas are dramas or melodramas, around 17% are comedies and the rest are mix of comedy and drama. On average, a telenovela airs around 16:38, consists of 129 episodes, and an IMDB user rating of 6.95 out of 10. The

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

Table 1: Descriptive Statistics of Telenovelas

	mean	sd	min	max
Drama (%)	34.39			
Air time	16:38	07:09	10:00	23:15
N 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, Argentina, Colombian, Panama, and Brazil. TV ratings refer to Nielsen ratings, which measure the percentage of households that, on average, tuned in to watch the show throughout its entire run. IMDb ratings represent user ratings for each telenovela, as listed on the IMDB website.

word count for each synopsis varies significantly, with a mean of around 2,000 words, but with some telenovelas having as few as 66 words and others exceeding 10,000 words.

The discrepancy in synopsis in word length might pose a challenge, as I aim to determine whether these narratives explore themes related to socio-economic differences among their characters. On average, a typical telenovela synopsis is extensive, averaging around 2,000 words, comparable to a couple of paragraphs. This length offers an overview of the plot with sufficient descriptions of the characters and the plot. However, there are notable exceptions within the dataset. Two telenovelas feature summaries that are less than 100 words, providing a brief overview of the plot, which may overlook subtler expressions of inequality. At the other end of the spectrum, the maximum word count for a synopsis reaches approximately 20,000 words, comparable to a small book. Four telenovelas have summaries exceeding 10,000 words, often including detailed descriptions of various seasons or encompass the entirety of the series. While rich in content, they introduce several themes making it challenging to extract a single central theme. Nevertheless, short and extensive synopses only account for less than 1% of all collected telenovelas, with the vast majority centered around the mean length.

2.3 How is inequality displayed in Telenovelas?

Telenovelas can visually show inequality through the choice of sets and wardrobe; audibly through the character's expressions and how they speak; and lastly, through the ideas with the script and dialogue. Pobre Rico, a popular comedy television series aired in Chile in 2012, narrates the story of two babies swapped at birth between a wealthy and a middle-class family, illustrating social class differences.[3] It juxtaposes the challenges a teen gas station worker faces with the lifestyle of the son of one of the most influential families in the country. The first paragraph of the translated synopsis can be found below:

First Paragraph of *Pobre Rico*'s Synopsis
Freddy Pérez and Nicolás Cotapos have led a normal life until now. They are two happy boys and are very comfortable in each of their worlds. Meanwhile, El Rucio, as Freddy is known, lives with his mother Eloísa and his sister Megan in poorer economic conditions. Nicolás lives with his parents Máximo and Virginia, and with his sister Julieta, in a wealthy family. The Cotapos are owners of the Cotapos Holding and Cotapos Airlines, while the Pérez are workers at a gas station.

This telenovela illustrates the economic gap between the Cotapos (rich) and Pérez (poor) families through their sets and script. Figure 2 displays screenshots taken from the first episode of the show.

[3]This telenovela draws inspiration from the classic tale of *The Prince and the Pauper*.

Figure 2: Visual Social Class Differences in *Pobre Rico*

Panel A: *Cotapos* (rich) house

Panel B: *Pérez* ("poor") house

Panel C: *Cotapos* (rich) living room

Panel D: *Pérez* ("poor") living room

Notes: these screenshots were taken from the first episode "Cambiados al Nacer" of Pore Rico available on YouTube

Panels A and C portray the Chilean wealth, characterized by a spacious residence adorned with several windows and a fence, alongside modernly decorated living rooms, ownership of at least two cars, and a maid. The Cotapos family represents the elite, belonging to the top 0.1% of the income distribution; they own one of the largest corporations and have a direct line to the president. In stark contrast, panel B and D present the reality of low-middle-income families, with modest house featuring only two small windows, the absence of cars, and worn-down living spaces. The Pérez family works at a gas station and, while facing their own struggles, live in a fully equipped home, owns a car, and maintains a complete nuclear family.

A common pattern across telenovelas is to show inequality by focusing on the struggle between the top 1% earners and the middle class, as producers often aim to appeal to a mass audience, particularly the working class. *Pobre Rico* does not adequately address the struggles faced by the most impoverished members of society and, instead, focus on the disparities between the wealthy elite and the gas station workers. By omitting the lowest social class, these telenovelas may present a distorted reality of the problem of inequality, which in turn can lead to the (mis)perception of inequality in a society.

3 Measuring Inequality in Telenovelas

To identify telenovelas that portray inequality, I utilize three natural language processing (NLP) techniques to analyze the information about the content of the telenovelas using their synopsis. The synopsis of telenovelas usually contains the main plot points of the story, the primary characters, and the challenges they face, providing a good representation of the content of the first few episodes.[4] Each method will categorize telenovelas into two groups. Inequality telenovelas are those telenovelas that each method predicts its plot contains social class conflict. No-inequality telenovelas are the ones in which social class conflict is not present or is not the central conflict.

Using NLP is essential for several reasons. Human categorization, while valuable, is often unreliable, as it can introduce subjective bias (Grimmer and Stewart, 2013; Evans and Aceves, 2016; Dhar *et al.*, 2021; Rathje *et al.*, 2024). Additionally, manually categorizing a large number of telenovelas is both time-consuming and costly. In contrast, machine-based categorization, though not without its imperfections, offers a more efficient and scalable solution. These techniques have been employed in previous research, and evidence suggests that they correlate with human categorization (Törnberg, 2023; Michalopoulos and Rauh, 2024; Rathje *et al.*, 2024). I employ three specific methods using the telenovela synopses as input: (i) a dictionary approach using related words extracted from ConceptNet, (ii) a vector embedding approach utilizing the Universal Sentence Encoder (USE), and (iii) categorization through ChatGPT, a generative AI capable of classifying with few or even zero examples.

3.1 ConceptNet

The first method is based on a simple idea: the frequency of words such as "poor," "humble," "rich," and "worker" indicates that the telenovela focuses on inequality. I employ a bag of words or dictionary approach to find out which telenovelas use these words the most. I clean the text of each synopsis by removing stop words, numbers, and proper nouns and extract the lemmas for each word[5]

The ConceptNet score is the frequency of the inequality words that appear in each telenovelas synopsis normalized by the text lemma length. Let D represent an inequality dictionary that contains words that represent socio-economic inequality, chosen by the researcher. The $\text{ConceptNet score}_t$ for any telenovela t with a cleaned synopsis of lemma length n_t, and each lemma within that synopsis is w_i where $i = 1, 2, ..., n_t$ is measured by

$$\text{ConceptNet score}_t = \sum_{i}^{n_t} \frac{I(w_i \in D)}{n_t}$$

where $I(w_i \in D)$ is an indicator function taking the value of 1 if the lemma w_i is part of the inequality

Table 2: Related Words to "poor" Using ConceptNet

root	related terms		
	en	es	pt
poor	poor	pobre	pobre
	poorer	pobrete	pobres
	slummy	pobretones	pobretão
	resourceless	pobretón	
		pobres	

[4]Other researches have relied on show synopsis to evaluate the content of the TV shows. Kim (2023) used the summaries to identify rags-to-riches contests shows (e.g. Shark Tank) which ordinary American participate to gain economic benefits through hard work and effort. However, telenovela synopses often provide incomplete plot overviews, as they may omit evolving conflicts and themes, leading to a fragmented understanding of the full story arc, but rather a representation of the starting conflict.

[5]Lemmatization is used to capture different forms of the same word. For example, "break," "breaks," "broke," "broken," and "breaking" are all variations of the lemma "break."

dictionary and 0 otherwise.[6]

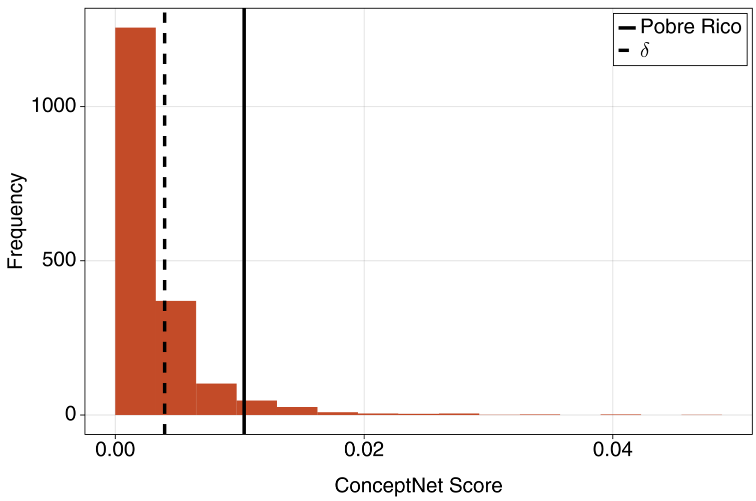
To build the dictionary D, I begin by compiling a list of root words that typically refer to social class conflict, such as "economic," "wealthy," "inequality," and "social class."[7] Using these root words, I then search for related terms in ConceptNet, a knowledge graph of natural languages that show the relations between words in any language. For example, in Table 2 contains the root word "poor" and some of the related terms found in English (EN), Spanish (ES) , and Portuguese (PT). The word "poor" is related to terms are "poor", "poorer," "resourceless," and "slummy," in English, as well as "pobre" and "pobretón" in Spanish which translates to "poor".

Any telenovela with a $\text{ConceptNet score}_t$ higher than a threshold δ will be considered an inequality telenovela. The chosen threshold is equal to 1.5 standard deviations bellow the $\text{ConceptNet score}_{pr}$ of *Pobre Rico*. This will capture more subtler ways in which inequality represented by less frequency of inequality words than *Pobre Rico*.[8]

Figure 3 shows the histogram of the ConceptNet scores across the telenovelas in the sample. The mean score of 0.0027 with a standard deviation of 0.0042. Pobre Rico has a score of 0.011, indicating that the average telenovela in the dataset is about two and a half standard deviations away from Pobre Rico in terms of social class-related content. The chosen threshold δ is set to 0.004, corresponding to 438 inequality telenovelas in all my sample.

To provide context for what these scores represent, consider the telenovela *Chipe Libre*, which has a score of 0. This story revolves around a couple taking a break and exploring new romances, leading to complicated relationships that challenge their views on love and fidelity, with no significant social class conflict depicted. A telenovela with social class conflict in their plot but doesn't have frequent inequality words is *Corazón Rebelde*. This telenovela has a score of 0.006 and tells the story of a group of scholarship students in an elite high school tormented by their wealthy classmates. *Asi en el Barrio Como en el Cielo* is the telenovela with the highest score of 0.039, with a plot paralleling Pobre Rico with a stronger emphasis on the class differences between the two families.

Figure 3: ConceptNet Score Distribution



Notes: This figure shows the ConceptNet scores of the telenovelas in my sample. The ConceptNet score is the frequency of all the words in the built inequality dictionary, normalized by the lemma count of each synopsis. Inequality telenovelas are those with a ConceptNet score $\geq \delta$. The threshold $\delta = 0.004$ was constructed taking 1.5 standard deviations away from the ConceptNet score of Pobre Rico.

[6]Example B.1 provides an example of how the ConceptNet score is calculated.
[7]For the full list of root words and their related terms please see Table Table B.3 and Table B.4 .
[8]I consider different δ's, with 2 and 1 standard deviations aways from *Pobre Rico*, in the robustness tests.

3.2 Vector Embeddings

The second method employs the Universal Sentence Encoder (USE) to assign a vector to each synopsis, enabling the identification of telenovelas that closely resemble Pobre Rico. The USE is a machine learning model developed by Google that effectively transforms sentences and phrases into vector representations (Cer *et al.*, 2018). It is part of a broader family of models designed to convert words, text, and documents into vectors, commonly referred to as vector embeddings. These vector representations play a crucial role in various of natural language processing tasks, including classification and semantic similarity. The algorithm processes text to produce a 512-dimensional vector, which captures the input's context and semantic information. This capability extends beyond simple word matching, a limitation exhibited by the ConceptNet method. Another distinguishing feature of the USE is its comprehensive approach; unlike traditional methods that concentrate on individual words, the USE considers the entire sentence. This allows the model to grasp subtleties and contextual elements more effectively.[9]

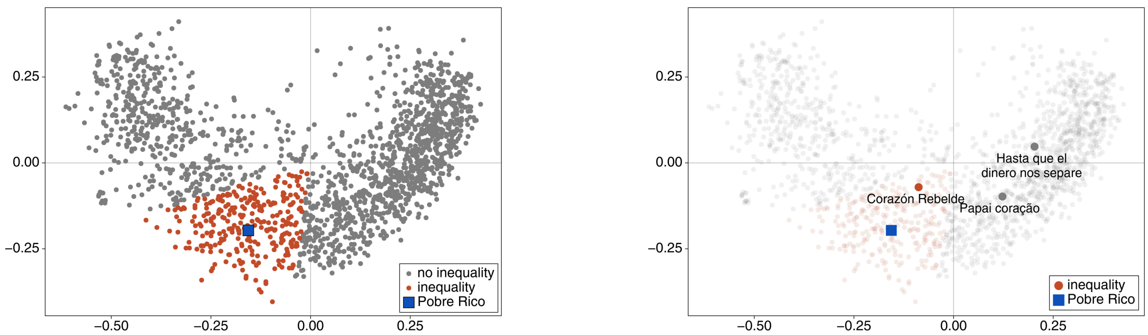
In this study, I apply the USE to the translated synopses of various telenovelas and keep the first 2 principal component using the Kernel PCA algorithm (Schölkopf, Smola and Müller, 1998).[10] Subsequently, I calculate the cosine similarity between the telenovela Pobre Rico and other telenovelas. The cosine similarity, cs, is formally defined as follows

$$cs\left(A, \overrightarrow{\text{Pobre Rico}}\right) = \frac{A \times \overrightarrow{\text{Pobre Rico}}}{\|A\| \times \| \overrightarrow{\text{Pobre Rico}} \|}$$

where A is the vector embedding of the synopsis of a telenovela, $\overrightarrow{\text{Pobre Rico}}$ is the 2-dimentional vector embedding for the Pobre Rico synopsis, and $\|A\|$ is the Euclidean norm of the 2-dimentional vector embedding of A. The cosine similarity ranges from -1 to 1, where the 1 shows that A has the same direction as the $\overrightarrow{\text{Pobre Rico}}$ vector, while a -1 shows that A points to the opposite direction, and a 0 indicates that A is perpendicular to $\overrightarrow{\text{Pobre Rico}}$. Although this metric has a mathematical interpretation, what does it mean for a telenovela to be opposite or perpendicular to *Pobre Rico*?

Figure 4 plots the principal components derived for each telenovela's vector embedding. The telenovela *Hasta que el dinero nos separe* is an opposite vector to *Pobre Rico* with a similarity score

Figure 4: Telenovela Data Latent Space *Pobre Rico*



Panel A: full sample

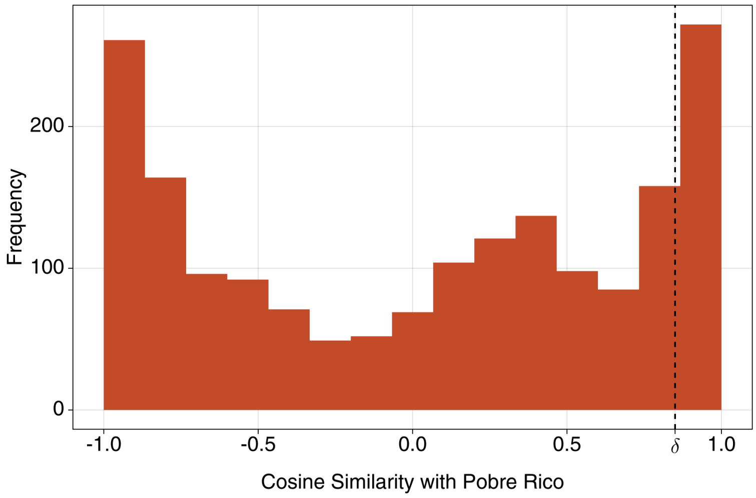
Panel B: select examples

Notes: This figure displays the first two principal components extracted form the vector embeddings of each telenovela and shows those telenovelas that are most similar to *Pobre Rico*. Inequality telenovelas are those with a similarity score $\geq .85$.

[9]For instance, the USE can differentiate between the various meanings of the word "capital" in sentences such as "The capital of Denmark is Copenhagen" and "Denmark will need to invest in capital to get out of the crisis." Other vector embedding models like GloVe or the TF-IDF assign the same vector to "capital" independent of the context (Pennington, Socher and Manning, 2014).

[10]Applying any dimensionality reduction techniques is a common practice in the machine learning literature that improving performance, visualization, and preventing overfitting (Roweis and Saul, 2000; Singh *et al.*, 2022).

Figure 5: Cosine Similarity with *Pobre Rico* Distribution



Notes: This figure displays the histogram of the similarity between a telenovela and the benchmark telenovela *Pobre Rico*. Inequality telenovelas are those with a similarity score $\geq .85$.

of -0.99. This telenovela is focused on the love of a poor merchant with a rich sales manager, and although there is an inequality element, it is not the main theme of the telenovela. Thus, opposite vectors to *Pobre Rico* might represent those stories with a component of inequality, but not at the center stage. An example of a telenovela with a perpendicular vector is *Papai coração*. This telenovela is centered around the struggles of a single father and his relationship with his children. *Corazon Rebelde* stands as a telenovela with vectors in the same direction, detailing the story of a group of poor students in an elite highschool that suffer from the harassment of their rich classmates.

Inequality telenovelas are those with a similarity score exceeding a threshold, delta, set to 0.85. Figure 5 illustrates the histogram of similarity scores, revealing a mean similarity of 0.03 with a standard deviation of 0.7. The distribution in the figure indicates that most telenovelas cluster at both extremes, either extremely similar to *Pobre Rico* or have opposite vectors.

Although *Pobre Rico* is centered around inequality issues, it's synopsis also incorporates law and family components which the vector embeddings also capture. Therefore, this method can mislabel a telenovela focused on inequality but in another setting, like Corazón Rebelde which achieved a high score with the ConceptNet score, but was labeled as no-inequality with vector embeddings with a similarity of 0.52 as its plot revolves around students and not families.

3.3 ChatGPT

Generative AI tools, such as OpenAI's gpt models, have shown remarkable capabilities in categorizing text into groups not present in their training data (Wang, Pang and Lin, 2023).[11] This versatility makes these models appropriate for identifying inequality in telenovelas. 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).[12]

[11]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 3Blue1Borwn's YouTube video "Large Language Models explained briefly", Welch Labs YouTube video The moment we stopped understanding AI and for a more detailed explanation to read Chapter 1.3 and 1.10 of (Jurafsky and Martin, 2024).

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

Example 1 *Chocolate com Pimenta*

ChatGPT CoT: The narrative frequently highlights socio-economic class differences. Ana Francisca is initially portrayed as poor, which leads to her being mistreated and humiliated by others, particularly by the wealthier characters like Olga and Danilo's family

Extract from synopsis: Danilo falls in love with Aninha [Ana Francisca], to the despair of Olga and his family, who reject her because she is poor. With the help of Bárbara, first lady of Ventura and Danilo's aunt, Olga plots to humiliate her rival.

In this project, I use OpenAI's gpt-4o model to classify telenovelas based on whether they portray social class conflict five times.[13] The model reads the translated synopses and outputs the classification result, a confidence score (50-100), and the model's chain of thought. Those telenovelas in which the model labels them as inequality with more than 90% confidence in 3 out of 5 times are considered as inequality telenovelas, and no-inequality telenovelas otherwise.[14]

I look at the provided chain of thought to understand the model's reasoning and performance. Figure 6 presents a word cloud that visually represents the vocabulary utilized in the reasoning during the categorization process. The most prominent words in the cloud, such as "socio," "economic," and "class," indicate that the model appropriately grasps the fundamental aspects of the topic at hand. Other frequently occurring words, including "difference," "topic," "conflict," "theme," "highlight," and "contrast," suggest that the model recognizes the importance of distinguishing between a display of different social classes conflict and conflict

A concern of generative AI models is their potential to hallucinate, make up a story that does not reflect the content in the synopsis. However, requesting the model's chain of thought mitigates this issue (Ashwin, Chhabra and Rao, 2023). In the task of identifying social class conflict, the model's reasoning matches extracts of the synopsis. For example, Example 1 shows the model's reasoning also matching with an excerpt of Chocolate com Pimenta. The fight betwee Ana Francisca and her mother-in-law was because of Ana's lower social class. Another example is Meu Pedacinho de Chão, that tells the story of a devoted teacher who arrives in a small village and challenges the oppressive

Figure 6: ChatGPT Reasoning Word Cloud

Notes: This figure displays the histogram of the similarity between a telenovela and the benchmark telenovela *Pobre Rico*. Inequality telenovelas are those with a similarity score $\geq .85$.

[13]Generative A.I. 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.

[14]The threshold δ for this methods is the confidence level acceptance set to 90%.

Example 2 *Meu Pedacinho de Chão*

ChatGPT CoT: The summary describes characters from different socio-economic backgrounds: the humble townspeople, the wealthy Colonel and his family, and an orphan boy. The disdain of the Colonel for the orphan boy and the conflicts arising are directly linked to social class differences.

Extract from synopsis: Teacher Juliana arrives in the small town of Santa Fé to teach the children and is faced with a humble people who are cowed by the excesses of Colonel Epaminondas Napoleão, an arrogant man who dictates the rules in the region and solves everything by shouting and using weapons.

rule of a powerful colonel while navigating romantic entanglements with both the colonel's son and a loyal farmhand amid rising tensions. The model's reasoning, detailed in Example 2, correctly identifies the conflict between the wealthy, oppressive colonel and the humble townspeople, thus categorizing it as an inequality telenovela.

3.4 Performance

Descriptive statistics for all inequality telenovelas identified by different methods is presented in Table 3. Overall, inequality telenovelas tend to have a slightly higher percentage of dramas, similar episode count, and air time compared to an average telenovela. Additionally, on average, telenovelas centered on inequality appear to achieve higher television ratings than the typical telenovela while maintaining comparable IMDb ratings. Synopsis word count does vary significantly across methods, with ConceptNet showing a similar length to an average telenovela, vector embeddings with less, and ChatGPT with more.

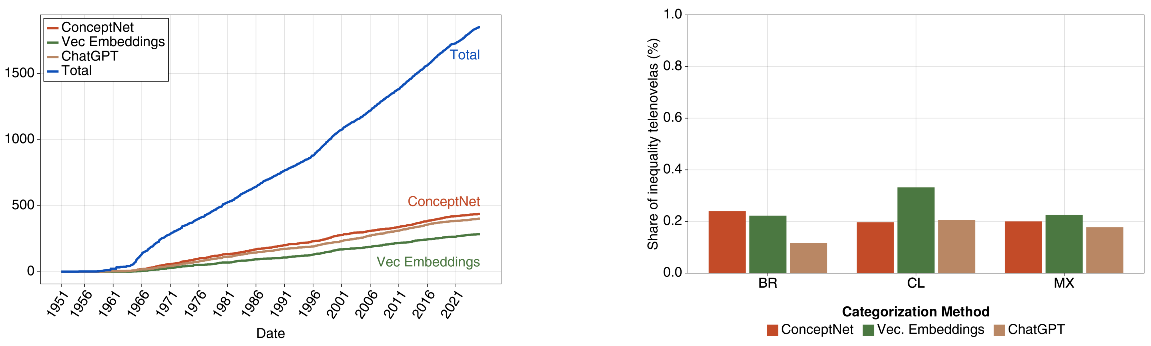
Inequality telenovelas are common in Latin America, but they are not the most frequent type of narrative. Panel A of Figure 7 shows the cumulative time series of all the telenovelas collected in the sample. The cumulative sum of social telenovelas by all categorization methods is well below the cumulative line of all telenovelas. This suggests that social class conflict is a common theme

Table 3: Inequality Telenovelas Descriptive Statistics by Method

Panel A: Full sample (N = 1852)					Panel B: ConceptNet (N = 438)				
	mean	sd	min	max		mean	sd	min	max
Drama (%)	34.39				Drama (%)	32.88			
Air time	16:38	07:09	10:00	23:15	Air time	16:09	07:36	14:00	23:15
N of episodes	129	76	1	1018	N of episodes	135	87	10	807
Synopsis word count	1719	1681	17	20784	Synopsis word count	1421	1081	75	9162
IMDB Rating	6.95	1.12	1.80	9.40	IMDB Rating	6.97	1.18	2.10	9.40
Panel C: Vector Embeddings (N = 284)					Panel D: ChatGPT (N = 402)				
	mean	sd	min	max		mean	sd	min	max
Drama (%)	29.58				Drama (%)	38.56			
Air time	16:58	06:51	16:00	23:10	Air time	17:11	06:34	14:00	23:00
N of episodes	130	75	10	604	N of episodes	140	65	2	459
Synopsis word count	956	271	407	1779	Synopsis word count	2423	2454	78	20784
IMDB Rating	6.95	1.06	3.20	9.40	IMDB Rating	6.98	1.08	2.10	9.10

Notes: This table provides descriptive statistics for all collected telenovelas between 1960-present in Chile, Mexico, Argentina, Colombian, Panama, and Brazil. TV ratings refer to Nielsen ratings, which measure the percentage of households that, on average, tuned in to watch the show throughout its entire run. IMDb ratings represent user ratings for each telenovela, as listed on the IMDB website.

Figure 7: Inequality Telenovelas Across Time and Countries



Panel A: Cumulative over time

Panel B: Across countries

Notes: Panel A shows the cumulative number of collected telenovelas and inequality telenovelas with all three categorization methods across time. Panel B shows the share of inequality telenovelas out of all telenovelas aired in a particular country.

when producing telenovelas in LATAM but is far from being the most common theme. Panel B of Figure 7 shows that inequality telenovelas as a percentage of all telenovelas is around 30% for Chile, Mexico, and Brazil. Thus, individuals in Latin America are not constantly exposed to these type of telenovelas.[15]

3.4.1 Model performance

To evaluate how each method performed in identifying inequality, I compare the model's label prediction to a personal categorization of 220 telenovelas from Chile and Brazil. 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 inequality label. In addition, let tn, fp, fn, bet he true negative, false positive, and false negative respectively. As an illustration, a false positive (fp) is when the model predicts a no-inequality telenovela while I label that telenovela as inequality. The formal definition of each metric is presented bellow.

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

$$\text{recall} = \frac{tp}{tp + fp} \qquad \text{f-score} = 2 \times \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

Accuracy measures the overall proportion of correct predictions, but it may be misleading if one class (inequality or not) dominates. Precision evaluates the quality of positive predictions, indicating how many telenovelas classified as inequality are correct, which is important to minimize false positives. Recall assesses the model's ability to identify all inequality-focused telenovelas, reflecting its sensitivity to true positives. The F-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 a telenovela as inequality when it is not) could bias the estimation of exposure to inequality-themed content.

Table 4 shows the performance metrics for each categorization method. ChatGPT has the best performance in precision and accuracy, which are most important for the empirical estimation. This method only produces false positives in 5% of the cases and incorrectly labels telenovelas in 25%. The

[15]These three countries contain 98% telenovelas in our sample. The remainder 2% are telenovelas from Colombia, Argentina, Peru and Venezuela which were omitted.

Table 4: Inequality Prediction Performance by Method

Categorization Method	Accuracy	Precision	Recall	F-Score
ConceptNet	0.70	0.70	0.41	0.52
Vec Embeddings	0.62	0.53	0.19	0.28
ChatGPT	0.75	0.80	0.48	0.60

Notes: This table shows the model performance based on my own categorization of telenovelas that display inequality aired in Chile. A higher number indicates better performance in all four measures.

other methods, ConceptNet and vector embeddings, have much lower precision and lower accuracy. Overall, the worst-performing model is the vector embeddings across all metrics.

Another way to evaluate the models' performance is by examining the most important words for each category and method. To measure word importance, I follow Kessler (2017), which considers the frequency of each word within specific categories. This method assigns greater importance to words that are frequent in one category but infrequent in another. Table 5 displays the top 10 important words for each category, inequality and no-inequality, according to the categorization method used. Both ConceptNet and ChatGPT highlight expected terms such as "poor," "rich," "neighborhood," "humble," and "social." However, with the vector embeddings model there does not appear to be a clear representation of the most significant words related to inequality. The most important words in the no-inequality category do not reflect inequality, with the exception of a few mentions in the vector embeddings category. Therefore, one could anticipate the impact on preferences for redistribution would be smaller using the vector embedding classifier.

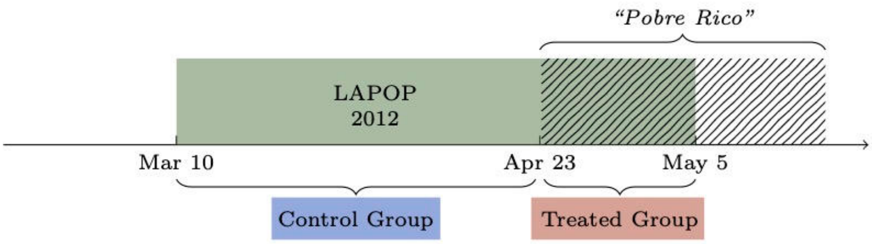
Each of these three methods each has clear advantages and limitations. The dictionary-based approach using ConceptNet offers simplicity and transparency, but it may potentially missed nuanced meanings and it's dependent on the researchers choice of the inequality dictionary. The vector embeddings provide a more sophisticated alternative, capturing contextual meanings and handling synonyms effectively, which allows for better generalization. But its complexity obscures the interpretation of what telenovelas are similar to *Pobre Rico*. In addition, it has the worst performance. Lastly, the ChatGPT method offers high flexibility and the ability to understand context and subtle language features with minimal examples. However, this method is somewhat more opaque than the vector embeddings, making it difficult to interpret its reasoning.

Table 5: Most Important Words by Label and Categorization Method

Inequality			No inequality		
ConceptNet	Vec. Embeddings	ChatGPT	ConceptNet	Vec. Embeddings	ChatGPT
economic	student	poor	people	return	mysterious
social	teacher	class	murder	mansion	passion
poor	case	refuse	travel	plan	past
rich	form	rich	search	receive	search
class	drug	neighborhood	grow	home	heart
employee	doctor	humble	little	go	happy
student	group	social	get	house	story
wealthy	hotel	house	truth	suspect	crime
own	personal	accept	go	promise	secret
neighborhood	obsess	money	find	accept	people

Notes: This figure shows the term (or word) importance for each category following Kessler (2017). Each method has a different definitions for the social category.

Figure 8: *Pobre Rico* Timeline



4 Empirical Strategy

To assess the impact of inequality telenovelas on preferences for redistribution, this study focuses on inequality telenovelas that were aired during the fieldwork of the Latin American Public Opinion Project (LAPOP) survey, ensuring that no other inequality telenovelas were on the air concurrently. This enables a comparison of individuals' responses to the redistribution question before and after the telenovela's debut.

For example, the 2012 LAPOP's Chilean fieldwork spanned from March 10th to May 5th, coinciding with the debut of Pobre Rico on April 23rd. Figure 8 illustrates the timeline of these events. The control group, or not exposed to an inequality telenovela, are those who were interviewed between March 10th and April 22nd. The treatment group, or those exposed to an inequality telenovela, are those who were interviewed between April 23rd and May 5th. Cases in which an inequality telenovela that ends within the same survey fieldwork as another inequality telenovela aired were discarded to avoid any contamination effect from the end of inequality telenovelas. This method provides a clear comparison, ensuring that individuals were not exposed to any other inequality telenovelas, at least since the start of the fieldwork.

Table A.1 presents descriptive statistics for each telenovela alongside the categorization methods employed for their identification in Panel A, and Panel B offers a concise one-line summary for each telenovela. *Os Dias Eram Assim*, *Pobre Rico*, and *Tengo todo excepto a ti* highlight the conflict between social classes, while *Tranquilo Papa* and *La fiscal de hierro* does not seem to have a strong inequality component. This shows the limitation of the vector embeddings method, which categorizes these telenovelas as similar to Pobre Rico in other aspects of the plot. Given this context, one would expect that inequality telenovelas identified by the vector embeddings method would attenuate the effect on support for redistribution than other methods.

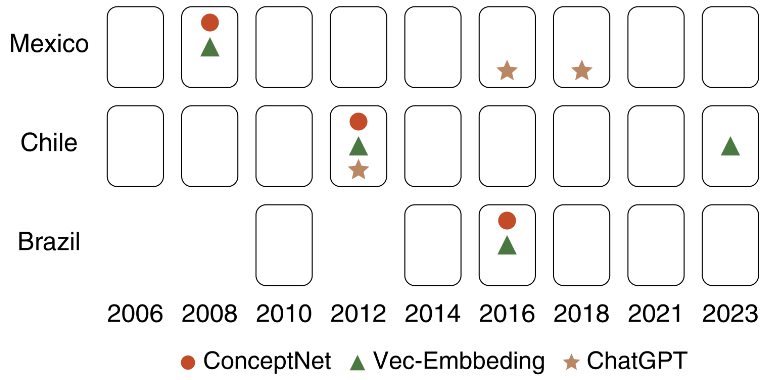
Figure 9 shows the categorization methods for each respective country and wave from the LAPOP survey that fulfills the previous identification criteria. The methods that exhibit the greatest variation across time and countries are ConceptNet and vector embeddings, which utilize 3 to 4 telenovelas in their respective analyses. Notably, the vector embeddings method identifies inequality telenovelas within the same survey wave across different coun- tries. ChatGPT only identified Pobre Rico as an inequality telenovela that meets the estab- lished identification criteria.

Empirical model: To estimate the causal effect of telenovelas that display social class differences and the preferences for redistribution, I run the following OLS regression:

$$Y_{icwt} = \alpha + \gamma \text{InequalityTV}_{cwt} + \beta X_i + \omega d_{t^*} + \delta_{cw} + \varepsilon_{icwt}$$

where Y_{icwt} is the individual answer to the support for redistribution question for individual i at interview date t residing in country c in the survey year wave w. The support for redistribution question is a 1-7 Likert scale question asking individuals how much they agree with the statement "The government should take measures to reduce differences in income levels." The answers are recoded to a 0-1 variable, where 1 indicates that the individual agrees with the statement and 0 oth-

Figure 9: Sources of Variation Across LAPOP Waves



Notes: This figure depicts the variation in the treatment across the LAPOP surveys. Each rectangle represents the fieldwork for a country and year, while the points indicate that at least one telenovela aired that year meets the identification criteria for that categorization method. For a full description of the telenovelas, please refer to Table A.1

erwise. X_i is a set of individual controls which include gender, age, if they live in an urban location, employment status, education, religion, relative income bracket, and marital status. Hausman and Rapson (2017) suggest to include a distance to the treatment date, d_{t^*}, in order to reduce time bias in this type of empirical framework. This control gives more weight to those individuals interviewed closer to the premiere date of the telenovela making for a more precise estimate of the short-term causal effect. δ_{ct} are wave $\times$ country fixed effects. My treatment variables are $\text{InequalityTV}_{cwt}$ that takes a value of 1 if a inequality telenovela is currently on the air at interview date t in country c and takes a value of 0 if there are no inequality telenovelas currently on the air at interview date t in country c.

With the established identification criteria, γ can be interpreted as the short-run effect of the exposure to an inequality telenovela compared to no exposure to inequality in telenovelas on preferences for redistribution.[16] This limitation is due to the LAPOP fieldwork's duration, which averages around two months. Given that inequality telenovelas typically start in the middle-end of the fieldwork, individuals in the treatment group are exposed to these telenovelas for at most one month. As a result, the findings reflect short-run effects on support for redistribution, primarily influenced by the impressions formed during the early episodes. The synopses of the identified telenovelas are well-aligned with the themes presented in the first episodes, ensuring that individuals are exposed the central theme of inequality identified by each categorization method.

The timing of the survey's fieldwork may lead to observable differences in between of individuals interviewed before and after the premiere of the inequality telenovela.[17] Table 6 presents a balance table for each categorization method, with each estimate corresponding to regressions using the set of individual characteristics as the dependent variable and the independent variable is the treatment variable, $\text{InequalityTV}_{cwt}$, with country and wave fixed effects. Individuals interviewed after the premiere of the social telenovela categorized by ConceptNet have a lower income levels, although no statistically significant differences between the two groups. With the vector embeddings method we find similar trends. Lastly, with the ChatGPT method we find that the treatment group

[16] Gulesci, Lombardi and Ramos (2024) uses a similar specification for their analysis. In particular, their treatment variable is the cumulative exposure to telenovelas with characters from the LGBTQ+ community. In contrast, I explore the idea of current exposure or salience effect, since inequality in telenovelas is more common than characters representing minorities.

[17] For instance, in the fieldwork of wave 2012 in Chile, they targeted Santiago, the main capital which houses almost half the population of the country, earlier than other regions in the country.

Table 6: Balance Table by Method

Method	Female	Age	Education	Married	Catholic	Employment	Urban	Income
ConceptNet (N=1636)	−0.008	−0.104	0.035	0.003	−0.005	0.003	−0.008	−0.182
	(0.024)	(0.889)	(0.270)	(0.010)	(0.031)	(0.034)	(0.036)	(0.246)
Vec. Embeddings (N=1071)	−0.006	−0.083	0.035	0.009	−0.010	0.004	0.014	−0.115
	(0.024)	(0.912)	(0.274)	(0.010)	(0.032)	(0.036)	(0.033)	(0.253)
ChatGPT (N=1476)	−0.048	−0.198	−0.025	0.003	−0.078***	−0.017	−0.011	0.065
	(0.033)	(0.850)	(0.336)	(0.009)	(0.028)	(0.019)	(0.052)	(0.251)
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wave Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country × Wave Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: catholics is a dummy for whether the individual is catholic or not. employment status refers to whether the individual is in the working or unemployed. Standard errors are clustered at interview date level.

has a slightly lower share of females, are located in rural areas, have higher income, and are less likely to be catholic. I directly control for these observable characteristics to avoid any bias due to the slight sample imbalance.

One might intuitively expect lower-income individuals would be more likely to support redistribution, suggesting a potential upward bias in my estimates. Similarly, a more rural population tends to be lower income, and thus, one should also expect an upward bias.[18] Considering all these potential biases, our estimates are most likely upwardly biased and the results still suggest that exposure to a social telenovela decreases support for redistribution.

5 Results

Rather than presenting results from each categorization method separately, I report an ensemble estimate that synthesizes the strengths of all three approaches. Each method captures distinct dimensions of inequality in telenovelas and has its own strengths and limitations—ConceptNet focuses on explicit terminology, vector embeddings capture contextual similarity, and ChatGPT leverages nuanced language understanding. By combining these methods, the ensemble approach provides a more robust and comprehensive estimate of the effect, reducing reliance on any single method and mitigating individual weaknesses. The ConceptNet approach focuses on explicit inequality-related terms, vector embeddings capture broader contextual similarities, and ChatGPT leverages nuanced language understanding. To combine their strengths and reduce reliance on any single method, I constructed an ensemble using principal component analysis (PCA) on the validation data, assigning weights to each method's coefficient. The PCA analysis indicated that most of the weight should be placed on ChatGPT (around 0.6), with about 0.3 assigned to ConceptNet and the remaining 0.1 to vector embeddings. This approach provides a more robust and balanced estimate of the effect.

Table 7 shows the ensemble estimates from the main specification. Across all methods and the ensemble, exposure to an inequality telenovela is associated with reduction in support for redistribution. This negative effect remain negative after including country-wave fixed effects, individual controls, and time distance to treatment shock, although the significance of the dissapears in some cases. The ConceptNet method shows a 4.9% decrease in support for redistribution, with higher sample size with other telenovelas and other countries. The vector embeddings we find the smallest and not statistically significant of 4.7% reduction. This is an expected result since this method

[18]There is no intuitive relation between support for redistribution and gender, however when you look at the data there is not a clear difference between mean and women. Empirically, there is is a difference of −0.02 points (se: 0.007, p_value = 0.02) in a scale from 1 to 7 between men and women.

Table 7: Exposure Effect of Inequality Telenovelas by Method

	Support for Redistribution (1-7)							
	ConceptNet		Vector Embeddings		ChatGPT		Ensamble	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exposure to Inequality TV	−0.433***	−0.275**	−0.290***	−0.243	−0.263***	−0.219	−0.332***	−0.245***
	(0.101)	(0.128)	(0.092)	(0.154)	(0.070)	(0.154)	(0.054)	(0.064)
Controls		Yes		Yes		Yes		Yes
Country FE		Yes		Yes		Yes		Yes
Wave FE		Yes		Yes		Yes		Yes
Country × Wave FE		Yes		Yes		Yes		Yes
N	3329	1636	4891	1071	4517	1476	12319	2443
Adjusted R^2	0.035	0.046	0.028	0.038	0.012	0.028	0.010	0.027
Control Mean	5.847	5.871	5.678	5.678	5.710	5.811	5.811	5.821

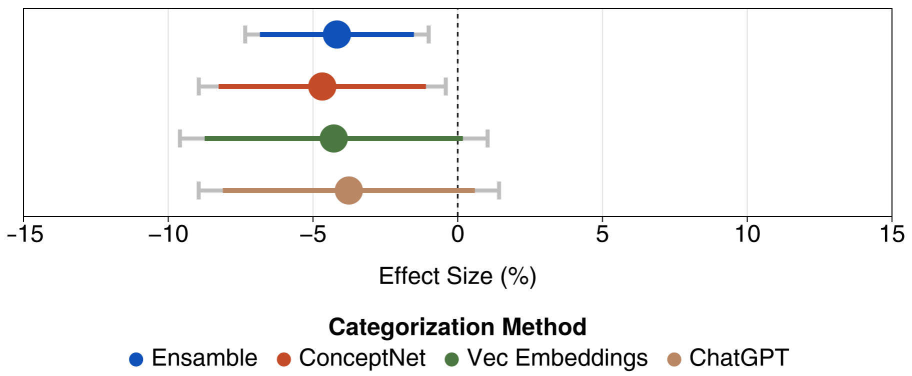
Notes: The outcome variable is the individuals answer to the following statement: "The (Country) government should implement strong policies to reduce income inequality between the rich and the poor." An answer of 7 indicates strong agreement and 1 strong disagreement with the previous statement. The ensemble refers to the weighted mean and standard deviation of the coefficients from the three methods using their PCA weights. The regression controls include gender, age, urban, employment status, education, religion, income, marital status, and a time trend. Clustered standard errors at the interview date level are showed in parenthesis.

misidentified some inequality telenovelas. The ChatGPT method show a 4.3% reduction in preferences for redistribution after being exposed to only Pobre Rico. Taking an ensamble of the three methods, I find a statistically significant 4.2% reduction in support for redistribution after being exposed to an inequality telenovela. This study's findings suggest that exposure to an inequality telenovela reduces individuals' preferences for redistribution in the short run. Although small, these results provide evidence that what is shown in the entertainment media influences policy relevant topics, like support for redistribution.

5.1 Intensive Treatments

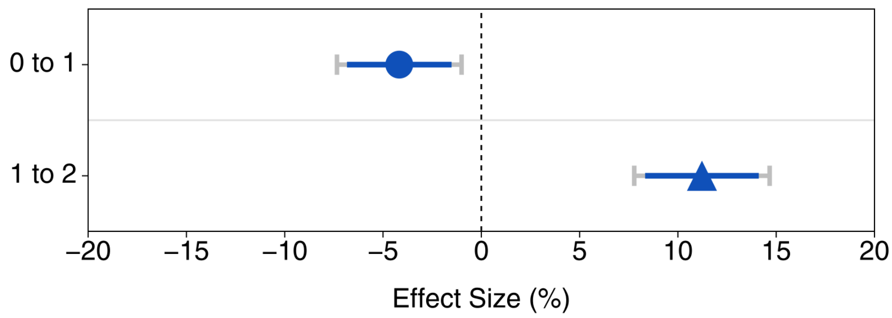
The empirical strategy employed to measure the intensive effects of adding extra inequality telenovelas follows a similar approach described previously. However, I now narrow the focus to cases where N inequality telenovelas are continuously airing throughout the entire fieldwork of the surveys. This restriction eliminates any cases in which one or more telenovelas finish mid-fieldwork.

Figure 10: Exposure Effect Coefficient Plot



Notes: This figure shows the effect size compared to the mean of the group with no exposure to an inequality telenovela. The outcome variable is the individual's answer to the following statement: "The (Country) government should implement strong policies to reduce income inequality between the rich and the poor." An answer of 7 indicates strong agreement and 1 strong disagreement with the previous statement. The ensemble refers to the weighted mean and standard deviation of the coefficients from the three methods using their PCA weights. Controls: gender, age, urban, employment status, education, religion, income, marital status, and a time trend. Standard errors are clustered at interview date level.

Figure 11: Intensive Exposure Effects of an Inequality Telenovela



Notes: This figure shows the effect size relative to the mean of the control group. For the "0 to 1" case, the control group consists of those with no exposure, while for the "1 to 2" case, the control group is composed of those with one exposure, and so on. The outcome variable is the individuals answer to the following statement: "The (Country) government should implement strong policies to reduce income inequality between the rich and the poor." An answer of 7 indicates strong agreement and 1 strong disagreement with the previous statement. Controls: gender, age, urban, employment status, education, religion, income, marital status, and a time trend. Standard errors are clustered at interview date level.

Then, I will analyze the effect of an additional inequality telenovela that debuts during the course of the fieldwork conditional of being exposed to N inequality telenovelas. For example, Figure A.1 shows that during the Mexican LAPOP 2021 survey, *Como Dice el Dicho* was in the air when *Si nos Dejan* premiered on June 1st. This situation stands in contrast to the case of Pobre Rico, which had no other social class-themed telenovelas airing concurrently during its premiere.

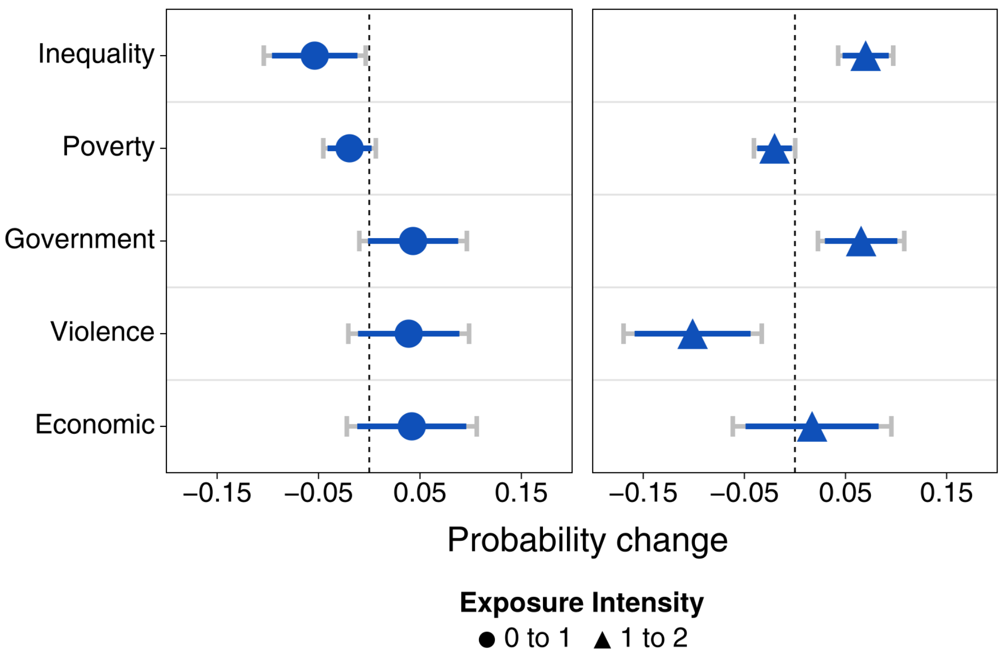
I will estimate a regression using the main specification replacing the treatment with $\text{InequalityTV}^{N}_{cwt}$ that takes a value of 1 if $N+1$ inequality telenovelas are currently on the air at interview date t in country c in the survey years w and takes a value of 0 if there are N inequality telenovelas currently on the air at interview date t in country c that meet the identification cases. Where γ in these cases will be the short-run effect of the exposure to an inequality telenovela compared to being exposed to N inequality telenovelas on preferences for redistribution.

Figure 11 displays the effect of introducing an additional inequality telenovela when individuals are already exposed to N inequality telenovelas. Only three telenovelas met the identification criteria when N = 1, one per method. No inequality telenovelas met the criteria when 2 or more inequality telenovelas were airing concurrently. In contrast to the negative effect observed with no prior exposure, adding an additional inequality telenovela is associated with a positive effect on support for redistribution. Specifically, the ensemble estimate suggests a 10% increase in support for redistribution. The cumulative effect from no exposure to two inequality telenovelas is a 6% increase in support for redistribution. This pattern may indicate that initial exposure to inequality telenovelas reduces support for redistribution, but further exposure in a context where inequality is already salient can actually increase support. However, this effect is not robust to changes in the thresholds used in each method, as documented in Figure 13.

5.2 Mechanisms

The reduction of support for redistribution after being exposed to an inequality telenovela is somewhat counterintuitive. Intuitively, when inequality is featured in media, individuals become more aware of its issues, prompting a shift in their preferences toward supporting redistributive policies and assuming that increased exposure to inequality would naturally lead to a heightened urgency to address it. However, the impact of media representation on public perception relies heavily on how inequality is portrayed. When telenovelas depict inequality as a significant and pressing issue, they have the potential to inform viewers and garner increased support for redistribution as the audience becomes more attuned to the severity of the problem. Conversely, if these shows downplay the challenges of inequality or offer overly simplistic solutions, viewers may absorb this distorted

Figure 12: Exposure Effects of Inequality Telenovelas on National Concerns



Notes: The outcome variables in the y axis is the individual's answer to the following statement: "What is the most important issue facing the country today?" The ensemble refers to the weighted mean and standard deviation of the coefficients from the three methods using their PCA weights. Controls: gender, age, urban, employment status, education, religion, income, marital status, and a time trend. Standard errors are clustered at interview date level.

portrayal. Such (mis)representation can lead individuals to underestimate the severity of inequality, ultimately lowering their priority to supporting redistributive policies. Thus, how inequality is represented in media is crucial in shaping public attitudes toward addressing it.

I will provide two pieces of evidence to support that inequality telenovelas give a distorted (mis)representation of inequality, which leads individuals to deprioritize redistribution policies, when there is no exposure. First, I will show that exposure to inequality telenovelas changes how important individuals perceive inequality as a main national issue. Second, I will explore the heterogeneous effects of exposure to an inequality telenovela, analyzing individual characteristics and political attitudes.

5.2.1 Interest in social problems

Individuals exposed to inequality telenovelas perceive social concerns as less significant at first impact and more significant in subsequent exposure. Figure 12 In fact, the magnitudes of the lower likelihood correspond to the magnitude of the effects on the support for redistribution. At first impact, individuals seems to substitute their social concerns with those about the government, violence, and the economy. This finding is consistent with the idea that inequality telenovelas do not frame inequality as a major problem, but rather as a conflict between the richest and the middle class. For instance, in the telenovela Pobre Rico, inequality is presented in a comedic light, highlighting the absurdities of a wealthy lifestyle from the perspective of a middle-class teenager. In addition, poverty concerns are unchanged after the exposure to inequality telenovelas since inequality in telenovelas focuses on the conflict between the richest and the middle class.

5.2.2 Perceived social class and interest in politics

Table 8 presents regression estimates of the interactions between socioeconomic status and political variables with the exposure to inequality telenovelas. The analysis reveals two key findings present in both first and second exposure. First, individuals' perceptions of their social class matter more than their actual income, as shown in Panel A. Income has little impact on attitudes toward redistribution, while individuals' social class perceptions do. Individuals who perceive themselves as belonging to a low social class are more likely to support redistribution than those who see themselves as middle or upper class. However, exposure to inequality telenovelas further reduces this support among the perceived low social class. The extent of this effect varies significantly: ConceptNet and vector embeddings estimate a 2% decrease in support, while ChatGPT estimates a much larger reduction of 24%. This suggests that telenovelas may provide (mis)information about social disparities, possibly making viewers feel their problems are less severe or leading them to identify with higher social classes, thereby reducing their concern about inequality in society. This finding aligns with existing literature indicating that (mis)perceptions of inequality are significant drivers of support for redistribution, as highlighted in (Stantcheva, 2021).

Second, interest in politics serves as a counterbalance to the (mis)information about inequality telenovelas. Panel B reveals the heterogeneous effects of political leaning and interest in politics on support for redistribution. The estimates suggest that left-leaning individuals are more inclined to advocate for redistribution policies, but there is no statistically significant change after the exposure. However, the estimates indicate that individuals with a high interest in politics show a -0.02% to -5% reduction in their support for redistributive policies after the exposure to inequality telenovelas. This finding implies that those who are more engaged and likely better informed about the actual state of inequality in their country are less affected by external narratives. Consequently, their policy preferences remain unchanged, or at least less affected by the (mis)representation of inequality.

5.3 Robustness

The results presented earlier depend heavily on the ability of the methods to identify inequality telenovelas. To enhance the validity of these findings, I conduct three robustness checks. I start by adjusting the thresholds applied in all three methods to determine whether the results remain consistent. Next, I examine whether telenovelas categorized as definitively not addressing inequality have any impact on support for redistribution. Lastly, I explore changes in other outcomes, such as attitudes toward a minority running for president with the exposure to inequality telenovelas.

5.3.1 Alternative Thresholds (δs)

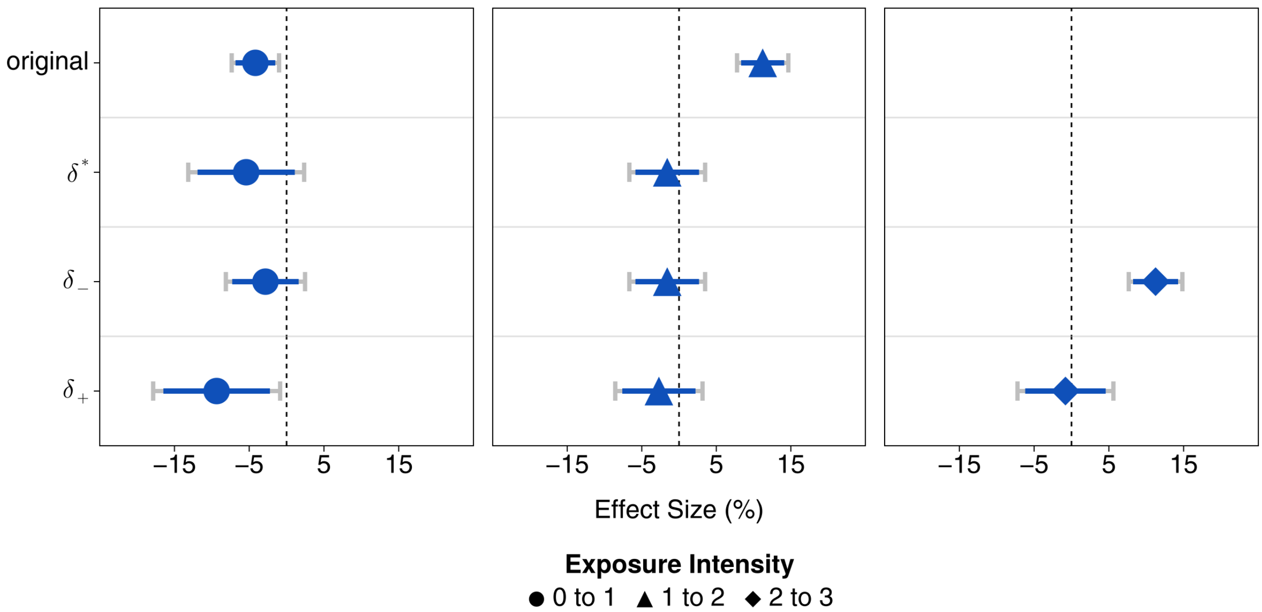
The categorization methods employed all incorporate specific thresholds. For the ConceptNet method, the threshold was set at a score of two standard deviations from the score of *Pobre Rico*. In the vector embeddings approach, inequality telenovelas are those with a similarity score to *Pobre Rico* greater than 0.85. Meanwhile, in the ChatGPT method, only those telenovelas that achieved a confidence level of 90, with the highest agreement among a set of five, were included.

I subsequently modified these thresholds in three ways: first, I optimized the threshold to maximize the precision of each model δ^*; second, I increased the threshold δ^+; and third, I reduced the threshold δ^-. The purpose of optimizing the threshold δ^* is to examine the scenario where false positives are minimized, while the latter two methods aim to assess the sensitivity of the findings to the original thresholds.[19]

Figure 13 compares results using both the original and modified thresholds for all intensive treatments. When $N = 0$, changes in the threshold maintain the negative effect, but with higher standard errors. In particular, increasing the threshold, making it more strict for a telenovela to be classified as inequality, yields a larger 8% decrease in support for redistribution, while reducing the threshold across all methods results in a smaller, non-significant negative effect. For the case of $N = 1$, the result loses its positive effect and significance, suggesting that the intensive effect is not robust and should be interpreted with caution. Lastly, although no telenovelas meet the identification criteria at the original threshold for higher N, some do with increased or decreased thresholds, but

[19] Table A.2 shows the different thresholds for each categorization method.

Figure 13: Exposure Effects to Inequality Telenovela with Alternative Thresholds



Notes: The outcome variable is the individual's answer to the following statement: "The (Country) government should implement strong policies to reduce income inequality between the rich and the poor." An answer of 7 indicates strong agreement and 1 strong disagreement with the previous statement. The original category displays the effect size with my chosen δ. The category δ^* optimizes the thresholds for each category to maximize the precision of each model. The category δ_- and δ_+ decreases and increases the original threshold δ, respectively. Controls: gender, age, urban, employment status, education, religion, income, marital status, and a time trend. Standard errors are clustered at interview date level.

the results are mixed. Overall, this suggests that the effects on redistribution occur primarily when there is no prior exposure to inequality telenovelas, and findings for intensive treatments should be taken with caution.

5.3.2 Placebo Treatments

Another test to enhance the validity of the findings is to examine whether those telenovelas predicted by my methods to exhibit no inequality have any effect on support for redistribution. These telenovelas include, for instance, those with a ConceptNet score of 0, those telenovelas with vector opposite or perpendicular to Pobre Rico, and telenovelas for which ChatGPT indicates no socio-economic class conflict with 90% confidence in at least 3 out of 5 evaluations. Figure 14 splays the effect of these placebo no-inequality telenovelas. As anticipated, the effects are statistically insignificant for both $N = 0$ and $N = 1$ cases. This suggests that the telenovelas identified by these methods provide a reasonably accurate representation of those narratives that do not depict inequality.

Table 8: Heterogeneous Effects of Exposure to an Inequality Telenovela

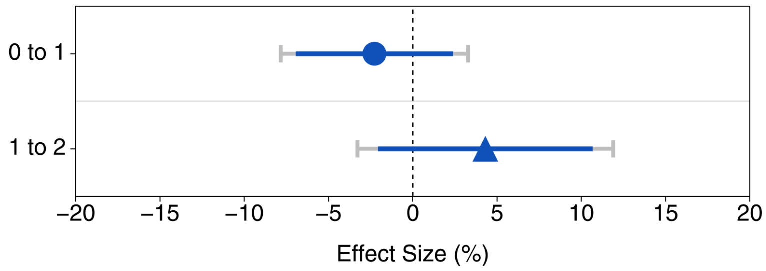
	Support for Redistribution									
	$N=0$					$N=1$				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Exposure to Inequality TV^N	−0.243*	−0.232	−0.374**	−0.299*	−0.238	0.655***	0.622***	1.010***	0.503***	0.525**
	(0.145)	(0.163)	(0.166)	(0.172)	(0.379)	(0.094)	(0.123)	(0.122)	(0.109)	(0.251)
Left Wing		0.118					0.159			
		(0.133)					(0.175)			
Exposure to Inequality TV^N × Left Wing		0.024					0.197			
		(0.199)					(0.158)			
High Pol. Interest			0.030					0.497***		
			(0.099)					(0.191)		
Exposure to Inequality TV^N × High Pol. Interest			0.395**					−0.870***		
			(0.184)					(0.216)		
Income ≤ \$100				−0.181					−0.247	
				(0.170)					(0.191)	
Exposure to Inequality TV^N × Income ≤ \$100				0.145					0.791***	
				(0.231)					(0.225)	
Perc. Low Class					0.414***					0.274*
					(0.135)					(0.163)
Exposure to Inequality TV^N × Perc. Low Class					−0.813**					−0.570
					(0.347)					(0.386)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country × Wave Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Distance to air date $t^{\{(*)\}}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Combined N	2443	2180	2432	2443	1271	2176	1855	2174	2176	908
Average R^2	0.039	0.036	0.045	0.042	0.065	0.029	0.036	0.049	0.031	0.029
Control Mean	5.821	5.756	5.827	5.821	5.887	5.563	5.550	5.563	5.563	5.688

Notes: The outcome variable is the individuals answer to the following statement: "The (Country) government should implement strong policies to reduce income inequality between the rich and the poor." An answer of 7 indicates strong agreement and 1 strong disagreement with the previous statement. I consider respondent with Income < 100 if they earn less than 100 U.S. dollars every month, and those that Perc. Low Class if they classify themselves as low-middle class or low class. Left Wing are those individuals that have a political leaning from 1 (left) to 10 (right) less than 4. High Political Interest refers to those individuals that have at least some interest in politics. Controls: gender, age, urban, employment status, education, religion, income, marital status, and a time trend. Clustered standard errors at the interview date level are showed in parenthesis.

5.3.3 Placebo Other Outcomes

I test whether exposure to inequality telenovelas affects support for other social issues unrelated to economic redistribution. Table 9 presents estimates from the main specification using outcome variables measuring individuals' willingness to support minority candidates for president. As expected,

Figure 14: Exposure Effect of Non-inequality Telenovelas on Support for Redistribution



Notes: This figure shows the effect size relative to the mean of the control group for those telenovelas considered a placebo for each method. A placebo in ConceptNet are those with a score of 0, for Vec. Embeddings is either having a similarity score near 0 or less than −0.85 (opposite), for ChatGPT will be those where the model agree's with confidence that they don't display any inequality. The outcome variable is the individuals answer to the following statement: "The (Country) government should implement strong policies to reduce income inequality between the rich and the poor." An answer of 7 indicates strong agreement and 1 strong disagreement with the previous statement. Controls: gender, age, urban, employment status, education, religion, income, marital status, and a time trend. Standard errors are clustered at interview date level.

exposure to inequality telenovelas shows no impact on these unrelated social issues—all coefficients remain small, centered around zero, and statistically insignificant across all levels of prior telenovela exposure.

These placebo tests provide additional validation that the categorization methods successfully isolate telenovelas with inequality as their central theme, rather than capturing broader social content. This specificity is important because previous research has shown that telenovela content can influence attitudes toward the social groups they portray. For example, Gulesci, Lombardi and Ramos (2024) found that exposure to telenovelas featuring LGBTQ+ characters decreases support for candidates from that community in Latin America. In contrast, my results show no effect of inequality telenovelas on support for minority presidential candidates, confirming that these shows specifically address economic themes rather than broader social representation issues.

6 Effect on Voting Behavior

This section examines whether the influence of inequality telenovelas extends beyond shaping attitudes toward specific policy issues, and instead has broader effects on political behavior. The portrayal of inequality in telenovelas introduces a politically charged narrative into popular media,

Table 9: Effects of Inequality Telenovela on Placebo Outcomes

	$N = 0$			$N = 1$		
	Approval	Running President		Approval	Running President	
	LGBTQ+	Women	LGBTQ+	LGBTQ+	Women	LGBTQ+
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure to Inequality TV^N	0.004	0.026	0.011	−0.065	−0.125	−0.052
	(0.044)	(0.036)	(0.036)	(0.041)	(0.150)	(0.048)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country × Wave Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Distance to air date t^*	Yes	Yes	Yes	Yes	Yes	Yes
Combined N	3348	4100	3269	2211	2211	925
Average R^2	0.092	0.081	0.036	0.129	0.088	0.046
Control Mean	0.508	0.643	0.225	0.511	0.591	0.262

Notes: The controls include gender, age, urban, employment status, education, religion, income, marital status, and a time trend. Clustered standard errors at the interview date level are showed in parenthesis.

which may impact how viewers engage with electoral processes. Potential mechanisms include increased salience of economic disparities, shifts in perceptions of fairness or legitimacy, and changes in how voters assess candidates or political parties. While these pathways are plausible, establishing a direct causal link between media exposure and voter behavior requires further research. Here, I provide empirical evidence on the association between exposure to inequality telenovelas and electoral outcomes, highlighting the need for future studies to unpack the underlying mechanisms.

To investigate whether exposure to inequality telenovelas influences electoral outcomes, I merged presidential election data from the National Elections Database (Marx, Pons and Rollet, 2025) with my telenovela dataset for Latin American countries. For each election, I calculated the number of inequality telenovelas airing in the 1, 3, and 6 months preceding election day.

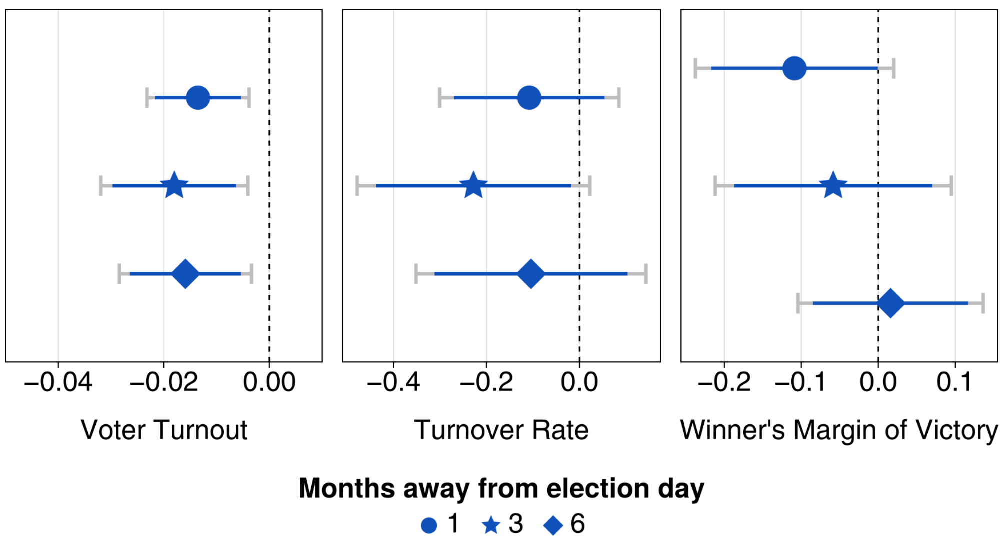
The estimated regression model is as follows:

$$Y_{ct} = \alpha + \gamma \text{IneTV}_{c,t} + \beta X_c + \delta_c + \varepsilon_{c,t}$$

where Y_{ct} is the electoral outcome for country c at time t, which can be total votes cast, vote share of the winning candidate, or the victory margin. The treatment variable, IneqTV_{ct}, indicates if there was an inequality telenovela airing in the months leading up to the election. X_c includes country-level controls such as GDP per capita and population size, and δ_c are country fixed effects to account for country differences. The coeficient γ will then measure the correlation between an electoral outcome and if there was an inequality telenovela airing in the months leading up to the election.

Figure 15 show the estimated coeficients for voter turnout, turnover rate, and the winner's margin of victory. The regressions reveal that the presence of a inequality telenovelas in the months leading up to elections has a correlation with electoral outcomes. An additional inequality telenovela airing in the months before an election is associated with a 0.02 percentage point decrease in voter turnout. This is consistent with the idea that individuals put less weight social concerns and those on the margin of voting are less likely to turn out. Exposure to inequality telenovelas leading up to election day is also associated with a narrower victory margin for the elected president and a reduction in incumbent turnover, suggesting that these shows may subtly influence electoral competitiveness and the likelihood of political change.

Figure 15: Exposure Effect and Voting Behavior



Notes: Voter Turnout is the number of voters divided by the voting-age population. Turnover rate is the likelihood of an incumbent being replaced. Winner's margin of victory refers to the difference in vote share between the winning candidate and the runner-up. Controls: GDP per capita, population size, and other macroeconomic variables for that country and year of the election. Clustered standard errors at the interview date level are showed in parenthesis.

7 Conclusion

This paper demonstrates that the fictional narratives of entertainment media can have tangible consequences for economic policy preferences and political outcomes. Using a novel dataset of Latin American telenovelas and a quasi-experimental design, I show that exposure to inequality-themed storylines paradoxically reduces public support for redistribution and is correlated with decreased voter turnout and tighter electoral competition. The proposed mechanism is that these portrayals distort viewers' understanding of economic realities. This effect is particularly pronounced among lower-class individuals and varies by narrative style. For instance, comedic representations appear to reduce the perceived severity of the issue. Methodological checks confirm the robustness of these findings, which are specific to inequality content. These findings contribute to our understanding of how culture shapes economic and political behavior, highlighting that popular media acts not just as a mirror of societal norms, but as a powerful force that can actively mold them. The stories we consume on screen can change minds, and in doing so, they can change the world outside the screen as well.

Overall, this research underscores the influential role of entertainment media in shaping public perceptions and policy preferences. It highlights the necessity for a deeper understanding of how narratives in popular culture can sway societal attitudes toward inequality and redistribution. Future research should concentrate on developing improved methods for analyzing the content of telenovelas. The behavioral economics framing literature (Tversky and Kahneman, 1981) suggests that how information is presented, whether in a positive or negative light, affects individual behavior differently. Consequently, discerning whether inequality is portrayed negatively or positively is crucial for understanding its impact on public attitudes.

Bibliography

Alesina, A. and Giuliano, P. (2011) "Preferences for Redistribution," *Handbook of Social Economics*. Edited by J. Benhabib, A. Bisin, and M.O. Jackson. North-Holland. Available at: https://doi.org/10.1016/B978-0-444-53187-2.00004-8.

Alesina, A., Miano, A. and Stantcheva, S. (2023) "Immigration and Redistribution," *The Review of Economic Studies*, 90(1), pp. 1–39. Available at: https://doi.org/10.1093/restud/rdac011.

Alesina, A., Stantcheva, S. and Teso, E. (2018) "Intergenerational Mobility and Preferences for Redistribution," *American Economic Review*, 108(2), pp. 521–554. Available at: https://doi.org/10.1257/aer.20162015.

Almås, I. *et al.* (2025) *Fairness Across the World*. Available at: https://econpapers.repec.org/RePEc:hhs:nhheco:2025_006 (Accessed: April 9, 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.

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.

Ashwin, J., Chhabra, A. and Rao, V. (2023) *Using Large Language Models for Qualitative Analysis can Introduce Serious Bias*. arXiv. Available at: https://doi.org/10.48550/arXiv.2309.17147.

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.

Cer, D. *et al.* (2018) *Universal Sentence Encoder*. arXiv. Available at: https://doi.org/10.48550/arXiv.1803.11175.

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).

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).

Coppini, D., Alvarez, G. and Rojas, H. (2018) “Entertainment, News, and Income Inequality: How Colombian Media Shape Perceptions of Income Inequality and Why It Matters,” *International Journal of Communication*, 12(0), p. 24. Available at: https://ijoc.org/index.php/ijoc/article/view/6205 (Accessed: December 31, 2024).

Dhar, A. *et al.* (2021) “Text categorization: past and present,” *Artificial Intelligence Review*, 54(4), pp. 3007–3054. Available at: https://doi.org/10.1007/s10462-020-09919-1.

Diermeier, M. *et al.* (2017) *Impact of inequality-related media coverage on the concerns of the citizens*. Available at: https://www.econstor.eu/handle/10419/226497 (Accessed: March 25, 2024).

Domènech-Arumí, G. (2023) *Neighborhoods, Perceived Inequality, and Preferences for Redistribution: Evidence from Barcelona*. Rochester, NY. Available at: https://doi.org/10.2139/ssrn.3863594.

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.

Evans, J.A. and Aceves, P. (2016) “Machine Translation: Mining Text for Social Theory,” *Annual Review of Sociology*, 42(Volume42, 2016), pp. 21–50. Available at: https://doi.org/10.1146/annurev-soc-081715-074206.

Fergusson, L., Robinson, J.A. and Torres, S. (2023) “The Interaction of Economic and Political Inequality in Latin America.,” *Latin America and Caribbean Inequality Review* [Preprint], (133). Available at: https://eprints.lse.ac.uk/122083/1/WP_133.pdf.

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.

Garcia-Jimeno, C. and Parsa, S. (2024) *Cultural Change through Writing Style: Gendered Pronoun Use in the Economics Profession*. Available at: https://github.com/parsahar/parsahar.github.io/blob/main/papers/diffusion_genderpronouns.pdf.

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.

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.

Grimmer, J. and Stewart, B.M. (2013) “Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts,” *Political Analysis*, 21(3), pp. 267–297. Available at: https://www.jstor.org/stable/24572662 (Accessed: June 6, 2024).

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.

Hausman, C. and Rapson, D.S. (2017) *Regression Discontinuity in Time: Considerations for Empirical Applications*. *Working Paper Series*, National Bureau of Economic Research. Available at: https://doi.org/10.3386/w23602.

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.

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.

Kessler, J.S. (2017) “Scattertext: a Browser-Based Tool for Visualizing how Corpora Differ,” *Association for Computational Linguistics* [Preprint]. Available at: https://arxiv.org/abs/1703.00565.

Kim, E. (2023) “Entertaining Beliefs in Economic Mobility,” *American Journal of Political Science*, 67(1), pp. 39–54. Available at: https://doi.org/10.1111/ajps.12702.

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

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.

Marx, B., Pons, V. and Rollet, V. (2025) *National Elections Dataset*. Available at: http://nationalelectionsdatabase.com/.

Mayer, V. (2003) “Living Telenovelas/Telenovelizing Life: Mexican American Girls' Identities and Transnational Telenovelas,” *Journal of Communication*, 53(3), pp. 479–495. Available at: https://doi.org/10.1111/j.1460-2466.2003.tb02603.x.

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

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.

Ocampo, J.A. and Gómez-Arteaga, N. (2018) “Social protection systems, redistribution and growth in Latin America,” *CEPAL Review*, 2017(122), pp. 7–30. Available at: https://doi.org/10.18356/6eb4de51-en.

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.globalmediajour nal.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).

Pennington, J., Socher, R. and Manning, C. (2014) “Glove: Global Vectors for Word Representation,” in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: Association for Computational Linguistics, pp. 1532–1543. Available at: https://doi.org/10.3115/v1/D14-1162.

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.

Roweis, S.T. and Saul, L.K. (2000) “Nonlinear Dimensionality Reduction by Locally Linear Embedding,” *Science*, 290(5500), pp. 2323–2326. Available at: https://doi.org/10.1126/science.290.5500.2323.

Salewski, L. *et al.* (2023) “In-Context Impersonation Reveals Large Language Models' Strengths and Biases,” in. Available at: https://openreview.net/forum?id=CbsJ53LdKc (Accessed: December 11, 2024).

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.

Schölkopf, B., Smola, A. and Müller, K.-R. (1998) “Nonlinear Component Analysis as a Kernel Eigenvalue Problem,” *Neural Computation*, 10(5), pp. 1299–1319. Available at: https://doi.org/10.1162/089976698300017467.

Singh, K.N. *et al.* (2022) “A novel approach for dimension reduction using word embedding: An enhanced text classification approach,” *International Journal of Information Management Data Insights*, 2(1), p. 100061. Available at: https://doi.org/10.1016/j.jjimei.2022.100061.

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.

Stantcheva, S. (2021) *Perceptions and Preferences for Redistribution*. *Working Paper Series*, National Bureau of Economic Research. Available at: https://doi.org/10.3386/w29370.

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.

Tufte, T. (2000) *Living with the Rubbish Queen: Telenovelas, Culture and Modernity in Brazil*. Indiana University Press.

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.

Table A.1: Description of Telenovelas that match the identification from 0 to 1

Panel A: Telenovela Statistics

Title	Country	Air Date	Episodes	Genre	ConceptNet	VecEmbeddings	ChatGPT
Os Dias EramAssim	brazil	2017-04-17	88	drama	x	x	
Pobre Rico	chile	2012-04-23	227	comedy	x	x	x
Tengo todo excepto a ti	mexico	2008-02-11	50	comedy-drama	x	x	
Un poquito tuyo	mexico	2019-02-25	79	comedy			x

Panel B: One-sentence Summary

Os Dias Eram Assim	*In 1970s Brazil, a young idealistic doctor, Renato, falls in love with a wealthy girl, Alice, only to be separated by political turmoil and deceit, leading to years apart until their love rekindles upon an unexpected reunion.*
Pobre Rico	*Freddy and Nicolás, teenagers from different socioeconomic backgrounds, discover they were swapped at birth, leading to a legal mandate to switch lives with their biological families, as they navigate complex family dynamics and the pursuit of acceptance.*
Tengo todo excepto a ti	*Years after parting ways due to economic hardships, former lovers Carlos and Rebeca, now both married with competing families and magazines, reignite their past connection, leading to conflicts and suspicions of infidelity that unravel both their households.*
Un poquito tuyo	*A successful family man, Antonio, realizes the detrimental effects of spoiling his family and decides they must fend for themselves while becoming entangled with Julieta, a woman with her own dramatic past.*

Notes: These telenovelas meet the identification criteria, provided that no other inequality telenovelas currently airing when these telenovelas started during the same fieldwork period. These telenovelas are used to measure the impact of increasing exposure from zero to one telenovelas that display inequality

8 Appendix

A. Additional Tables and Figures

B. NLP Method Details

B.1 ConceptNet

Tables Table B.3 and Table B.4 present the root words selected for the dictionary and the corresponding related terms found in English (EN), Spanish (ES), and Portuguese (PT), utilizing the ConceptNet database. To access the database I sent requests to ConceptNet API, where you can search for the word "poor" and their related words in Spanish by visiting http://api.conceptnet.io/c/en/poor?filter=/c/es using the following syntax

```
http://api.conceptnet.io/c/ROOT_LANG/WORD?filter=/c/TARGET_LANG
```

Figure A.1: *Si nos Dejan* timeline

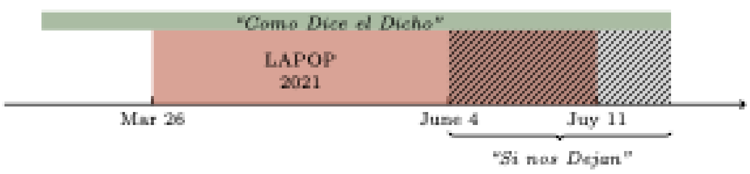


Table A.2: Thresholdss by Method

Categorization Method	**original**	δ^*	δ_-	δ_+
ConceptNet	0.900	0.927	0.850	0.950
Vec Embeddings	0.0040	0.0104	0.0018	0.0061
ChatGPT	0.850	0.931	0.750	0.950

Where ROOT_LANG is the language of the root word, WORD is the root word itself, and TARGET_LANG is the language you want to get related words from. As an example, for the word "poor" in Spanish, I used the URL `http://api.conceptnet.io/c/en/poor?filter=/c/es`. On the website you can find a json file structure and under the key related you get the list of all related words for poor in Spanish with a certain relationship weight, which I chose to be ≥ 0.5.

The chosen root words embody themes of social class conflict depicted in popular telenovelas such as *Pobre Rico* and *Rebelde*. These shows vividly portray the stark inequalities present in various social classes. For example, from *Pobre Rico*, I extracted terms including "poor," "wealthy," "rich," "owner," and "worker," all of which highlight the dynamics of social disparity. Similarly, words like "social class," "elite," and "distinguished" were sourced from *Rebelde* further illustrating the narratives of class distinction and conflict. Most of the Spanish and Portuguese terms identified through ConceptNet are synonyms of their English counterparts, mirroring the translated root words. For instance, "pobre" in both Spanish and Portuguese serves as the literal translation of the word "poor." Additionally, other terms in Spanish, such as "pobrete" or "pobreton," represent alternate expressions of "pobre."

Example B.1 illustrates the ConceptNet score in action. Suppose that the chosen inequality dictionarty $D = \{$poor, poorer, slummy, resourceless$\}$, and take a sentence from the synopsis of *Pobre Rico* and *Mama Mechona*, a telenovela about a mother and son going to college at the same time which does not display social class differences. The telenovela *Pobre Rico* then gets a score of $\frac{1}{6}$, since only one word from the dictionary (the word poor) matches among 6 total unique words. The telenovela *Mama Mechona* will get a score of zero since neither "poor", "poorer," "resourceless," nor "slummy" appear in it's synopsis among 10 unique words.

Example B: ConceptNet Score	
Pobre Rico score = $\frac{1}{6}$	*Mama Mechona* score = $\frac{0}{10}$
Freddy lives in **poorer** economic conditions.	Macarena, the mother of a freshman, is also a student.

B.2 Vector Embeddings

Vector embeddings are a type of representation for words in a continuous vector space, where each word is mapped to a unique vector of real numbers. This technique helps capture the semantic meanings of words based on their contexts in which they appear. Essentially, words that have similar meanings or are used in similar ways are positioned closer together in this multidimensional space.

Common techniques for generating word embeddings include Word2Vec, GloVe, Transformers, and even large language models like OpenAI's ChatGPT (Jurafsky and Martin, 2024). These methods employ large datasets containing text to learn the relationships between words based on co-occurrences and context. Once generated, these embeddings can be used in various natural language processing tasks, such as sentiment analysis, translation, and information retrieval, enhancing machines' understanding of human language.

Take for example Figure B.2 where the words "apple," "tree," "grape," and "vine" are encoded in a 2 dimensional space. Vector embeddings are designed to maintain some semantic relationship between words. In this case, it aims to maintain the following relationship: "a apple grows from a tree, and a grape grows from a vine." Taking the relationship GROWS by substracting tree from apple, and adding grape we can get vine. Another common example is the relationship between the words "king," "male", "queen," and "female", with vectors queen $\approx$ king $-$ male $+$ female.

Table B.3: ConceptNet Inequality Terms List

root		related terms		
		EN	ES	PT
poor	→	poor	pobre	pobre
		poorer	pobrete	pobres
		slummy	pobretones	pobretão
		resourceless	pobretón	
			pobres	
inequality	→	inequality	desigualdad	desigualdade
		inequity	inequidad	inequação
		disparity	inecuaciones	eqüidade
		unequal	disparidad	igualdade
		gini coefficient	igualdad	disparidade
economic	→	economic	económicas	econômica
		noneconomic	economía	economia
		economy	fisiócratas	ecónomo
		macroeconomic	macroeconomico	macroeconómica
		socio economic	socioeconómico	financeiros
rich	→	rich	ricas	rica
		superrich	rico	rico
		richen	opulento	opulento
		richer	pudiente	
wealthy	→	wealthy	adinerado	ricas
		inferior	acaudalada	rico
			acaudalado	ricos
			pudiente	
			adinerada	

Vector embeddings have also been used to keep relationship with other objects in economics. Garcia-Jimeno and Parsa (2024) uses the academic network of economists and assigned them a vector to each theorist to find who are the other economists are most likely to be their co-author. Another example is Hansen uses a big data set of firm characteristics and input output structures to assign a vector to each firm. This exercise surprisingly captures the geographical distances of each firm without having that as an input.

B.3 ChatGPT

The prompt used for this categorization exercise was:

Prompt 1: Inequality Extraction Prompt

```
You are a helpful research assistant. Classify each TV show summary if the plot is
centered around socio-economic class differences. Please show your reasoning step by
step of the categorization. Also provide a confidence level where 50 (uncertain) and 100
(perfectly certain).
```

Table B.4: ConceptNet Inequality Terms List Continued

root		related terms		
		EN	**ES**	**PT**
owner	→	owner	propietarias	proprietária
		proprietor	dueño	dono
		possessor	empresaria	proprietario
		landlord	poseedor	possuidor
		owns	amo	amo
worker	→	worker	cuadrillero	operário
		laborer	laburante	obrador
		employee	trabajadora	trabalhador
		farmworker	obrero	assalariado
		manual_laborer	empleado	empregado
elite	→	elite	elite	elite
		élite	élite	escol
		elitist	elitista	nata
social class	→	social	social	social
		sociability	macrosocial	sociabilidade
		societal	ecosocial	socialização
		social_network	socialista	socialismo
		socialist	socialismo	sociabilizar
distinguished	→	distinguished	distinguido	distinta
		illustrious	egregio	assinalada
		esteemed	prestigioso	eminente
		renowned	destacado	notável
		prestigious	ilustre	marcante

Figure B.2: ConceptNet Inequality Terms List

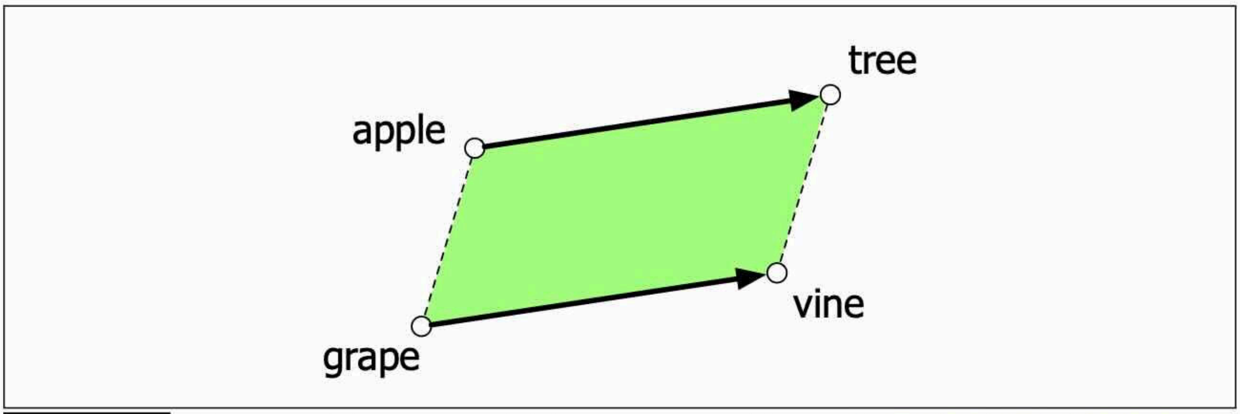


Figure 6.15 The parallelogram model for analogy problems (Rumelhart and Abrahamson, 1973): the location of $\overrightarrow{\text{vine}}$ can be found by subtracting $\overrightarrow{\text{apple}}$ from $\overrightarrow{\text{tree}}$ and adding $\overrightarrow{\text{grape}}$.

The prompt can be divided into three distinct parts that each serve a specific purpose for maximizing the model's efficacy.

First, the instruction, "You are a helpful research assistant," while seemingly lacking contextual depth, is significant in enhancing the performance of generative AI. Research conducted by Salewski *et al.* (2023) indicates that assigning a particular role to the model enables it to concentrate on specific segments of its knowledge base while adopting an expert perspective on the topic at hand. This role-based approach helps align the AI's responses with the desired format and depth of information. Second, my request for the model to explore socio-economic class differences, rather than just issues of inequality, is intentional. My aim is to highlight how characters from distinct economic backgrounds are portrayed in conflict with one another, thereby reflecting the nuances of inequality that arise from these disparities. This focus ensures that the theme of conflict stemming from economic differences remains central to the narrative at least at the beginning of each episode of the telenovelas under analysis. Third, I request that the model provide a confidence level alongside its reasoning, employing a chain of thought (COT) approach. Although this method is informal, the inclusion of confidence levels allows me to gauge the model's certainty regarding whether it identifies inequality or not. This mechanism is helpful in filtering out cases where the model expresses uncertainty. While a formal definition of this confidence could be derived from the log-probabilities of the categorization, implementing such a framework has proven challenging. Nevertheless, encouraging the model to articulate its chain of thought can enhance its performance and mitigate the incidence of hallucinations, leading to more accurate and reliable outputs (Ashwin, Chhabra and Rao, 2023).

Generative AI models are inherently stochastic, which means they can produce different responses to the same prompt, even when the temperature setting is fixed at 0. This unpredictability raises interesting questions about the underlying mechanisms of these models. Although experts have not fully pinpointed the reasons for this behavior, two primary factors are often considered the most likely culprits Chann (2023).

First, computational floating-point inaccuracies in GPUs could contribute to the variability in responses. These inaccuracies arise from the way numerical calculations are performed on GPUs, which can lead to slight discrepancies in the results. Even minimal differences in calculations can propagate through the model's computations, resulting in varied outputs. Second, randomness in the tie-breaking rules for words with equal probabilities is another potential source of this stochastic behavior. When the model encounters multiple words that share the same probability for selection, it may rely on randomization techniques to make a decision. This randomness can result in different words being chosen each time, contributing to the overall variability of the responses.

To mitigate the randomness inherent in generative AI models, or even better, to harness it effectively, I maintain OpenAI's default temperature setting. This choice allows the model to retain a level of creativity in its analyses, which is essential for nuanced interpretations. To further structure the results and enhance reliability, I ask the same prompt five times. By doing so, I can identify the most agreed-upon categorization from the responses provided. This approach not only enables me to aggregate insights from multiple runs but also ensures that the findings are robust and reflective of a consensus among the varied outputs. Additionally, I keep detailed records of each iteration's results, allowing others to verify the outcomes. This transparency fosters trust in the analytical process and enables further scrutiny or replication of the findings.