

Dynamic Models for Bilingualism, Language Adoption, and Language Dynamics

Riccardo Del Gratta¹

¹CNR Institute for Computational Linguistics, Pisa, Italy
riccardo.delgratta@ilc.cnr.it

Draft

Contents

1 Primary purpose of the paper and its structure	5
2 Introduction	6
3 Input Data	6
4 Language dynamics: models	7
4.1 Basic Abram-Strogatz model	8
4.1.1 Some comments on prestige (s) and volatility (a)	8
4.2 Running the Abrams-Strogatz model on experimental data	9
4.2.1 Data preparation	9
4.2.2 Experiments	9
4.2.3 Experiments: comments and insights on optimal and initial parameters	14
4.2.4 Experiments: comments and insights on RMSEs	14
4.2.5 Experiments: model comparison and external factors	15
4.2.6 Fixed points and stability	15
4.3 Abrams-Strogatz Model with a varying population	22
4.3.1 The Bilingual group: two distinct typologies of increments	23
4.4 Modeling the likelihood of a natural increase in the Bilingual group.	24
4.5 Fixed points and stability	25
4.6 Comparison of fixed points	27
4.7 XXX	29
A A comprehensive discussion on the Abrams-Strogatz Model	32
A.1 Fixed points of Abrams-Strogatz model	32
A.2 Stability Analysis of the Abrams-Strogatz Model	33
A.2.1 Jacobian approach	33
A.2.2 Reduced Ordinary Differential Equation (ODE)	37
A.2.3 An alternative approach to stability Analysis of fixed points	39
A.2.4 Evaluation of stability	44
A.3 Special cases	45
A.3.1 The Abrams-Strogatz model in the linear case: $a=1$	45
A.3.2 The Abrams-Strogatz model in the quasi-linear case: $\mathbf{a} = \mathbf{1} + \boldsymbol{\mu}$	46
A.4 Numerical simulations	49
A.4.1 Evaluation of stability	49
A.4.2 Perturbation theory	51
A.4.3 Evaluation of stability in critical dynamics	51
A.4.4 Numerical calculation of the derivative at the fixed points	55
A.5 Abrams-Strogatz Model with a varying population	56
B Data Pre-processing	57
B.1 Data Pre-fitting	57
B.2 Managing “I do not know (DK)” categories	58
C Data Processing	60
D Data Analysis	63
D.1 Plots interpretations and insights	64
D.2 Trends analysis	66
D.3 Metrics and correlations	67
D.3.1 M_{si} : Spanish to Indigenous metric	67
D.3.2 M_{pi} : Population to Indigenous metric	69
D.4 Data Analysis: conclusions	71

E Parameters estimation and optimization strategies	71
E.1 Volatility estimation	72
E.2 Prestige estimation	73
E.2.1 s_x estimation algorithm	73
E.3 Normalized prestige estimation	75
E.4 Loss functions and execution strategies	77
E.5 Error estimation for relevant model parameters	78

List of Figures

1 A schematic view of system X	5
2 In the last five surveys, we have observed a stabilization trend. The percentages of bilingual and monolingual speakers are decreasing, with slopes of -0.0143 and -0.0140 , respectively. Additionally, the Indigenous population is declining at a slope of -0.0330 . Meanwhile, the Spanish population is increasing at the same slope of 0.0330	7
3 Abrams-Strogatz model and actual data. Initial parameters and bounds independent from a . Full data range from 1895 to 2020 (top); two separate blocks: 1895 – 1970 and 1970 – 2020 (bottom).	10
4 Abrams-Strogatz model and actual data. Initial parameters and bounds depend on a . Full data range from 1895 to 2020 (top); two separate blocks: 1895 – 1970 and 1970 – 2020 (bottom).	11
5 Abrams-Strogatz model and actual data. Initial parameters and bounds independent from a . Full data range from 1895 to 2020 (top); two separate blocks: 1895 – 1970 and 1970 – 2020 (bottom).	12
6 Abrams-Strogatz model and actual data. Initial parameters and bounds depend on a . Full data range from 1895 to 2020 (top); two separate blocks: 1895 – 1970 and 1970 – 2020 (bottom).	13
7 Phase space, fixed points, and stability for the entire data range from 1895 to 2020 and for parameters $a = 1.2519$, $s_l = 0.0055$, and $s_o = 0.0349$	17
8 Phase space, fixed points, and stability for the entire data range from 1970 to 2020 and for parameters $a = 0.6553$, $s_l = 0.0149$, and $s_o = 0.0469$	18
9 Data points, ASM simulation, and perturbation theory for the early period from 1970 to 2020 with extrapolation until 2700.	19
10 Phase space, fixes points, and stability for the early period (1970-2020) and for parameters $a = 0.9789$ and $s_l = 0.4806$. ASM standard with single prestige value.	20
11 Phase space, fixes points, and stability for the early period (1970-2020) and for parameters $a = 0.9789$ and $s_l = 0.4806$. ASM quasi-linear with single prestige value.	20
12 ε^* in function of a and ν for the early period (1970-2020) and for parameters $a = 0.9789$ and $s_l = 0.4806$. ASM standard with single prestige value and $\nu = 0.0388$	22
13 ε^* in function of a and ν for the early period (1970-2020) and for parameters $a = 0.9789$ and $s_l = 0.4806$. ASM quasi-linear with single prestige value and $\nu = 0.0388$	22
14 Abrams-Strogatz oscillation and global percentage trends for a 125-year period.	23
15 Probability for a new Indigenous to be Bilingual.	25
16 ASM dynamics for two the overall periods: 1895 – 2020.	27
17 Vector field for x_o at $t = 2020$. Census period: 1895 – 2020.	28
18 ASM dynamics for two distinct periods: 1895 – 1970 and 1970 – 2020.	28
19 Vector field for x_o at $t = 2020$. Census period: 1895 – 2020.	29
A.1 As a function of a , ν , and ε , λ'_o shows that as ε approaches zero, the region where the function is positive tends to encompass the entire plot. When $\varepsilon < \varepsilon_o^*$, λ_o is positive, denoting a change of stability.	50
A.2 As a function of r , μ , and ε , λ_o shows that as ε approaches zero, the region where the function is positive gradually increases. When $\varepsilon < \varepsilon_o^*$, λ_o is positive, denoting a change of stability.	52

A.3	As a function of a , ν , and ε , λ'_o shows that as ε approaches zero, the region where the function is positive tends to encompass the entire plot. When $\varepsilon < \varepsilon_o^*$, λ_o is positive, denoting a change of stability.	53
A.4	As a function of a and ν , ε^* (from λ_o) decreases when ν approaches 1. The quasi-linear model has steeper curves, but the values tend to converge near $a = 1$	54
A.5	At $\nu = 0.5$, $r = 3$, The quasi-linear and the standard models have similar values when a tends to 1.	54
A.6	As a function of a and ν , x^* decreases similarly in both models, but the quasi-linear model reaches zero for smaller μ , compared with the standard model.	55
B.1	Missing data (red x) and logistic fitting for Population, Bilingual, Indigenous, and Spanish.	58
B.2	Trends in bilingual, monolingual, and “don’t know” (DK) responses within the Indigenous category.	59
B.3	Comparison of Original and Redistributed Trends for Bilingual and Monolingual Speakers within the Indigenous category.	59
C.1	Logistic curve and actual values for Population, Spanish, Bilingual, and Indigenous from 1895 to 2020.	60
C.2	Logistic curve and actual values for Population, Spanish, Bilingual, and Indigenous from 1500 to 2300.	61
C.3	Pre-fitted vs available data for Population. In 1939, they differ by no more than 8 percent.	62
C.4	Pre-fitted vs available data for Population. In 1932, they differ by no more than 3 percent.	63
D.1	Three possible trends for bilingual speakers within the Indigenous category.	64
D.2	Bilingual vs Indigenous. The upper left plot (normalized to 1) shows how Bilingual values fit the logistic curve of Spanish. The lower right plot shows the Q-Q plot.	65
D.3	Bilingual vs Spanish. The upper left plot (normalized to 1) shows how Bilingual values fit the logistic curve of Spanish. The lower right plot shows the Q-Q plot.	66
D.4	Bilingual and monolingual percentages calculated on Indigenous (upper plot) and Population (lower plot) categories.	67
D.5	Correlation between M_{si} and bilingual percentage, 1895 – –2020.	68
D.6	Correlation between M_{si} and bilingual percentage, last five censuses.	69
D.7	Correlation between M_{pi} and bilingual percentage w.r.t the overall population, 1895 – –2020.	70
D.8	Correlation between M_{pi} and monolingual percentage w.r.t the overall population, 1895 – –2020.	70
D.9	Comparison of bilingual and monolingual percentages: historical vs fitted time series	71
E.1	Monolingual percentages trends over the entire 1895 – –2020 period.	72
E.2	Polynomial fit of actual data.	72
E.3	Plots of s_l and s_o (scaled by k_{max}) for $a \in [0.5, 2]$	74
E.4	Plots of s_l and s_o (scaled by k_{max}) for $a \in [0.5, 2]$	75
E.5	Plots of s and $1 - s$ for $a \in [0.5, 2]$	76
E.6	Plots of s and $1 - s$ (scaled by k_{max}) for $a \in [0.5, 2]$	77

List of Algorithms

1	Estimation of s_l and s_o	73
2	Estimation of s and $1 - s$	76
3	Minimization process	77
4	Error estimation process. Optimal parameters, bounds, and loss function are calculated in 3.	79

Abstract

This article analyzes data from Mexican censuses from 1895 to 2020, focusing on the growth of *bilingual* speakers. We aim to understand the trends and drivers behind the increasing number of *bilingual* speakers in Mexico by employing mathematical models for Language Adoption and Language Dynamics.

1 Primary purpose of the paper and its structure

This paper examines the transition to and from Bilingualism in the Mexican indigenous community as a dynamic system X . In this system, the possible states are B and M . The state M represents the monolingual Indigenous community, while B represents the portion of the Indigenous who speak Spanish and other Indigenous languages. Refer to Figure 1 for a visual representation of these transitions.

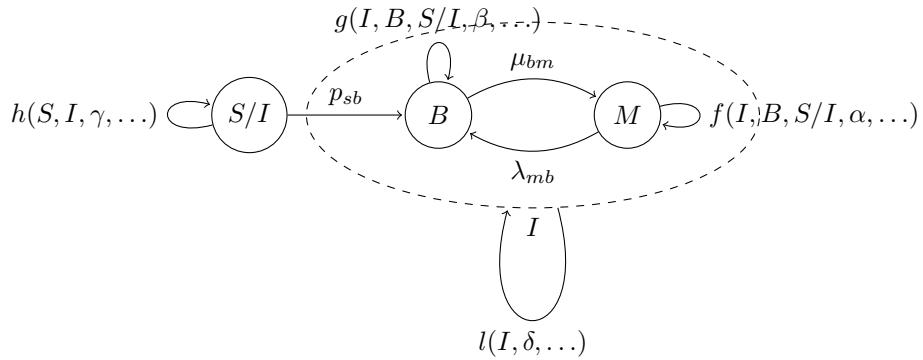


Figure 1: A schematic view of system X .

Figure 1 contains various interesting aspects summarized below:

Dashed ellipse: The dashed ellipse contains the two populations B and M . Both have growth dynamics, denoted by the functions f and g . The two populations respect the constraint:

$$B + M = I$$

The function l drives the dynamic of I ;

Action of node S/I : The node S/I is interpreted by an external action that modifies the dynamic of B .

We analyze two types of dynamics: the first type examines the internal changes of individual populations, M and B . Here, the factors λ_{mb} and μ_{bm} account for the transition probability $M \rightarrow B$ and $B \rightarrow M$, respectively. In other words, we focus on the dynamics within the ellipse. The functions f and g do not depend on S/I . In addition, we address this system using two different approaches:

- We consider I constant and use the percentages $b = B/I$ and $m = M/I$. In this approach, the constraint becomes $b + m = 1$. We employ: $\lambda_{mb}, \mu_{bm}, f(I, B, \alpha, \dots)$, and $g(I, B, \beta, \dots)$;
- We consider a varying population I and use the row numbers B and M . In this approach, the constraint stays $B + M = I$. We employ: $\lambda_{mb}, \mu_{bm}, f(I, B, \alpha, \dots), g(I, B, \beta, \dots)$, and $l(I, \delta, \dots)$.

The second type captures the action of S/I , p_{sb} on the number of B . The dynamics involve the entire set of transition probabilities and functions f, g, h , and l .

The organization of the paper is the following: the Introduction sets the context for this study and offers additional references for readers interested in exploring further; Section 3 describes the input data used in this research; Section B outlines the pre-processing tasks necessary to clean and normalize the data; Section C explains the data processing algorithms; and Section D introduces the data analysis.

2 Introduction

From a historical perspective, Mexico is a country where colonization, migratory phenomena, and intense urbanization, especially in recent times and in specific areas such as the capital and some specific tourist areas, have contributed to a stimulating situation from the point of view of the languages present, their integration, along with their adoption.

While migration phenomena have contributed to an increase in the number of speakers of European languages, Spanish colonization has undoubtedly contributed to the Spanish language's predominant status in Mexico.

Recent studies [1]–[3] emphasize how factors such as social integration, urbanization, and access to economic and educational opportunities have influenced many Indigenous speakers to adopt Spanish as a second language. This shift toward Bilingualism poses considerable challenges for the preservation of Indigenous languages. Understanding the dynamics of language adoption is essential for maintaining linguistic diversity and planning for sustainable cultural integration.

This paper first examines the available data on bilingualism trends in Mexico to propose dynamic systems that can capture many of the typical characteristics of language adoption patterns.

3 Input Data

The data for this analysis originates from the INEGI¹ (*El Instituto Nacional de Estadística y Geografía*). The organization conducts periodic censuses to monitor demographic changes in Mexico. The datasets found in the section titled *Censos y Conteos de Población y Vivienda* contain extensive information on education levels, ages², urbanization, migration, and spoken languages, among other topics. We are particularly interested in details about the Indigenous and Spanish-speaking populations and the proportion of bilingual individuals over time.

To accomplish this goal, we compiled a dataset that includes the following categories Unknown_Monolingual and Unknown_Spanish are examples of “I do not know (DK)” responses to surveys.:

Year Temporal dimension covering from 1895 to 2020;

Population Total population of Mexico;

Spanish Total number of people from families of Spanish genealogy;

Indigenous Total number of Indigenous people. Indigenous has three distinct sub-categories:

Bilingual Indigenous people who speak both Indigenous and Spanish languages;

Monolingual Indigenous people who speak Indigenous languages only;

Unknown_Monolingual Indigenous people who did not provide information on their status;

Unknown_Spanish Collect people who speak languages other than Spanish and/or Indigenous.

The relations among the categories are straightforward:

$$\text{Population} = \text{Spanish} + \text{Indigenous} + \text{Unknown_Spanish}$$

$$\text{Indigenous} = \text{Bilingual} + \text{Monolingual} + \text{Unknown_Monolingual}$$

The surveys reflect the unique characteristics of Mexico, where it is reasonable to assume that transitions from Indigenous languages to Bilingualism occur more frequently than transitions from Spanish to Bilingualism. Indeed, it is highly likely that an Indigenous individual will learn Spanish for work or educational reasons (see Section 2). Consequently, social researchers are interested in monitoring the relationship between bilinguals and monolinguals within the Indigenous population to provide valuable support for policy decisions to prevent causes of endangerment to Indigenous languages.

¹<https://www.inegi.org.mx/programas/>

²We filtered the data to include only individuals of both sexes who are older than five years.

4 Language dynamics: models

The analysis conducted in the previous sections indicates that the total population, monolingual Spanish speakers, bilingual speakers, and the Indigenous population exhibit growth patterns that resemble logistic curves. Notably, the growth rate of the bilingual population is slightly higher than that of the total population and significantly higher than that of the indigenous population, as shown in Table C.1. However, in recent years, the growth of the bilingual population has closely mirrored that of the indigenous population, as illustrated in Figure D.2. These findings and the methodology used in the surveys (which consider bilinguals as a subset of the Indigenous population) suggest that many Indigenous individuals are born into families that were already bilingual. Additionally, if bilingualism was adopted, it likely occurred before 1970.

We identified two significant correlations. The first shows a positive relationship between the rise in the ratio of Spanish speakers to indigenous speakers and the increase in bilingual speakers. The positive correlation suggests that the extent of contact the indigenous population has with Spanish speakers influences the likelihood of adopting Spanish as a second language, see Figure D.5. The second negative correlation indicates that the total percentage of bilingual and monolingual indigenous speakers decreases as the total ratio of these speakers to the indigenous population increases, cf. Figures D.8 and D.7.

We present the images that illustrate the relevant trends.

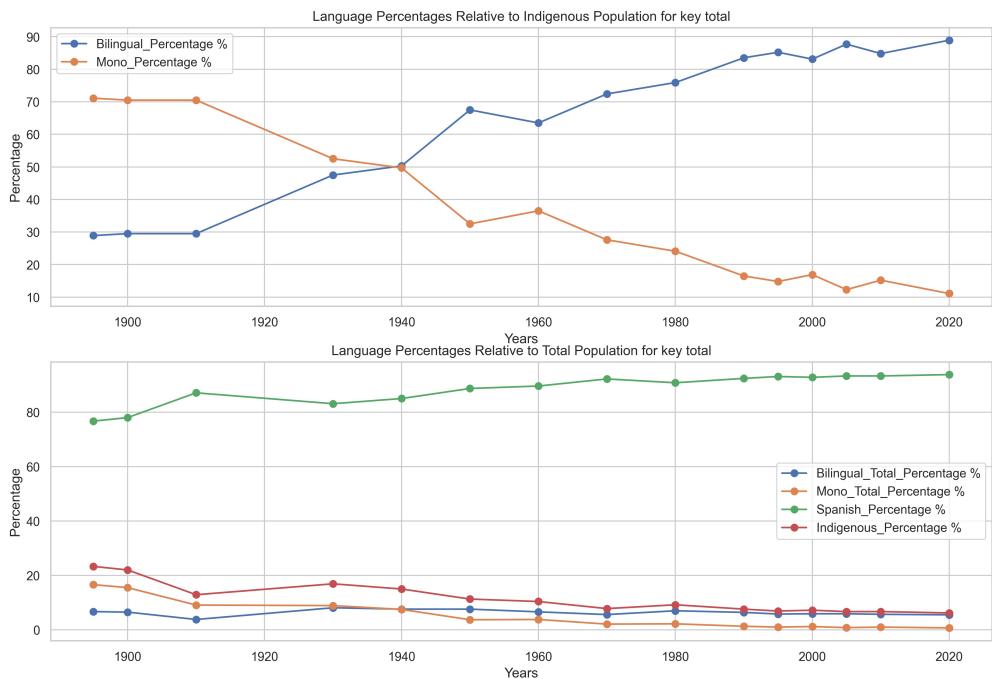


Figure 2: In the last five surveys, we have observed a stabilization trend. The percentages of bilingual and monolingual speakers are decreasing, with slopes of -0.0143 and -0.0140 , respectively. Additionally, the Indigenous population is declining at a slope of -0.0330 . Meanwhile, the Spanish population is increasing at the same slope of 0.0330 .

- A pattern of language adoption between Indigenous monolingual speakers and Spanish monolingual speakers;
- The total monolingual and bilingual percentages tend to stabilize;
- People become bilingual for two reasons: (i) People adopt Spanish as a second language, and (ii) are born into bilingual families;
- It is possible to revert from Bilingualism to Indigenous monolingualism.

4.1 Basic Abram-Strogatz model

The Abrams-Strogatz model (ASM) [13] addresses the competition between speakers of different languages and aims to predict the extinction of one of those languages. Defining what a language's death means is quite complex from linguistic, philological, and historical viewpoints. In the context of AMS, the death of a language is defined as the point at which the number of speakers of that language falls below a specific threshold. The model illustrates the competition between two languages by focusing on the proportion x_l of speakers of a language L , while $x_o = 1 - x_l$ represents the proportion of the remaining population. The model considers two key factors: prestige (s), which represents the status of a language and affects individuals' inclination to adopt it, and volatility (a), which adjusts individuals' sensitivity to differences in prestige between languages.

Mathematically, we can formalize the model as follows:

Let L and O two languages with speaker proportions x_l and x_o , respectively. ASM imposes $x_l + x_o = 1$. The rates of change of x_l and x_o are given by:

$$\frac{dx_l}{dt} = P_{OL}x_o - P_{LO}x_l \quad (2a)$$

$$\frac{dx_o}{dt} = P_{LO}x_l - P_{OL}x_o \quad (2b)$$

Where P_{LO} is the transition rate from language L to language O , or in other words, the rate of adoption of the language O by speakers of the language L . Similarly, P_{OL} is the transition rate from language O to language L . The transition rates P_{XY} are influenced by the prestige (s) of language Y and the number of speakers of Y , $P_{XY} = s_y Y^a$, as in (3)

$$P_{OL} = s_l x_l^\alpha \quad (3a)$$

$$P_{LO} = s_o x_o^\alpha. \quad (3b)$$

Substituting (3) into (2) we obtain

$$\frac{dx_l}{dt} = P_{OL}x_o - P_{LO}x_l = s_l x_l^\alpha x_o - s_o x_o^\alpha x_l \quad (4a)$$

$$\frac{dx_o}{dt} = P_{LO}x_l - P_{OL}x_o = s_o x_o^\alpha x_l - s_l x_l^\alpha x_o \quad (4b)$$

and since $x_l + x_o = 1$:

$$\frac{dx_l}{dt} = s_l x_l^\alpha (1 - x_l) - s_o (1 - x_l)^\alpha x_l \quad (5a)$$

$$\frac{dx_o}{dt} = s_o x_o^\alpha (1 - x_o) - s_l (1 - x_o)^\alpha x_o \quad (5b)$$

The Abrams-Strogatz Model has been studied and applied across various scenarios with different parameter choices and topologies, as noted in [6], [7], [10], [12].

4.1.1 Some comments on prestige (s) and volatility (a)

In the original paper by Abrams and Strogatz ([13]), the prestige of the languages sums to 1: $s_l + s_o = 1$. Therefore, when the language l has high prestige, the other language o has low prestige. Equations in (5) assume a symmetric form:

$$\frac{dx}{dt} = sx^\alpha(1 - x) - (1 - s)(1 - x)^\alpha x \quad (6)$$

In the same paper, the authors define a as a measure of the adoption process's non-linearity. When $a > 1$, the transitions between languages are nonlinear, and speakers tend to remain loyal to the majority language. This results in a stable state where one language dominates, which is referred to in the literature as a "low volatile" or "consensus" system. A "high volatility" or "coexistence" state occurs.

We can utilize the Abrams-Strogatz equations in equation (5) to derive valuable insights into the model's significant parameters, a and s , see Section E.

4.2 Running the Abrams-Strogatz model on experimental data

In this section, we apply the Abrams-Strogatz model to data from the Mexican census. Our focus is on the upper portion of Figure 2, where we utilize the ASM to understand the internal dynamics of the language adoption process among monolingual and bilingual speakers.

4.2.1 Data preparation

According to Sections B, the census data undergoes preprocessing, and any missing data is estimated using a specific logistic curve defined by the following equation:

$$L(t, K, r, N0) = \frac{K}{1 + \left(\frac{K-N0}{N0}\right) e^{-rt}}$$

In this equation, the parameters correspond to the growth rate (r), the initial population ($N0$), and the carrying capacity (K) for specific missing categories, such as Population, Spanish, Indigenous, etc. Data preprocessing also involves redistributing “I don’t know” responses. For the Indigenous category, these responses average 2%, so it is reasonable to redistribute them between Monolingual and Bilingual speakers. The conclusions presented in the C section indicate that the augmented data of missing values are as good as those that exclude them. This result allows for a broader analysis over a wider range of years.

4.2.2 Experiments

Section D.3 shows how the metric $m_{si} = \frac{N_S}{N_I}$ measures the correlation between the ratio of Spanish speakers to Indigenous speakers and the trend in the percentage of bilingual speakers within the Indigenous community. Even if m_{si} does not directly determine the prestige values s_o and s_l in the Abrams-Strogatz Model, it measures Spanish speakers’ (averaged) density per Indigenous. A higher m_{si} indicates more Spanish neighbors and a stronger external influence for indigenous speakers to adopt Spanish as a second language. A higher m_{si} drives a higher prestige value for the majority language, s_o , relative to the minority’s value s_l , changing the transition probabilities, see (3).

We report here equations (4) and (6) for the minority language x_l :

$$ASM = \begin{cases} \frac{dx_l}{dt} = s_l x_l^\alpha (1 - x_l) - s_o (1 - x_l)^\alpha x & \text{corresponding to an ASM with double prestige values} \\ & \text{and used for the first experiment.} \\ \frac{dx_l}{dt} = s_l x_l^\alpha (1 - x_l) - \overbrace{(1 - s_l)}^{s_o} (1 - x_l)^\alpha x & \text{corresponding to an ASM with single prestige} \\ & \text{and used for the second experiment.} \end{cases} \quad (7)$$

We completed two experiments corresponding to the model with double and single prestige values. Each experiment is conducted on the entire data interval from 1895 to 2020 and on two distinct time frames: 1895 to 1970 and 1970 to 2020. The methodology applies the algorithm presented in (Alg.4.1) to establish the initial parameters and their limits. Then, the methodology utilizes the algorithm described in (Alg.4.3) to improve the optimization of the fitting process. As suggested in Section E.4, an unbounded method is also tested to verify to which extent the optimal parameters resulting from the fitting processes are sensitive to initial parameters and their bounds. In addition, we tried a gradient-free minimization for loss functions, cf. again Section E.4.

Each experiment consists of two sub-experiments that use different relationships between the volatility parameter, the prestige values, and their bounds.

In the first sub-experiment, the initial prestige values and their associated bounds are independent of the volatility parameter a . In contrast, in the second sub-experiment, the initial prestige values and their associated bounds depend on the volatility parameter a .

By structuring experiments in this way, we can compare the results from the Abrams-Strogatz Model with double and single prestige values and evaluate the influence of the initial parameters on the fitting

process. Moreover, we can verify if the prestige values' dependence on the volatility parameter a affects the model performance.

Figures 3 and 4 present the results of the fitting processes for the first macro experiment, while Tables 1 and 2 summarize these results.

In a similar manner, Figures 3 and 4 also display the fitting process results for the second macro experiment, with Tables 3 and 4 providing a summary of those results.

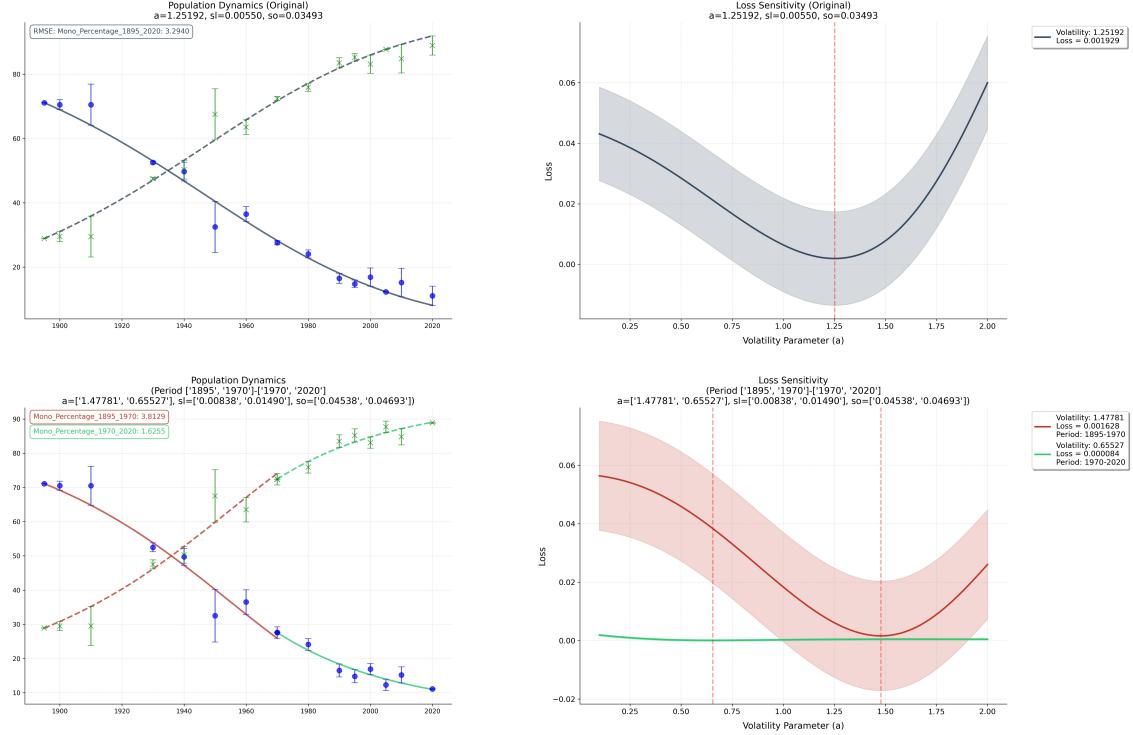


Figure 3: Abrams-Strogatz model and actual data. Initial parameters and bounds independent from a . Full data range from 1895 to 2020 (top); two separate blocks: 1895 – – 1970 and 1970 – 2020 (bottom).

Period (From, To)	Optimal Params	Initial Params	Bounds	Loss function	Min. Loss value
(1895, 2020)	$a = 1.2519 \pm 0.3130$ $s_l = 0.0055 \pm 0.0007$ $s_o = 0.0349 \pm 0.0087$	$a = 1.1$ $s_l = 0.0131$ $s_o = 0.0267$	$a \in (0.3, 2.2)$ $s_l \in (0.0046, 0.0867)$ $s_o \in (0.0058, 0.0602)$	weighted	$1.93 \cdot 10^{-3}$
(1895, 1970)	$a = 1.4778 \pm 0.3653$ $s_l = 0.0084 \pm 0.0020$ $s_o = 0.0454 \pm 0.0094$	$a = 1.5$ $s_l = 0.0131$ $s_o = 0.0267$	$a \in (0.3, 2.2)$ $s_l \in (0.0046, 0.0867)$ $s_o \in (0.0058, 0.0602)$	weighted	$1.63 \cdot 10^{-3}$
(1970, 2020)	$a = 0.6553 \pm 0.1638$ $s_l = 0.0149 \pm 0.0037$ $s_o = 0.0469 \pm 0.0092$	$a = 1.3$ $s_l = 0.0131$ $s_o = 0.0267$	$a \in (0.3, 2.2)$ $s_l \in (0.0046, 0.0867)$ $s_o \in (0.0058, 0.0602)$	weighted	$8.4 \cdot 10^{-5}$
(1970, 2020) ^(*)	$a = 0.5963 \pm 0.1486$ $s_l = 0.0112 \pm 0.0028$ $s_o = 0.0423 \pm 0.0092$	$a = 0.7$ $s_l = 0.0131$ $s_o = 0.0267$	$a \in (0.3, 2.2)$ $s_l \in (0.0046, 0.0867)$ $s_o \in (0.0058, 0.0602)$	weighted	$8.4 \cdot 10^{-5}$
(1970, 2020) ^(**)	$a = 0.6463 \pm 0.1616$ $s_l = 0.0142 \pm 0.0036$ $s_o = 0.0461 \pm 0.0093$	$a = 0.9$ $s_l = 0.0131$ $s_o = 0.0267$	$a \in (0.3, 2.2)$ $s_l \in (0.0046, 0.0867)$ $s_o \in (0.0058, 0.0602)$	weighted	$8.4 \cdot 10^{-5}$

Table 1: Report on the fitting process for the first sub-experiment: prestige parameters and bounds are independent of a . Values are obtained using unbounded strategies in (*) and SciPy’s *differential_evolution* as the optimizer in (**).

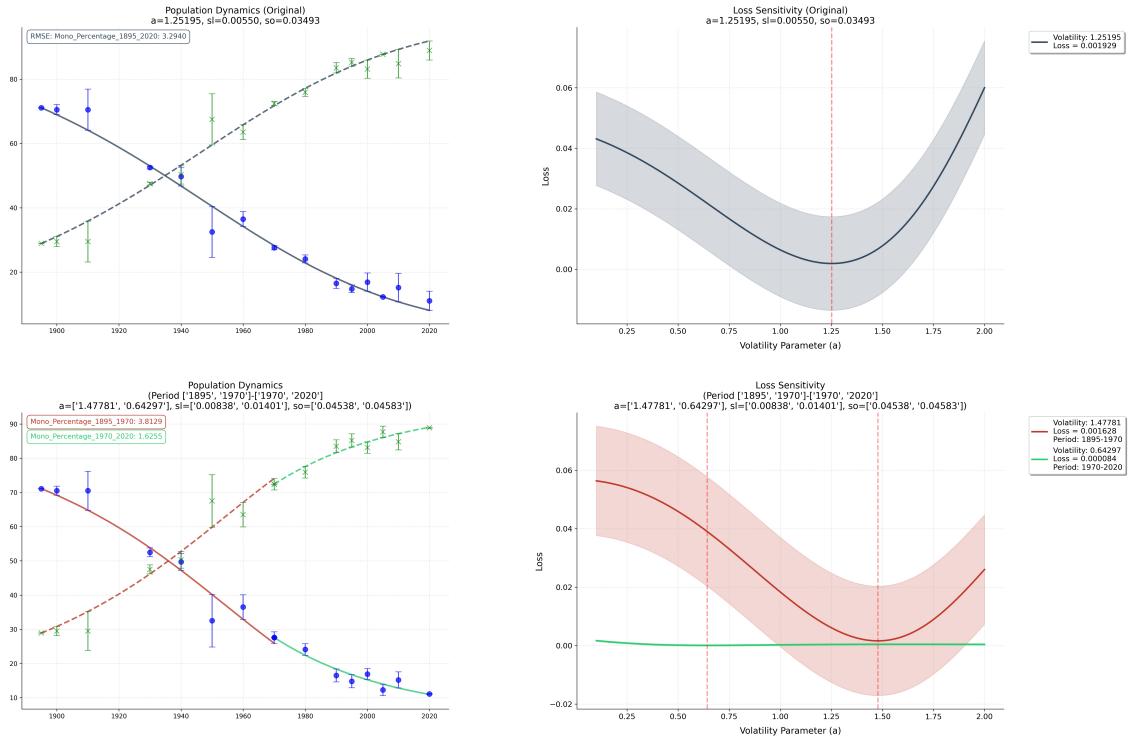


Figure 4: Abrams-Strogatz model and actual data. Initial parameters and bounds depend on a . Full data range from 1895 to 2020 (top); two separate blocks: 1895 – 1970 and 1970 – 2020 (bottom).

Period (From, To)	Optimal Params	Initial Params	Bounds	Loss function	Min. Loss value
(1895, 2020)	$a = 1.2520 \pm 0.3130$ $s_l = 0.0055 \pm 0.0007$ $s_o = 0.0349 \pm 0.0087$	$a = 0.5$ $s_l = 0.0131$ $s_o = 0.0267$	$a \in (0.3, 2.2)$ $s_l \in (0.0046, 0.0284)$ $s_o \in (0.0058, 0.0602)$	weighted	$1.93 \cdot 10^{-3}$
(1895, 1970)	$a = 1.4778 \pm 0.3653$ $s_l = 0.0084 \pm 0.0018$ $s_o = 0.0454 \pm 0.0094$	$a = 0.5$ $s_l = 0.0131$ $s_o = 0.0267$	$a \in (0.3, 2.2)$ $s_l \in (0.0046, 0.0284)$ $s_o \in (0.0058, 0.0602)$	weighted	$1.63 \cdot 10^{-3}$
(1970, 2020)	$a = 0.6430 \pm 0.1607$ $s_l = 0.0140 \pm 0.0035$ $s_o = 0.0458 \pm 0.0104$	$a = 0.7$ $s_l = 0.0195$ $s_o = 0.029$	$a \in (0.3, 2.2)$ $s_l \in (0.0055, 0.0415)$ $s_o \in (0.0064, 0.0645)$	weighted	$8.4 \cdot 10^{-5}$
(1970, 2020) ^(*)	$a = 0.6508 \pm 0.1627$ $s_l = 0.0146 \pm 0.0036$ $s_o = 0.0465 \pm 0.0114$	$a = 0.9$ $s_l = 0.0292$ $s_o = 0.0316$	$a \in (0.3, 2)$ $s_l \in (0.0067, 0.0622)$ $s_o \in (0.007, 0.069)$	weighted	$8.4 \cdot 10^{-5}$
(1970, 2020) ^(**)	$a = 0.6457 \pm 0.1614$ $s_l = 0.0142 \pm 0.0036$ $s_o = 0.0461 \pm 0.0093$	$a = 0.5$ $s_l = 0.0131$ $s_o = 0.0267$	$a \in (0.3, 2)$ $s_l \in (0.0067, 0.0622)$ $s_o \in (0.007, 0.069)$	weighted	$8.4 \cdot 10^{-5}$

Table 2: Report on the fitting process for the second sub-experiment: prestige parameters and bounds depend on a . Values are obtained using unbounded strategies in (*) and SciPy’s *differential_evolution* as the optimizer in (**).

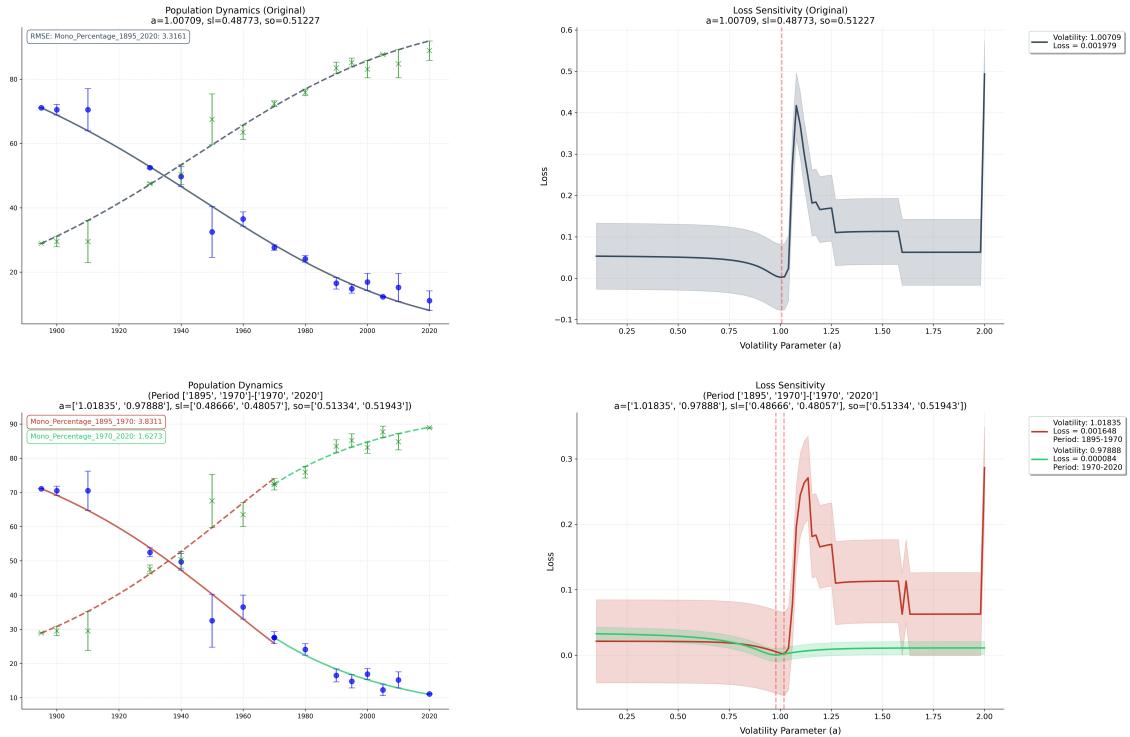


Figure 5: Abrams-Strogatz model and actual data. Initial parameters and bounds independent from a . Full data range from 1895 to 2020 (top); two separate blocks: 1895 – – 1970 and 1970 – 2020 (bottom).

Period (From, To)	Optimal Params	Initial Params	Bounds	Loss Function	Min. Loss Value
(1895, 2020)	$a = 1.0071 \pm 0.1610$ $s_l = 0.4877 \pm 0.0848$	$a = 0.5$ $s_l = 0.2943$	$a \in (0.3, 2.2)$ $s_l \in (0.1668, 0.4972)$	Weighted	$1.9 \cdot 10^{-3}$
(1895, 1970)	$a = 1.0184 \pm 0.2546$ $s_l = 0.4867 \pm 0.0635$	$a = 0.5$ $s_l = 0.2943$	$a \in (0.3, 2.2)$ $s_l \in (0.1668, 0.4972)$	Weighted	$1.65 \cdot 10^{-3}$
(1970, 2020)	$a = 0.9789 \pm 0.2447$ $s_l = 0.4806 \pm 0.0642$	$a = 0.5$ $s_l = 0.2943$	$a \in (0.3, 2.2)$ $s_l \in (0.1668, 0.4972)$	Weighted	$8.4 \cdot 10^{-5}$
(1970, 2020) ^(*)	$a = 0.9789 \pm 0.2447$ $s_l = 0.4806 \pm 0.0642$	$a = 0.9$ $s_l = 0.2943$	$a \in (0.3, 2.2)$ $s_l \in (0.1668, 0.4972)$	Weighted	$8.4 \cdot 10^{-5}$
(1970, 2020) ^(**)	$a = 0.9789 \pm 0.2447$ $s_l = 0.4806 \pm 0.0642$	$a = 0.7$ $s_l = 0.2943$	$a \in (0.3, 2.2)$ $s_l \in (0.1668, 0.4972)$	Weighted	$8.4 \cdot 10^{-5}$

Table 3: Report of the fitting process for the first sub-experiment ($s_0 = 1 - s_l$). Prestige parameters and bounds are independent of a . Values are obtained using unbounded strategies in (*) and SciPy’s *differential_evolution* as the optimizer in (**).

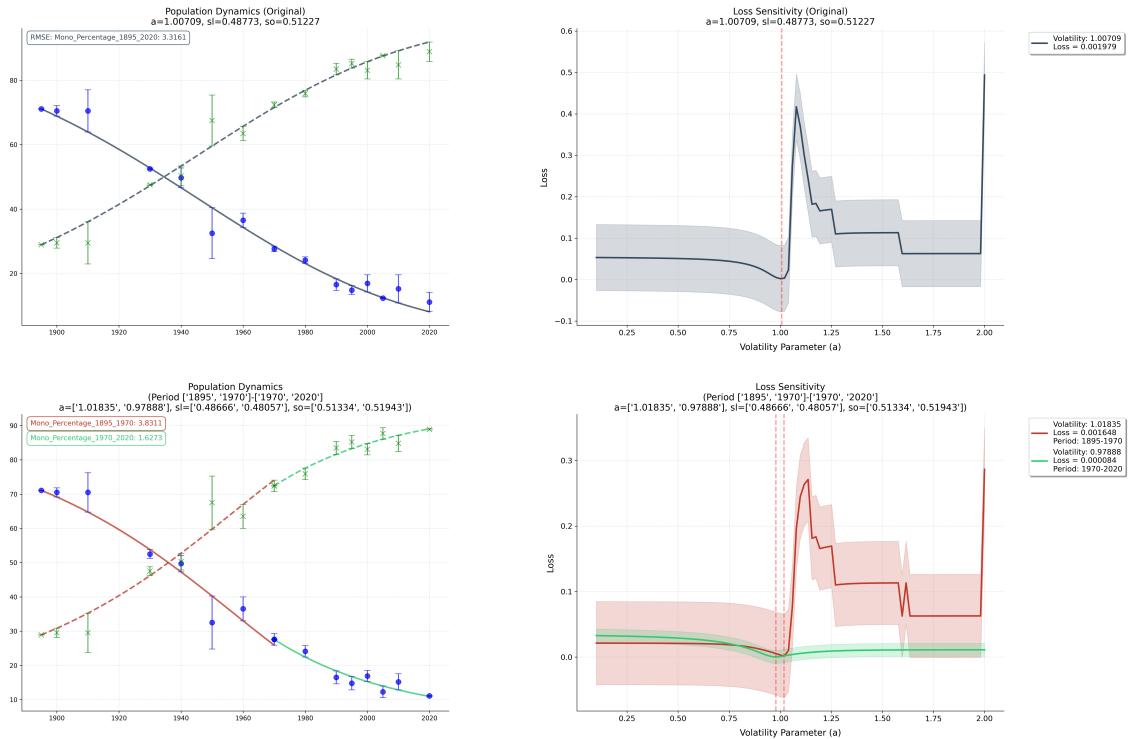


Figure 6: Abrams-Strogatz model and actual data. Initial parameters and bounds depend on a . Full data range from 1895 to 2020 (top); two separate blocks: 1895 – 1970 and 1970 – 2020 (bottom).

Period (From, To)	Optimal Params	Initial Params	Bounds	Loss Function	Min. Loss Value
(1895, 2020)	$a = 1.0071 \pm 0.1610$ $s_l = 0.4877 \pm 0.0848$	$a = 1.1$ $s_l = 0.4972$	$a \in (0.3, 2.2)$ $s_l \in (0.4521, 0.4972)$	Weighted	$1.9 \cdot 10^{-3}$
(1895, 1970)	$a = 1.0184 \pm 0.2546$ $s_l = 0.4867 \pm 0.0151$	$a = 1.1$ $s_l = 0.4775$	$a \in (0.3, 2.2)$ $s_l \in (0.4521, 0.4972)$	Weighted	$1.65 \cdot 10^{-3}$
(1970, 2020)	$a = 0.9789 \pm 0.2447$ $s_l = 0.4806 \pm 0.0299$	$a = 1.3$ $s_l = 0.4292$	$a \in (0.3, 2.2)$ $s_l \in (0.3673, 0.4869)$	Weighted	$8.4 \cdot 10^{-5}$
(1970, 2020) ^(*)	$a = 0.9789 \pm 0.2447$ $s_l = 0.4806 \pm 0.0355$	$a = 0.7$ $s_l = 0.4292$	$a \in (0.3, 2.2)$ $s_l \in (0.3426, 0.4846)$	Weighted	$8.4 \cdot 10^{-5}$
(1970, 2020) ^(**)	$a = 0.9789 \pm 0.2447$ $s_l = 0.4806 \pm 0.0355$	$a = 0.7$ $s_l = 0.4292$	$a \in (0.3, 2.2)$ $s_l \in (0.3426, 0.4846)$	Weighted	$8.4 \cdot 10^{-5}$

Table 4: Report of the fitting process for the first sub-experiment ($s_0 = 1 - s_l$). Prestige parameters and bounds depend on a . Values are obtained using unbounded strategies in (*) and SciPy’s *differential_evolution* as the optimizer in (**).

4.2.3 Experiments: comments and insights on optimal and initial parameters

This section summarizes the salient facts that emerge from analyzing the values of a , s_l , and s_o in the various sub-experiments described in Tables 1, 2, 3, and 4.

Tables 1 and 2 show similar values for both the optimal parameters and the associated uncertainties. Recall that the optimal parameters are computed according to the algorithm (Alg.4.3). Still, the initial parameters on which the minimization is computed differ between the sub-experiments.

Table 1 shows the results of the first sub-experiment, in which the only varying parameter is the volatility a . The experiments conducted over the entire duration (as shown in the first row) and during the period from 1895 to 1970 (as shown in the second row) yield optimal values for the volatility a close to the respective initial values: 1.2519 from 1.1 and 1.4778 from 1.5. In contrast, the last three rows of Table 1 demonstrate that the initial value of s_l is the primary factor in the optimization processes. This finding holds true regardless of the approach used: whether it is bounded (third row), unbounded (fourth row), or based on the SciPy differential evolution method. When we analyze the results in Table 2, we notice that the optimal parameters do not show an unambiguous dependence on one of the initial parameters, resulting in minimization processes able to investigate the entire combination of the initial parameters. Nevertheless, it is noticeable that the values of the loss functions’ minima are the same in both sub-experiments.

The scenario changes dramatically in the sub-experiments run with the Abrams-Strogatz model, which assumes the sum of the prestige values s_l and s_o equals 1. Tables 3 and 4 show little or no variability in the values of a and s_l independently of the sub-experiments conducted, including the last three rows, whose values, as we mentioned, are obtained with three different approaches.

As a final comment on the optimization processes, we note that the values of the optimal values in the last three rows of Tables 1 and 2 are very similar but not exactly the same, confirming how fitting and optimization processes can identify local minima of loss functions, but that these are reached with different combinations. On the other hand, this is not the case for the second experiment, where the additional constraint on prestige values limits the optimizers’ freedom to range in parameter space, allowing them to work on fewer combinations.

4.2.4 Experiments: comments and insights on RMSEs

This section describes the models’ performances in the first and second experiments based on the root mean square errors (RMSEs). As we see from the plots reported in Figures 3 and 6 for the first and second experiments, the provided RMSE are comparable.

The model for the 125-year period shows an RMSE of approximately 3.3. When we contextualize this error relative to the total percentage change decrease ($p_c = 71\% - 11\% = 60\%$), we find the RMSE represents only 5.5% of the total change, indicating that the model’s error is somewhat small when compared to the overall trend. If the error is evenly distributed (and this assumption is sound since RMSE is a mean square error), an RMSE of 3.3 corresponds to an average error of $3.3/\sqrt{125} \approx 0.29$ per year. This RMSE effectively captures the general trend of the decline in the percentage of monolingual speakers over the timeframe.

For the early period (1895-1970, 75 years), the model produces an RMSE of 3.8. In this timeframe, $p_c \approx 43.4\%$. The resulting ratio is 8.8%, higher than the entire period. The increased relative error arises primarily from considerable residuals in the 1950 and 1960 data points, where the model shows a more significant deviation from observed values. Even the normalized error per year is higher than the full-period model (0.44 vs. 0.29), suggesting a reduced accuracy. Notably, the model extrapolates a lower value for 1970 than the one observed in the data, indicating a potential deceleration (caused by external factors) in language shift that this period's parameterization does not entirely capture.

The recent 50-year period (1970-2020) exhibits a lower absolute RMSE of 1.6. During this time, $p_c \approx 16.6\%$, and the ratio is about 9.6%, comparable to the early period. The smaller RMSE mainly reflects the smaller scale of percentage change occurring during this period rather than a better model performance.

4.2.5 Experiments: model comparison and external factors

Figure 3 indicates a significant shift in the volatility parameter, a . In the full period model, $a = 1.2519$, which is relatively close to the values of 1.31 and 1.34 reported in the literature [13]. The value of 1.47781 found for the early period (1895-1970) is moderately higher than typical literature values, potentially demonstrating a stronger trend towards “consensus”. In contrast, $a = 0.65527$ shows a notable deviation from typical findings, indicating a trend towards a status in which monolingual and bilingual speakers coexist.³

The prestige value parameters have shown significant evolution when we compare the early period with the more recent one. The prestige of minority languages has changed from $s_l = 0.00838$ to $s_l = 0.01490$. On the other hand, the prestige of the majority languages, s_o , remains relatively stable, with values of 0.04538 compared to 0.04693, but it continues to dominate overall.

Below, we compare the optimal parameters obtained from the second experiment (refer to Figure 6) with those from the first experiment. For the entire period, the model shows $a = 1.00709$ (compared to 1.26515); in the early period (1895-1970), $a = 1.01835$ (was 1.47781); and in the recent period (1970-2020), $a = 0.97888$ (was 0.65527).

The most remarkable difference is that all volatility parameters are now close to 1, representing a critical threshold in the Abrams-Strogatz model, cf. Section A.3.2. Values around 1 reveal a competition where prestige effects and frequency (number of speakers around) effects similarly influence language selection.

The difference in volatility between the early and recent periods is significantly smaller in the second experiment compared to the first. The second experiment presents the same scenario from “consensus” to “coexistence” but with a more balanced transition.

The values of the prestige status parameters, s_l and s_o , indicate a much more balanced status relationship between the competing languages than observed in the previous model. Interestingly, the status parameters remain stable across different periods.

4.2.6 Fixed points and stability

This section analyzes some of the fixed points of the four sub-experiments and their stability. Please refer to Section A.4 for details on the algorithms and methodology used.

From the first experiment, we focus on the first sub-experiment. In this experiment, the Abrams-Strogatz model employs a double prestige whose values and bounds are independent of the volatility parameter a . The fitting process on the entire period (from 1895 to 2020) returns the following optimal parameters, see Table 1:

$$a = 1.2519 \pm 0.3130, s_l = 0.0055 \pm 0.0007, \text{ and } s_o = 0.0349 \pm 0.0087$$

The equation

$$\frac{dx_l}{dt} = s_l x_l^\alpha (1 - x_l) - s_o (1 - x_l)^\alpha x$$

has three fixed points, cf. Section A:

$$x_l = 0, x_l = 1, \text{ and } x_l = x_l^*$$

³A “coexistence” status reflects language protection policies and a cultural revaluation of the minority language.

where

$$x_l^* = \frac{\beta}{1 + \beta} \text{ with } \beta = \left(\frac{s_o}{s_l} \right)^{\frac{1}{a-1}}$$

The numerical calculation of the derivatives at the fixed points provides the values in (8). Moreover, according to (A.6) in Section A.2, the fixed points $x_l = 0$ and $x_l = 1$ are stable, while x_l^* is unstable.

$$\frac{dx_l}{dt} = \begin{cases} \left. \frac{dx_l}{dt} \right|_{x_l=0} = -0.0348 \pm 0.0108 \text{ corresponding at } x_l = 0 \\ \left. \frac{dx_l}{dt} \right|_{x_l=1} = -0.0050 \pm 0.0013 \text{ corresponding at } x_l = 1 \\ \left. \frac{dx_l}{dt} \right|_{x_l=x_l^*} = 0.0014 \pm 0.0004 \text{ corresponding at } x_l \approx 0.9994 \end{cases} \quad (8)$$

These values agree with the theoretical eigenvalues predicted in (A.6), which for $a > 1$ are the following:

$$\lambda_{a>1} = \begin{cases} -s_o = -0.0349 \pm 0.0087 \text{ corresponding at } x_l = 0 \\ -s_l = -0.0055 \pm 0.0007 \text{ corresponding at } x_l = 1 \\ s_o(a-1) \left(\frac{1}{1 + \left(\frac{s_o}{s_l} \right)^{\frac{1}{a-1}}} \right)^{a-1} = 0.00155 \pm 0.000133 \text{ corresponding at } x_l \approx 0.9994 \end{cases} \quad (9)$$

Please note that the bare calculation of the stable point is:

$$s_o(a-1) \left(\frac{1}{1 + \left(\frac{s_o}{s_l} \right)^{\frac{1}{a-1}}} \right)^{a-1} = 0.00138$$

The value of 0.00155 ± 0.000133 results from a Monte-Carlo simulation with 100000 samples.

Figure 7 reports the fixed points (red for unstable, green for stable fixed points), the trend of the derivative of x_l (light blue), and the data points (blue points) mapped on the derivative trend.

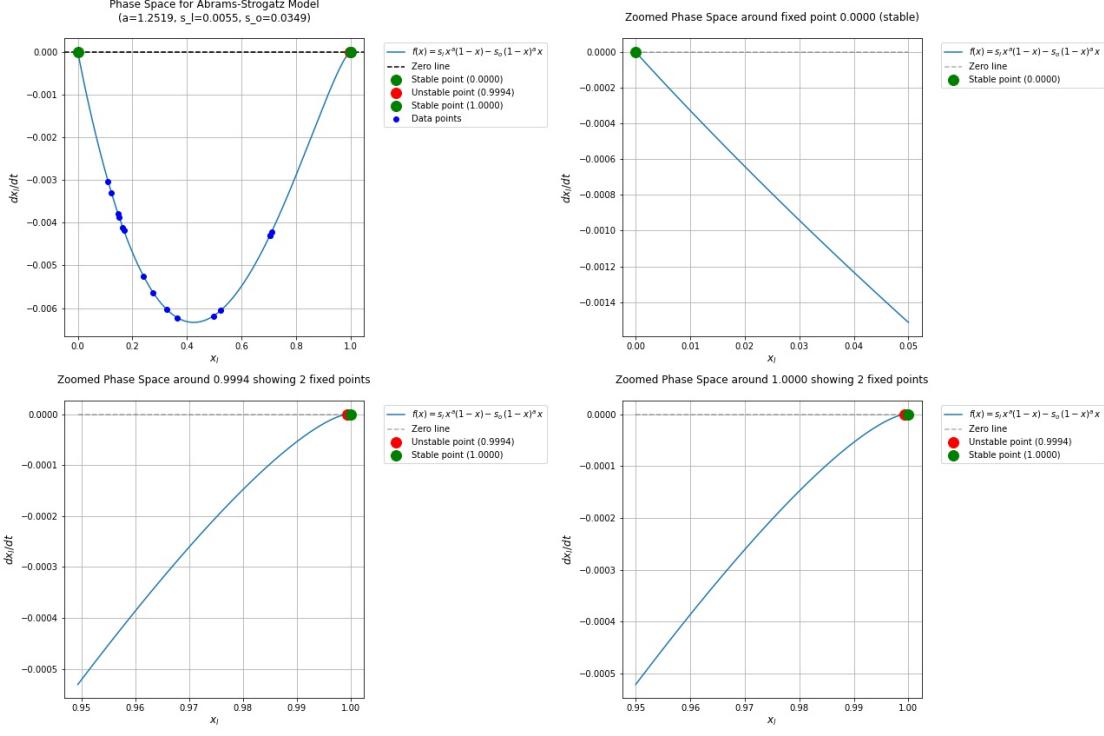


Figure 7: Phase space, fixed points, and stability for the entire data range from 1895 to 2020 and for parameters $a = 1.2519$, $s_l = 0.0055$, and $s_o = 0.0349$

Similarly, the fitting process on the early period (from 1970 to 2020) returns the following optimal parameters, see third row of Table 1:

$$a = 0.6553 \pm 0.1638, s_l = 0.0149 \pm 0.0037, \text{ and } s_o = 0.0469 \pm 0.0092$$

The numerical calculation of the derivatives at the fixed points provides the values in (10): the fixed points $x_l = 0$ and $x_l = 1$ are unstable, x_l^* is stable. The values of the derivatives are calculated according to (A.49) in Section A.4.4.

$$\frac{dx_l}{dt} = \begin{cases} \left. \frac{dx_l}{dt} \right|_{x_l=0} & = 2.7845 \pm 0.1217 \text{ corresponding at } x_l = 0 \\ \left. \frac{dx_l}{dt} \right|_{x_l=1} & = 8.8975 \pm 0.1884 \text{ corresponding at } x_l = 1 \\ \left. \frac{dx_l}{dt} \right|_{x_l=x_l^*} & = -0.0164 \pm 0.0005 \text{ corresponding at } x_l \approx 0.0347 \end{cases} \quad (10)$$

Figure 8 reports the fixed points (red for unstable, green for stable fixed points), the trend of the derivative of x_l (light blue), and the data points (blue points) mapped on the derivative trend for the early period.

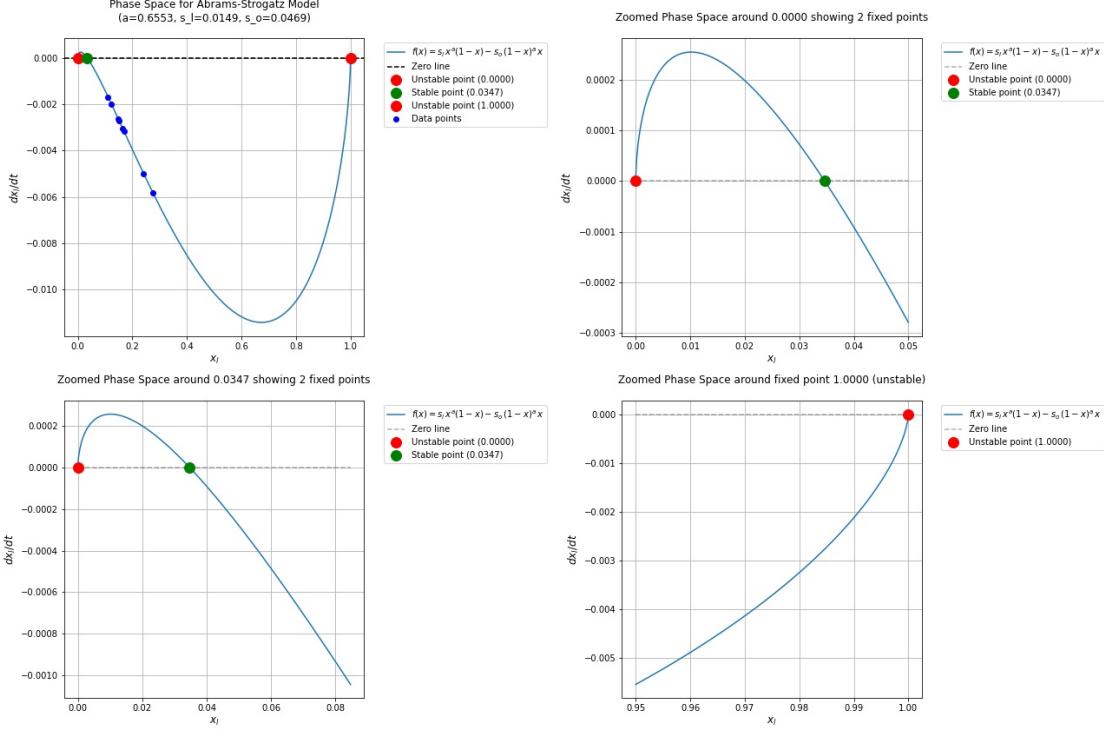


Figure 8: Phase space, fixed points, and stability for the entire data range from 1970 to 2020 and for parameters $a = 0.6553$, $s_l = 0.0149$, and $s_o = 0.0469$

Since $x_l = 0.0347$ is a stable fixed point, it represents a basin of attraction. This means trajectories starting from points close to $x_{th} = 0.0347$ will converge to 0.0347. We verified this statement through the following steps:

1. We numerically simulated the dynamic equation of the ASM using the volatility and prestige values from the third row of Table 1. As a starting value, we set $x_l(0) = 0.276$, which represents the percentage of initial monolingual Indigenous speakers in 1970, based on census data, and then calculated the time t_f at which $x_l \approx x_{th}$, obtaining $t_f \approx 2684$;
2. We extrapolated the time t_p at which $x_l = 2x_{th}$, obtaining the value of ≈ 2093 ;
3. We selected five x_0 between $\frac{x_{th}}{2}$ and $2x_{th}$; fixed point from $x_l^* = \frac{\beta}{1+\beta}$ obtaining $x_l^* =$
4. Starting from $t_p = 2093$, we applied perturbation theory using:

$$x_l(t) = x_{th} + (x_o(t_p) - x_{th}) e^{-\lambda^*(t-t_p)}$$

where $\lambda^* = -0.0164$.

Figure 9 plots the percentage of monolingual speakers, the ASM simulation, and perturbation theory for the early period from 1970 to 2020 with extrapolation until 2700.

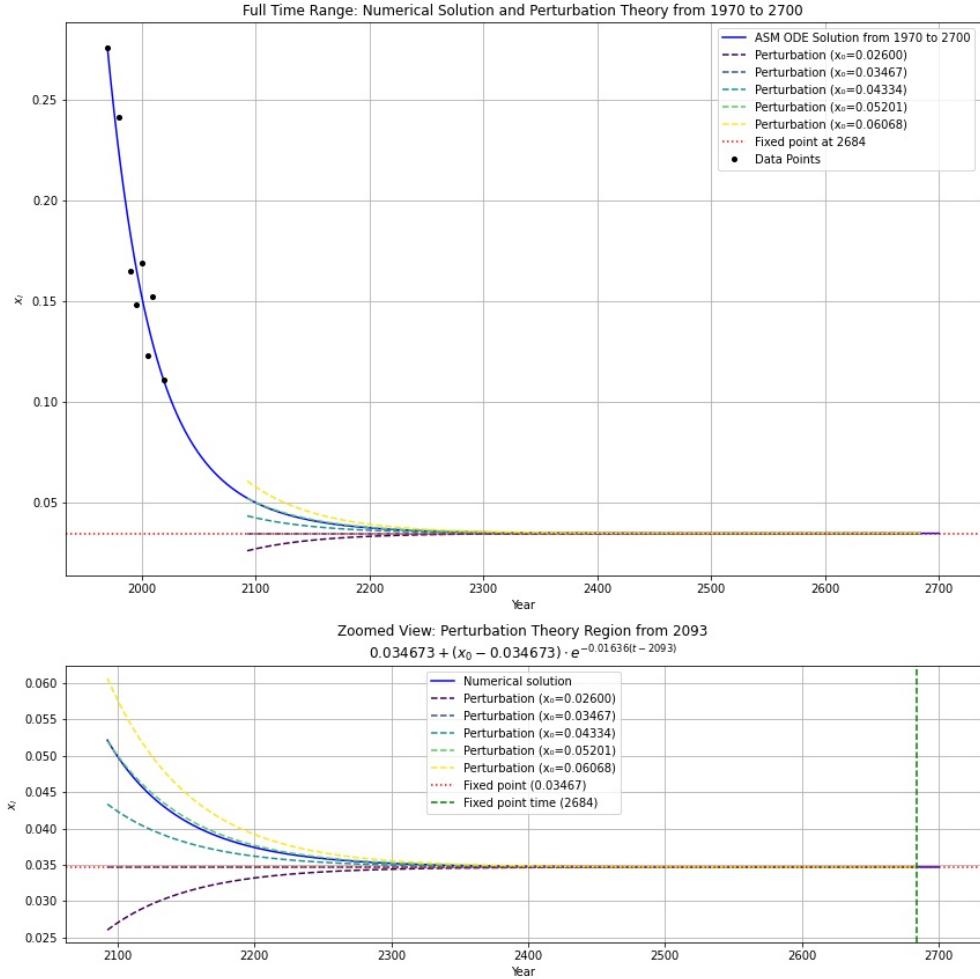


Figure 9: Data points, ASM simulation, and perturbation theory for the early period from 1970 to 2020 with extrapolation until 2700.

We compare the fixed point x^* results with parameters $a = 0.9789$ and $s_l = 0.4806$, obtained using standard and quasi-linear single prestige AS models. The parameters are the following:

$$a = 0.9789 \quad (\mu = -0.0211) \pm 0.2447 \text{ and } s_l = 0.4806 \pm 0.0299$$

The following figures show the closeness of the fixed points $x = 0$ and $x = x^*$ in the case of standard and quasi-linear models:

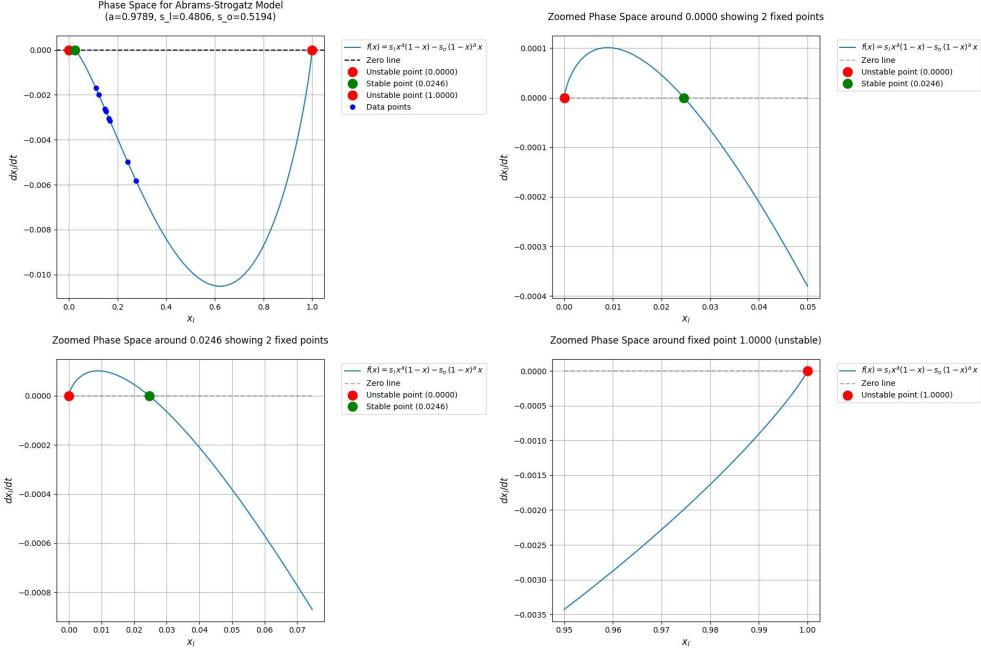


Figure 10: Phase space, fixes points, and stability for the early period (1970-2020) and for parameters $a = 0, 9789$ and $s_l = 0.4806$. ASM standard with single prestige value.

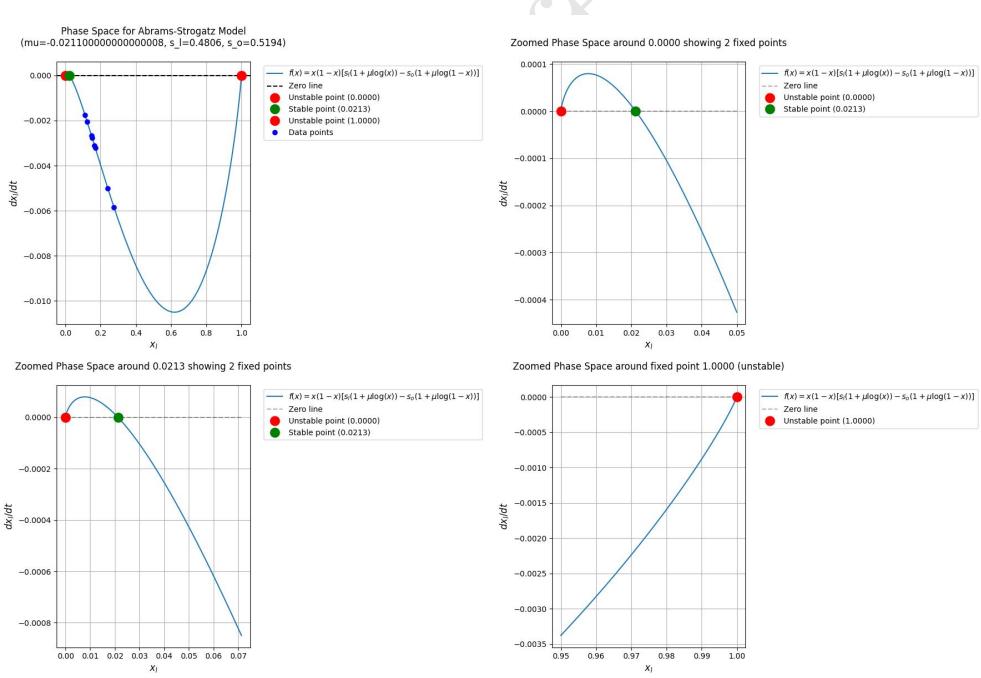


Figure 11: Phase space, fixes points, and stability for the early period (1970-2020) and for parameters $a = 0, 9789$ and $s_l = 0.4806$. ASM quasi-linear with single prestige value.

The numerical simulation extrapolates a $t_p \approx 4100$ for both x^* (in the standard and quasi-linear model); the perturbation theory is not shown here. The numerical calculation of the derivatives at the fixed points, according to the standard Abrams-Strogatz model, provides the values in (11): the fixed points $x_l = 0$ and

$x_l = 1$ are unstable, x_l^* is stable. The values of the derivatives are calculated according to (A.49) in Section A.4.4.

$$\frac{dx_l}{dt} = \begin{cases} \left. \frac{dx_l}{dt} \right|_{x_l=0} & 0.1551 \pm 0.0244 \text{ corresponding at } x_l = 0 \\ \left. \frac{dx_l}{dt} \right|_{x_l=1} & = 0.2484 \pm 0.0470 \text{ corresponding at } x_l = 1 \\ \left. \frac{dx_l}{dt} \right|_{x_l=x_l^*} & = -0.0110 \pm 0.0029 \text{ corresponding at } x_l \approx 0.0246 \end{cases} \quad (11)$$

We calculated the derivative values at the fixed points for the quasi-linear model, as stated in (A.50) in Section A.4.4.

$$\frac{dx_l}{dt} = \begin{cases} \left. \frac{dx_l}{dt} \right|_{x_l=0} & 0.1145 \pm 0.0642 \text{ corresponding at } x_l = 0 \\ \left. \frac{dx_l}{dt} \right|_{x_l=1} & = 0.2045 \pm 0.0642 \text{ corresponding at } x_l = 1 \\ \left. \frac{dx_l}{dt} \right|_{x_l=x_l^*} & = -0.0102 \pm 0.0027 \text{ corresponding at } x_l \approx 0.0213 \end{cases} \quad (12)$$

In the case of the standard model, the fixed point $x = 0$ becomes unstable under very small perturbations. By setting $\lambda_0 = 0$ in equation (A.6), and using the parameters $a = 0.9785$ and $s_l = 0.4806$, we find that $\varepsilon^* \approx 9.1 \times 10^{-3}$.

Conversely, for the quasi-linear model, when we set $\lambda_0 = 0$ in equation (A.43) and use the same parameters, we obtain a slightly smaller value of $\varepsilon_{ql}^* \approx 8 \times 10^{-3}$.

The following figures report the values of ε^* x^* in the case of standard and quasi-linear models. We notice that those values are not so different between the two models, as described in Section A.3.2. The pictures are generated using $\nu = 0.0388$ which corresponds to a ratio $r = 0.5194/0.4806 \approx 1.081$

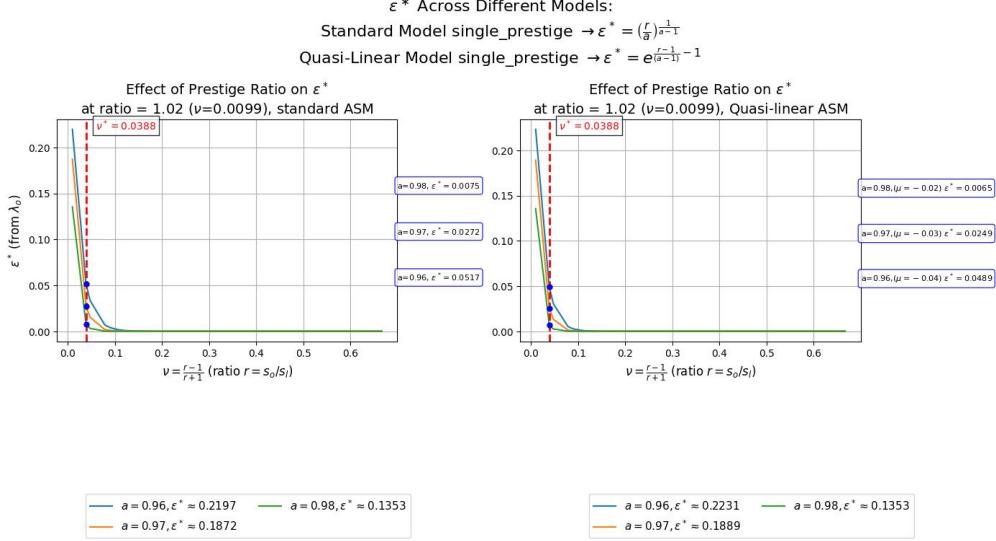


Figure 12: ε^* in function of a and ν for the early period (1970-2020) and for parameters $a = 0, 9789$ and $s_l = 0.4806$. ASM standard with single prestige value and $\nu = 0.0388$

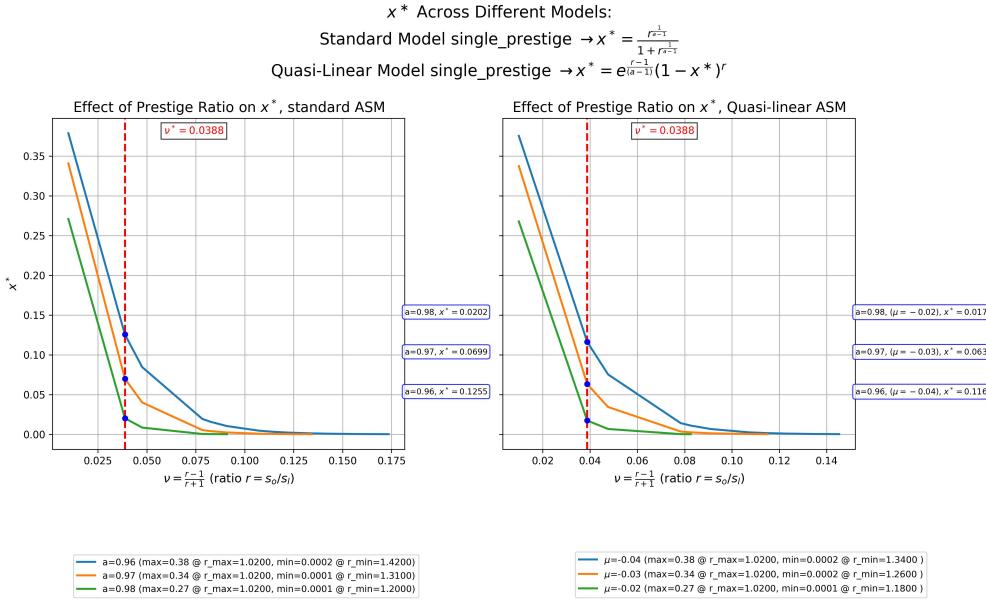


Figure 13: x^* in function of a and ν for the early period (1970-2020) and for parameters $a = 0, 9789$ and $s_l = 0.4806$. ASM quasi-linear with single prestige value and $\nu = 0.0388$

4.3 Abrams-Strogatz Model with a varying population

This section describes an extension of the Abrams-Strogatz model that includes the variation of Indigenous population I .

From Section D on data analysis, it resulted that each group, including B , follows a logistic growth pattern:

$$X(t, K_X, r_X, X_0) = \frac{K_X}{1 + \left(\frac{K_X - X_0}{X_0}\right) e^{-r_P t}} = \frac{K_X}{1 + C_X e^{-r_X t}} \quad (13)$$

Where X is either B, S or I . K_X is the carrying capacity for group X , r_X is the growth rate for group X , and X_0 is the initial size of X . It is appropriate to note that the carrying capabilities of the model are not independent since they should obey the constraint $P = I + S$ (cf. Section 3; growth rates are not independent for the same reason, and from the data results that $r_I < r_S, r_B$, and r_P .

Fig. 14 shows the percentages of Spanish, Indigenous Monolingual, and Bilingual calculated over the entire population P for 125 years.

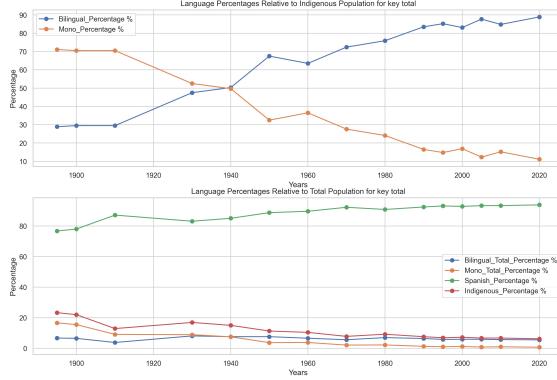


Figure 14: Abrams-Strogatz oscillation and global percentage trends for a 125-year period.

The upper image depicts a typical Abrams-Strogatz model (ASM) oscillation between the groups B and M within the Indigenous community I . If x_o represents the percentage of Indigenous Bilingual individuals, and the ASM equation is the usual:

$$\left. \frac{dx_o}{dt} \right|_{ASM} = s_o x_o^a (1 - x_o) - s_I x_o (1 - x_o)^a \quad (14)$$

The lower part of the image describes the percentage trends relative to the entire population P . Except for group S , the percentages of the other groups tend to decrease.

We observe two distinct dynamics: an Abrams-Strogatz oscillation and the overall trends of the macro groups S and I .

The objective is to determine whether the macro dynamics, specifically the metrics m_{si} and m_{pi} (detailed below), significantly influence the micro-dynamic oscillatory behavior. The metric $\mathbf{m}_{\mathbf{si}}$ represents the ratio between the number of Spanish speakers and the number of Indigenous speakers, denoted as $m_{si} = \frac{N_S}{N_I}$, where N_X is the number of elements in sets S (Spanish speakers) and I (Indigenous speakers). In a lattice of social interaction, this metric is connected to the number of Spanish-speaking neighborhoods associated with an Indigenous speaker.

Using equation (13), we can express a closed form for m_{si} :

$$m_{si}(t) = \frac{K_S}{K_I} \cdot \frac{1 + C_I e^{-r_I t}}{1 + C_S e^{-r_S t}} \quad (15)$$

The data analysis shows a positive correlation between m_{si} and the percentage of the subset B within I . Specifically, as m_{si} increases, there is also an increase in the percentage of Indigenous Bilingual speakers.

4.3.1 The Bilingual group: two distinct typologies of increments

Two distinct types of increments contribute to the variation of the Bilingual group. The first typology emerges from (14) and involves the percentage of Monolingual speakers who adopt Spanish as a second language within the Indigenous group. The second increment turns on the probability that out of n new Indigenous, k individuals are natively Bilingual. The increments of population X_o are described below:

$$X_o(t + dt) - X_o(t) = \Delta X_o \Big|_{tot} = \Delta X_o \Big|_{ASM} + \Delta X_o \Big|_{growth} \quad (16)$$

The first term is directly connected to the ASM dynamics:

$$\Delta X_o \Big|_{ASM} = \Delta(Ix_o) = I(t)\Delta x_o + x_o \underbrace{\Delta I}_{=0}$$

The second term measures the Bilingual natural increment:

$$\Delta X_o \Big|_{growth} = \Delta(p_o I) = p_o \Delta + I \Delta p_o$$

In the limit as $dt \rightarrow 0$, by combining the equation governing the ASM dynamics ((14)) with the explicit dependence of p on I , expressed as $dp/dt = (dp/dI) \cdot (dI/dt)$, and recalling that the proportion x_o is defined as X_o/I , we obtain the following formulation for the complete dynamics:

$$\begin{aligned} \frac{dx_o}{dt} \Big|_{tot} &= s_o x_o^a (1 - x_o) - s_l (1 - x_o)^a x_o \\ &+ \frac{1}{I} \left[p_o - x_o + I \frac{dp_o}{dI} \right] \frac{dI}{dt} \end{aligned} \quad (17)$$

4.4 Modeling the likelihood of a natural increase in the Bilingual group.

The probability p_o in (A.55) is modeled as a binomial likelihood. This approach assumes that each new individual in the population I independently faces a probability of becoming Bilingual, influenced by whether they are born into a Bilingual family. If we denote by p the probability of an individual joining the Bilingual subgroup B , and by $1 - p$ the probability of remaining in the Monolingual subgroup M , the number of new Bilingual individuals, k , among n new individuals follows a binomial distribution:

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

However, our interest lies in the expected proportion of individuals joining B , which is proportional to p :

$$\mathbb{E}(k) = np.$$

For a short time interval Δt , the number of new individuals in the group I can be approximated by the formula $n = \frac{dI}{dt} \Delta t$. Therefore, the expected number of new Bilinguals per unit of time is defined by this relationship.

$$\frac{\mathbb{E}(k)}{\Delta t} = p \frac{dI}{dt}.$$

This formulation expresses that the overall growth of I and the probability p of transitioning to B influence the growth of the Bilingual population. It remains to determine the functional form for p . Since m_{si} is positively correlated with the increase of Bilingual speakers, p should be an increasing monotonic function of m_{si} . We considered different functional forms for $p(m_{si}(t))$ and selected the scaled logistic as the best-fitting model. This functional form of $p(t)$ facilitates the stability, especially as $t \rightarrow \infty$.

$$p(t) = \frac{p_{max}}{1 + Ae^{-\nu m_{si}(t)}} \quad (18)$$

Where $A \approx 5.47$, $\nu \approx 0.271$, and $p_{max} \approx 0.97$.

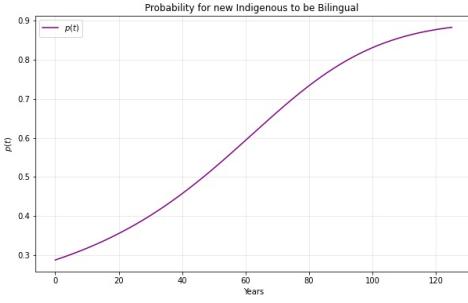


Figure 15: Probability for a new Indigenous to be Bilingual.

The exponential component, $e^{-\nu m_{si}(t)}$, stabilizes at K_S/K_I , ensuring that $p(t)$ approaches its asymptotic value without oscillation or divergence.

4.5 Fixed points and stability

The following couple of equations describe the complete dynamics of the system:

$$\begin{aligned} \frac{dx_o}{dt} \Big|_{tot} &= s_o x_o^a (1 - x_o) - s_l (1 - x_o)^a x_o \\ &\quad + \frac{1}{I} \left[p_o - x_o + I \frac{dp_o}{dI} \right] \frac{dI}{dt} \end{aligned} \quad (19a)$$

$$\frac{dI}{dt} = r_I I \left(1 - \frac{I}{K_I} \right) \quad (19b)$$

Equation (19b) has two fixed points, $I = 0$ or $I = K_I$, which, supposing $r_I > 0$, have the following stability:

$$\begin{cases} I = 0 \text{ unstable} \\ I = K_I \text{ stable} \end{cases}$$

The fixed point $I = 0$ implies both $x_o = 0$ and $x_l = 0$ and identifies the total extinction of the group of Indigenous. On the contrary, when $I = K_I$, (19a) reduces to (14) showing the same fixed points as the standard Abrams-Strogatz model:

$$\begin{cases} p_0(x_l, x_o) = (0, 1), \text{ stable if } a > 1 \text{ unstable if } a < 1 \\ p_1(x_l, x_o) = (1, 0), \text{ stable if } a > 1 \text{ unstable if } a < 1 \\ p^*(x_l^*, x_o^*) = \left(\frac{\beta}{1+\beta}, \frac{1}{1+\beta} \right) \text{ where } \beta = \left(\frac{s_o}{s_l} \right)^{\frac{1}{a-1}}, \\ \text{unstable if } a > 1 \text{ stable if } a < 1 \end{cases} \quad (20)$$

We examine the behavior of (A.55) under the conditions $I \neq 0$ and $I \neq K_I$. The term involving I :

$$\frac{1}{I} \left[p_o - x_o + I \frac{dp_o}{dI} \right] \frac{dI}{dt}$$

has a temporary effect when $I < K_I$, causing deviations in system trajectories before reaching equilibrium. The key dependencies within the system can be summarized as follows.

The system's dynamics is influenced by the interaction between internal competition, represented by ASM dynamics, and the percentage of new individuals introduced, determined by the probability p_o . In fact, during the phase when I is still increasing, the distribution of new individuals —shaped by p_o — can significantly

impact the evolution of x_o and x_l .

To study fixed points in x_o when the total population is changing, we need to solve (A.55) by setting $\frac{dx_o}{dt} = 0$.

$$\begin{aligned} & s_o x_o^a (1 - x_o) - s_l (1 - x_o)^a x_o \\ & + \frac{1}{I} \left[p_o - x_o + I \frac{dp_o}{dI} \right] \frac{dI}{dt} = 0 \end{aligned}$$

According (18), p depends on I through $m_{si} = \frac{S(t)}{I(t)}$. We use the chain rule to calculate the derivative of p_o with respect to I :

$$\frac{dp_o}{dI} = -\nu p_o \left(1 - \frac{p_o}{p_{max}} \right) \left(\frac{m_{si}}{I} \right)$$

The calculation proceeds as follows:

$$\begin{aligned} \frac{dx_o}{dt} &= s_o x_o^a (1 - x_o) - s_l (1 - x_o)^a x_o \\ &+ \frac{1}{I} \left[p_o - x_o + I \frac{dp_o}{dI} \right] \frac{dI}{dt} \\ &= s_o x_o^a (1 - x_o) - s_l (1 - x_o)^a x_o + \\ &+ \left[p_o - x_o - p_o \left(1 - \frac{p_o}{p_{max}} \right) \nu m_{si} \right] r_I \left(1 - \frac{I}{K_I} \right) \end{aligned}$$

The complete system used to analyze the fixed points is given by:

$$\left\{ \begin{array}{l} (i) \quad \frac{dx_o}{dt} = \overbrace{s_o x_o^a (1 - x_o) - s_l (1 - x_o)^a x_o}^{\text{ASM}} + \\ \qquad \underbrace{\left[p_o - x_o - p_o \left(1 - \frac{p_o}{p_{max}} \right) \nu m_{si} \right] r_I \left(1 - \frac{I}{K_I} \right)}_{\text{natural increment}}, \\ (ii) \quad \frac{dI}{dt} = r_I I \left(1 - \frac{I}{K_I} \right) \end{array} \right. \quad (21)$$

$$\left\{ \begin{array}{l} p_o = \frac{p_{max}}{1 + A e^{-\nu m_{si}}}, \\ m_{si} = \frac{S(t)}{I(t)}. \end{array} \right. \quad (22)$$

In the non-equilibrium case where $\frac{dI}{dt} \neq 0$ (i.e., $I \neq 0$ and $I \neq K_I$), determining the points at which (i) is equal to zero is intricate. To simplify the analysis, we introduce the functions $f(x_o)$, and $g(x_o, I)$:

$$\begin{aligned} f(x_o) &= s_o x_o^a (1 - x_o) - s_l (1 - x_o)^a x_o, \\ g(x_o, I) &= \left[p_o - x_o - p_o \left(1 - \frac{p_o}{p_{max}} \right) \nu m_{si} \right] \cdot \\ &\quad \cdot r_I \left(1 - \frac{I}{K_I} \right) \\ &= \underbrace{\left[p_o (1 - \nu m_{si}) + \nu \frac{m_{si}}{p_{max}} p_o^2 - x_o \right]}_{h(x_o, I)} \cdot \\ &\quad \underbrace{r_I \left(1 - \frac{I}{K_I} \right)}_{y(I)} \end{aligned}$$

which allows us to express (i) in the compact form

$$f(x_o) = -g(x_o, I) = -h(x_o, I)y(I)$$

For each value of I within the interval $(0, K_I)$, we fix $I^*(t^*)$ and determine the corresponding value of $S^*(t^*)$. The metric $m_{si}(I^*)$ from (iv) is then utilized to compute the proportion of Bilingual speakers, denoted as $p_o(I^*)$, according to (iii). Solving equation (i), where $dx_o/dt = 0$, is equivalent to identifying the family of curves in the (x_o, I^*) phase plane that satisfies these conditions:

$$f(x_o^*) = -g(x_o^*, I^*) = -h(x_o^*, I^*)y(I^*).$$

The interaction between trajectories and these curves of “quasi-equilibrium” points, where $dx_o/dt = 0$, while the group of Indigenous (I) continues to evolve ($dI/dt \neq 0$), controls the overall system dynamics.

The stability of the family of fixed points is determined by the Jacobian of the system.⁴ The elements of the Jacobian are reported below:

$$J_{00} = \frac{\partial}{\partial x_o} (f(x_o) + h(x_o, I)y(I)) = A_{00} + B_{00} + C_{00}$$

Where

$$\begin{aligned} ASM &= \begin{cases} A_{00} = (1 - 2x_o) [s_o x_o^{a-1} - s_l (1 - x_o)^{a-1}] \\ B_{00} = x_o (1 - x_o) \\ \cdot [(a-1)(s_o x_o^{a-2} + s_l (1 - x_o)^{a-2})] \end{cases} \\ nat. \ inc. &= \left\{ C_{00} = -r_I \left(1 - \frac{I}{K_I} \right) \right. \\ J_{01} &= \frac{\partial}{\partial I} (f(x_o) + h(x_o, I)y(I)) = A_{01} + B_{01} \\ &\left. \begin{cases} A_{01} = - \left[p_o \left(1 - \nu \frac{S}{I} \right) + \frac{\nu}{p_{max}} \frac{S}{I} p_o^2 - x_o \right] \cdot \frac{r_I}{K_I} \\ B_{01} = r_I \left(1 - \frac{I}{K_I} \right) \cdot \nu^2 \frac{S^2}{I^3} p_o \left(1 - \frac{p_o}{p_{max}} \right) \left(1 - 2 \frac{p_o}{p_{max}} \right) \end{cases} \right. \end{aligned}$$

In the previous calculations, we utilized the definition of $m_{si} = \frac{S}{I}$ as stated in (22). Finally, the other two elements of the Jacobian are:

$$J_{10} = \frac{\partial}{\partial x_o} \left(\frac{dI}{dt} \right) = 0, J_{11} = \frac{\partial}{\partial I} \left(\frac{dI}{dt} \right) = r_I \left(1 - 2 \frac{I}{K_I} \right)$$

4.6 Comparison of fixed points

This section compares the fixed points of the standard Abrams-Srogatz model with those derived from considering both population variation and the natural increase of Bilingual speakers. Fig. 16 reports the fitting results of census data where the volatility parameter is approximately $a \sim 1.25$, reflecting a “consensus” status.

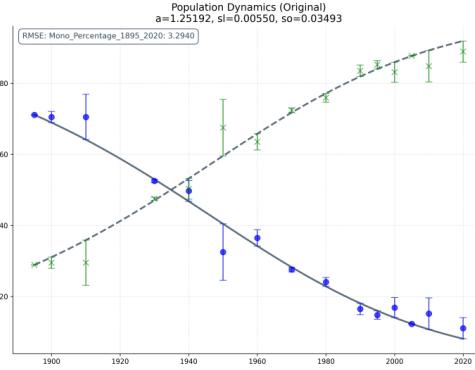


Figure 16: ASM dynamics for two the overall periods: 1895 – 2020.

⁴The complete derivation of the Jacobian is available at https://github.com/riccardodog/language_dynamics/blob/main/cist_cs_jacobian.pdf

According to (20), $p_0 = (0, 1)$ is stable, with a derivative of -0.0050 , while $p^* = (0.001, 0.999)$ is unstable, with a derivative 0.0014 .

The complete system (see (21)), ASM plus the natural increment of Bilingual speakers, exhibits stable points for $x_o < 1$. Fig. 17 displays the vector fields around a sample of 5 stable fixed points around $x_o = 1$.

The fixed point p_0 began to stabilize around $t = 2013$, with the average value for x_0 approximately 0.8720 , which is well below the threshold of $x_0 \sim 1$. The system's eigenvalues are -0.0014 and -0.0383 . The eigenvalue with the largest magnitude, 0.0383 , primarily influences the internal dynamics (ASM), while the eigenvalue with the largest real part, -0.0014 , is associated with the process of natural increment. Similarly to Fig. 19, Fig. 19 displays the vector fields around a sample of 5 stable fixed points around $x_o = 1$.

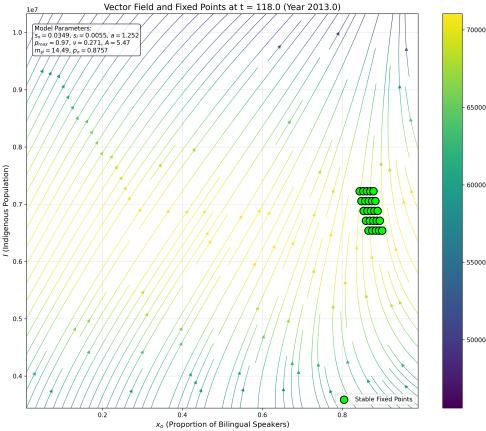


Figure 17: Vector field for x_o at $t = 2020$. Census period: 1895 – 2020.

The analysis of census data reveals that the internal dynamics of the AS model exhibit two distinct periods: 1895 – 1970 and 1970 – 2020. The first period indicates a “consensus” status, while in the second, a volatility parameter of approximately $a \sim 0.655$ reflects a “coexistence” status; see Fig. 18.



Figure 18: ASM dynamics for two distinct periods: 1895 – 1970 and 1970 – 2020.

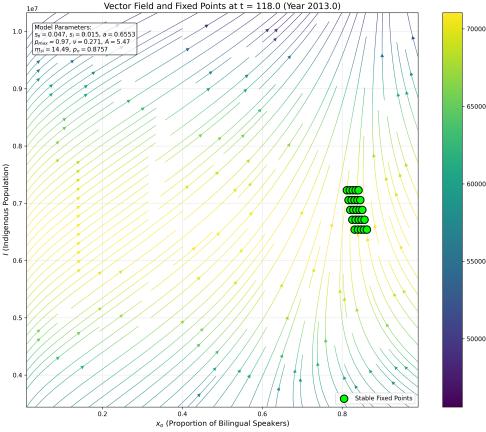


Figure 19: Vector field for x_o at $t = 2020$. Census period: 1895 – 2020.

For the time period 1970 – 2020, the mean value of x_o at $t = 2013$ is ≈ 0.8360 . By $t = 2020$, the value of x_o is 0.85, which is 4.5% lower than the actual value from the data, approximately 0.89. It is important to note that the complete system does not alter the stability of the fixed points; rather, it affects their values and, most importantly, the timing of when those values are reached.

The fixed point for “coexistence” using the standard model of Abrams-Strogatz is achieved around $t = 2700$ and approximates $x_o \approx 0.97$. In contrast, the complete system demonstrates stable “coexistence” much earlier, around $t = 2400$, and behaves as a stable system starting from $t = 2020$.

4.7 XXX

- 1 Normalization of N The normalized population N is given by:

$$\tilde{N} = \frac{N}{K}.$$

The logistic differential equation in normalized form is:

$$\frac{d\tilde{N}}{dt} = r\tilde{N}(1 - \tilde{N}),$$

where $\tilde{N} \in [0, 1]$.

- 2. **Abrams-Strogatz Dynamics: For the two subpopulations x_1 and x_2 , the Abrams-Strogatz model is:

$$\begin{aligned}\frac{dx_1}{dt} &= (1 - x_1)x_1(s_1x_1^a - s_2x_2^a), \\ \frac{dx_2}{dt} &= (1 - x_2)x_2(s_2x_2^a - s_1x_1^a),\end{aligned}$$

where: - s_1 and s_2 are the status factors, - a is the exponent capturing nonlinear influence.

- 3. **Combining Abrams-Strogatz with Logistic Growth The total population N evolves according to the logistic equation:

$$\frac{dN}{dt} = rN(1 - \frac{N}{K}),$$

or, in normalized form:

$$\frac{d\tilde{N}}{dt} = r\tilde{N}(1 - \tilde{N}).$$

The subpopulations x_1 and x_2 are expressed as fractions of N :

$$x_1 = p_1N, \quad x_2 = p_2N, \quad \text{with } p_1 + p_2 = 1.$$

Using the chain rule:

$$\begin{aligned}\frac{dx_1}{dt} &= \frac{dp_1}{dt}N + p_1\frac{dN}{dt}, \\ \frac{dx_2}{dt} &= \frac{dp_2}{dt}N + p_2\frac{dN}{dt}.\end{aligned}$$

- 4. **Dynamics of p_1 and p_2 The proportions p_1 and p_2 evolve as:

$$\begin{aligned}\frac{dp_1}{dt} &= (1 - p_1)p_1(s_1p_1^a - s_2p_2^a), \\ \frac{dp_2}{dt} &= (1 - p_2)p_2(s_2p_2^a - s_1p_1^a).\end{aligned}$$

- 5. **Final Differential Equations for x_1 and x_2 : Substituting the dynamics of p_1 , p_2 , and N into the expressions for x_1 and x_2 :

$$\begin{aligned}\frac{dx_1}{dt} &= [(1 - p_1)p_1(s_1p_1^a - s_2p_2^a)]N + p_1rN\left(1 - \frac{N}{K}\right), \\ \frac{dx_2}{dt} &= [(1 - p_2)p_2(s_2p_2^a - s_1p_1^a)]N + p_2rN\left(1 - \frac{N}{K}\right).\end{aligned}$$

Here: - $p_1 = \frac{x_1}{N}$, $p_2 = \frac{x_2}{N}$, - $p_1 + p_2 = 1$, - N evolves according to the logistic equation.

This completes the normalized logistic growth and Abrams-Strogatz framework.

$$\frac{dB}{dt} = (I - B)B^a s_B - B(I - B)^a s_M + \alpha \frac{dI}{dt}$$

Abram-Strogatz Model Derivation

Starting with the original model in terms of population fractions:

$$\frac{dx_1}{dt} = (1 - x_1)s_1x_1^a - x_1s_2x_2^a \quad (23)$$

Converting to absolute populations where $N_1 + N_2 = P(t)$:

$$\begin{aligned}x_1 &= \frac{N_1}{P(t)} \\ x_2 &= \frac{N_2}{P(t)} = \frac{P(t) - N_1}{P(t)}\end{aligned}$$

Substituting these into the original equation:

$$\frac{d}{dt}\left(\frac{N_1}{P(t)}\right) = \left(1 - \frac{N_1}{P(t)}\right)s_1\left(\frac{N_1}{P(t)}\right)^a - \left(\frac{N_1}{P(t)}\right)s_2\left(\frac{N_2}{P(t)}\right)^a \quad (24)$$

Using the chain rule on the left side:

$$\frac{1}{P(t)}\frac{dN_1}{dt} - \frac{N_1}{P(t)^2}\frac{dP}{dt} = \left(1 - \frac{N_1}{P(t)}\right)s_1\left(\frac{N_1}{P(t)}\right)^a - \left(\frac{N_1}{P(t)}\right)s_2\left(\frac{N_2}{P(t)}\right)^a \quad (25)$$

Multiplying all terms by $P(t)$:

$$\frac{dN_1}{dt} - \frac{N_1}{P(t)}\frac{dP}{dt} = (P(t) - N_1)s_1\left(\frac{N_1}{P(t)}\right)^a - N_1s_2\left(\frac{N_2}{P(t)}\right)^a \quad (26)$$

Final form of the differential equation:

$$\frac{dN_1}{dt} = (P(t) - N_1)s_1 \left(\frac{N_1}{P(t)} \right)^a - N_1 s_2 \left(\frac{N_2}{P(t)} \right)^a + \frac{N_1}{P(t)} \frac{dP}{dt} \quad (27)$$

where:

- N_1, N_2 are the populations of speakers of each language
- $P(t)$ is the total population function
- $\frac{dP}{dt}$ is the rate of change of total population
- s_1, s_2 are the status (prestige) parameters of each language
- a is the volatility parameter

Additional Material

A A comprehensive discussion on the Abrams-Strogatz Model

A.1 Fixed points of Abrams-Strogatz model

The model's differential equations are as follows:

$$\frac{dx_l}{dt} = s_l x_l^a x_o - s_o x_o^a x_l \quad (\text{A.1a})$$

$$\frac{dx_o}{dt} = -s_l x_l^a x_o + s_o x_o^a x_l \quad (\text{A.1b})$$

At fixed points, both derivatives are zero:

$$\frac{dx_l}{dt} = s_l x_l^a x_o - s_o x_o^a x_l = 0 \quad (\text{A.2a})$$

$$\frac{dx_o}{dt} = -s_l x_l^a x_o + s_o x_o^a x_l = 0 \quad (\text{A.2b})$$

$$(\text{A.2c})$$

We begin by factorizing the first equation (A.2a) to obtain:

$$s_l x_l^a x_o - s_o x_o^a x_l = x_l x_o (s_l x_l^{a-1} - s_o x_o^{a-1}) = 0 \quad (\text{A.3})$$

The equation (A.3) identifies three fixed points.

$$p(x_l, x_o) = \begin{cases} p_0(x_l, x_o) = (0, 1) & \text{corresponding to } x_l = 0 \text{ and } x_o = 1 \\ p_1(x_l, x_o) = (1, 0) & \text{corresponding to } x_l = 1 \text{ and } x_o = 0 \\ p^*(x_l, x_o) = (x_l^*, x_o^*) & \end{cases} \quad (\text{A.4})$$

The p^* fixed point is determined using the factor from (A.3):

$$\frac{x_l^{a-1}}{x_o^{a-1}} = \frac{s_o}{s_l} = r.$$

We define β as the $(a-1)$ -th root of r , which can be expressed as $\beta = r^{\frac{1}{a-1}}$, so:

$$x_l = \beta x_o.$$

Using the normalization constraint $x_l + x_o = 1$, we have:

$$p^*(x_l^*, x_o^*) = \left(\frac{\beta}{\beta + 1}, \frac{1}{\beta + 1} \right) \quad (\text{A.5})$$

$$x_l = \frac{\beta}{\beta + 1}.$$

Therefore, the fixed points are:

$$x_l^* = \frac{\beta}{\beta + 1}, \quad x_o^* = \frac{1}{\beta + 1},$$

where $r = \frac{s_o}{s_l} > 1$ and $a > 0$. Please note that for prestige factors $s_l = s, s_o = 1 - s$, the value of β is as follows:

$$\beta = \left(\frac{1-s}{s} \right)^{\frac{1}{a-1}}$$

In conclusion, p_0 and p_1 represent a “consensus” status where either x_o or x_l populations survive. In contrast, the third case, denoted by p^* , corresponds to a non-trivial fixed point where both x_l and x_o populations are positive. This scenario is referred to as “coexistence”.

A.2 Stability Analysis of the Abrams-Strogatz Model

This section discusses the stability of the fixed points of the Abrams-Strogatz model. The stability of the three fixed points mentioned in (A.4) relies on s_l , s_o , and predominantly on a .

- for $a > 1$, p_0 and p_1 are stable, while p^* is unstable, leading to “consensus” status with the dominance of either x_l or x_o ;
- For $a < 1$, p_0 and p_1 become unstable and one “coexistence” status is possible. Indeed, p^* is stable.

After completing the calculation, we determined that the eigenvalues influencing the stability of the fixed points are as follows:

$$\lambda(s_l, s_o, a) = \begin{cases} \lambda_o \approx -s_o + s_l a \varepsilon^{a-1} & \text{corresponding to } p_0 \\ \lambda_l \approx -s_l + s_o a \varepsilon^{a-1} & \text{corresponding to } p_1 \\ \lambda_* \approx (a-1)s_o \left(\frac{1}{1+\beta}\right)^{a-1} \left[1 + a \varepsilon \frac{1 - \left(\frac{s_o}{s_l}\right)^{\frac{2}{a-1}}}{\left(\frac{s_l}{s_o}\right)^{\frac{1}{a-1}}} \right] & \text{corresponding to } p^* \end{cases} \quad (\text{A.6})$$

We derive equation (A.6) using three distinct approaches. The first approach employs the Jacobian formulation, which is the most general method available. The second approach simplifies the pair of differential equations for x_l and x_o by applying the normalization constraint $x_o = 1 - x_l$. The final approach introduces the concept of magnetization, defined as $m = x_o - x_l$.

A.2.1 Jacobian approach

The equation referenced in (A.1) is equivalent to the equation below, (A.7).

$$\frac{dx_l}{dt} = f(x_l, x_o) \quad (\text{A.7a})$$

$$\frac{dx_o}{dt} = g(x_l, x_o) \text{ where} \quad (\text{A.7b})$$

$$f(x_l, x_o) = s_l x_l^a x_o - s_o x_o^a x_l \quad (\text{A.7c})$$

$$g(x_l, x_o) = -s_l x_l^a x_o + s_o x_o^a x_l \quad (\text{A.7d})$$

The Jacobian Matrix is

$$J = \begin{bmatrix} \frac{\partial f}{\partial x_l} & \frac{\partial f}{\partial x_o} \\ \frac{\partial g}{\partial x_l} & \frac{\partial g}{\partial x_o} \end{bmatrix}.$$

The partial derivatives:

$$\frac{\partial f}{\partial x_l} = a s_l x_l^{a-1} x_o - s_o x_o^a, \quad \frac{\partial f}{\partial x_o} = s_l x_l^a - a s_o x_o^{a-1} x_l,$$

$$\frac{\partial g}{\partial x_l} = -a s_l x_l^{a-1} x_o + s_o x_o^a, \quad \frac{\partial g}{\partial x_o} = -s_l x_l^a + a s_o x_o^{a-1} x_l.$$

By utilizing the partial derivatives, we can determine the Jacobian as follows:

$$J = \begin{bmatrix} a s_l x_l^{a-1} x_o - s_o x_o^a & s_l x_l^a - a s_o x_o^{a-1} x_l \\ -a s_l x_l^{a-1} x_o + s_o x_o^a & -s_l x_l^a + a s_o x_o^{a-1} x_l \end{bmatrix}. \quad (\text{A.8})$$

The conservation law $x_l + x_o = 1$ reduces the dimension of the system. As a result, the Jacobian matrix is singular, with the determinant $\text{Det}(J) = 0$. Its eigenvalues are $\lambda_0 = 0$ and $\lambda_1 = \text{Tr}(J)$.

At p_0 (where x_o is fully dominant), the trace of the Jacobian is calculated using $x_l = \varepsilon$ and $x_o = 1 - \varepsilon$.

$$J = \begin{bmatrix} a s_l \varepsilon^{a-1} (1 - \varepsilon) - s_o (1 - \varepsilon)^a & [.] \\ [.] & -s_l \varepsilon^a + a s_o (1 - \varepsilon)^{a-1} \varepsilon \end{bmatrix}.$$

The trace $Tr(J)$ assumes the following values:

$$\begin{aligned}
Tr(J) \Big|_{p_0=(\varepsilon, 1-\varepsilon)} &= J_{00} + J_{11} \\
&= as_l\varepsilon^{a-1}(1-\varepsilon) - s_o(1-\varepsilon)^a - s_l\varepsilon^a + as_o(1-\varepsilon)^{a-1}\varepsilon \\
&\approx s_l(a(1-\varepsilon)\varepsilon^{a-1} + \varepsilon^a) - s_o((1-\varepsilon)^a - a(1-\varepsilon)^{a-1}\varepsilon) \\
&\approx s_l\varepsilon^{a-1}(a(1-\varepsilon) - \varepsilon) - s_o(1-\varepsilon)^{a-1}((1-\varepsilon) - a\varepsilon) \\
&\approx -s_o + as_l\varepsilon^{a-1}
\end{aligned} \tag{A.9}$$

The eigenvectors are:

$$\begin{cases} v_0 = [0, 1] & \text{corresponding to } \lambda_0 = 0 \\ v_1 = [-1, 1] & \text{corresponding to } \lambda_1 = -s_o + as_l\varepsilon^{a-1} \end{cases} \tag{A.10}$$

In the direction of v_0 , the solutions remain constant, indicating that the x_o axis is a stable direction. In contrast, in the direction of v_1 , solutions decay to zero at a rate of $e^{-s_o t}$ if $a > 1$ and any initial condition will eventually approach the direction of v_0 as the components corresponding to v_1 decay. On the contrary, when $a < 1$ the rate $e^{\lambda_1 t}$ starts increasing for $\varepsilon < \left(\frac{as_l}{s_o}\right)^{\frac{1}{1-a}}$ and the direction v_1 is unstable.

Similarly, at p_1 (full dominance of x_l) the trace of the Jacobian is calculated with $x_l = 1 - \varepsilon$ and $x_o = \varepsilon$:

$$J = \begin{bmatrix} as_l(1-\varepsilon)^{a-1}\varepsilon - s_o\varepsilon^a & [.] \\ [.] & -s_l(1-\varepsilon)^a + as_o\varepsilon^{a-1}(1-\varepsilon) \end{bmatrix}.$$

The trace $Tr(J)$ assumes the following values:

$$\begin{aligned}
Tr(J) \Big|_{p_0=(1-\varepsilon, \varepsilon)} &= J_{00} + J_{11} \\
&= as_l(1-\varepsilon)^{a-1}\varepsilon - s_o\varepsilon^a - s_l(1-\varepsilon)^a + as_o\varepsilon^{a-1}(1-\varepsilon) \\
&\approx s_l(1-\varepsilon)^{a-1}(a\varepsilon - (1-\varepsilon)) - s_o\varepsilon^{a-1}(\varepsilon - as_o(1-\varepsilon)) \\
&\approx -s_l + as_o\varepsilon^{a-1}
\end{aligned} \tag{A.11}$$

The eigenvectors are:

$$\begin{cases} w_0 = [1, 1] & \text{corresponding to } \lambda_0 = 0 \\ w_1 = [-1, 1] & \text{corresponding to } \lambda_1 = -s_l + as_o\varepsilon^{a-1} \end{cases} \tag{A.12}$$

The solutions remain constant in the direction of w_0 , making the x_l axis stable. In contrast, in the direction of w_1 , the solutions decay to zero at a rate of $e^{-s_l t}$ if $a > 1$ and, regardless of the initial condition, the system will eventually approach the stable direction w_0 as the other component decays. On the contrary, when $a < 1$, the rate $e^{\lambda_1 t}$ starts increasing for $\varepsilon < \left(\frac{as_o}{s_l}\right)^{\frac{1}{a-1}}$ and the direction w_1 is unstable.

Finally, the eigenvalue of p^* is

$$\begin{aligned}
Tr(J) \Big|_{p^*=(x_l^*, x_o^*)} &= J_{00} + J_{11} \approx \frac{(a-1)s_o}{\left(1 + \left(\frac{s_o}{s_l}\right)^{\frac{1}{a-1}}\right)^{a-1}} + as_o(a-1)\varepsilon \left(\frac{1}{\left(1 + \left(\frac{s_o}{s_l}\right)^{\frac{1}{a-1}}\right)^{a-1}} \right) \left[\frac{1}{\beta} - \beta \right] \\
&\approx s_o(a-1) \left(\frac{1}{\left(1 + \left(\frac{s_o}{s_l}\right)^{\frac{1}{a-1}}\right)} \right)^{a-1} \left[1 + a\varepsilon \left(\frac{1 - \left(\frac{s_o}{s_l}\right)^{\frac{2}{a-1}}}{\left(\frac{s_o}{s_l}\right)^{\frac{1}{a-1}}} \right) \right]
\end{aligned} \tag{A.13}$$

The fixed point p^* results stable if $a < 1$ and unstable if $a > 1$.

In the following, we will analyze how J_{00} and J_{11} depend on ε to ensure that neither exhibits divergences that could potentially affect the stability of p^* . We know that:

- The variable r measures the ratio between s_o and s_l : $r = \frac{s_o}{s_l}$;
- The variable β represents the ratio between the populations x_l and x_o : $\beta = \frac{x_l}{x_o}$;
- β and r are related: $\beta^{a-1} = r$;
- The fixed point p^*, x_l^*, x_o^* in terms of β is $x_l^* = \frac{\beta}{\beta+1}$, $x_o^* = \frac{1}{\beta+1}$.

The Derivation of $J_{00}(x_l^*, x_o^*) \approx J_{00}^{(0)}(x_l^*, x_o^*) + J_{00}^{(1)}(x_l^*, x_o^*, \varepsilon)$ is as follows:
First, we report the definition of J_{00} :

$$J_{00}(x_l, x_o) = as_l x_l^{a-1} x_o - s_o x_o^a$$

Next, we apply the transformation $x_l \rightarrow x_l^* + \varepsilon$. Given the normalization constraint, it follows that $x_o \rightarrow x_o^* - \varepsilon$. Expanding to first order in ε :

$$\begin{aligned} J_{00} &= as_l(x_l^* + \varepsilon)^{a-1}(x_o^* - \varepsilon) - s_o(x_o^* - \varepsilon)^a \\ &= as_l(x_l^*)^{a-1} \left(1 + \frac{\varepsilon}{x_l^*}\right)^{a-1} x_o^* \left(1 - \frac{\varepsilon}{x_o^*}\right) - s_o(x_o^*)^a \left(1 - \frac{\varepsilon}{x_o^*}\right)^a \\ &\approx as_l(x_l^*)^{a-1} x_o^* \left(1 + \frac{a-1}{x_l^*} \varepsilon\right) \left(1 - \frac{\varepsilon}{x_o^*}\right) - s_o(x_o^*)^a \left(1 - \frac{a}{x_o^*} \varepsilon\right) \end{aligned}$$

As the next step, we will collect the terms corresponding to the orders ε^0 and ε^1 :

$$\begin{aligned} J_{00} &\approx as_l(x_l^*)^{a-1} x_o^* \left(1 + \frac{a-1}{x_l^*} \varepsilon\right) \left(1 - \frac{\varepsilon}{x_o^*}\right) - s_o(x_o^*)^a \left(1 - \frac{a}{x_o^*} \varepsilon\right) \\ &= as_l(x_l^*)^{a-1} x_o^* \left[1 + \varepsilon \left(\frac{a-1}{x_l^*} - \frac{1}{x_o^*}\right)\right] - s_o(x_o^*)^a + as_o(x_o^*)^{a-1} \varepsilon \\ &= \underbrace{as_l(x_l^*)^{a-1} x_o^* - s_o(x_o^*)^a}_{\text{order } \varepsilon^0} + \underbrace{\left[as_l(x_l^*)^{a-1} x_o^* \left(\frac{a-1}{x_l^*} - \frac{1}{x_o^*}\right) + as_o(x_o^*)^{a-1}\right]}_{\text{order } \varepsilon^1} \varepsilon \end{aligned}$$

Finally, we set:

$J_{00} = J_{00}^{(0)} + J_{00}^{(1)}$, where $J_{00}^{(0)} = as_l(x_l^*)^{a-1} x_o^* - s_o(x_o^*)^a$. For $J_{00}^{(1)}$ we proceed in the following manner:

$$\begin{aligned} J_{00}^{(1)} &= as_l(x_l^*)^{a-1} x_o^* \left(\frac{a-1}{x_l^*} - \frac{1}{x_o^*}\right) + as_o(x_o^*)^{a-1} = (a-1) as_l(x_l^*)^{a-2} x_o^* + \\ &- a \underbrace{(s_l(x_l^*)^{a-1} - s_o(x_o^*)^{a-1})}_{=0 \text{ in } p^*(x_l^*, x_o^*)} \end{aligned}$$

In conclusion:

$$J_{00} = as_l(x_l^*)^{a-1} x_o^* - s_o(x_o^*)^a + [(a-1) as_l(x_l^*)^{a-2} x_o^*] \varepsilon$$

Similarly, we calculate the derivation of $J_{11}(x_l^*, x_o^*) \approx J_{11}^{(0)}(x_l^*, x_o^*) + J_{11}^{(1)}(x_l^*, x_o^*, \varepsilon)$: We start from the definition of J_{11} :

$$J_{11}(x_l, x_o) = -s_l x_l^a + as_o x_o^{a-1} x_l$$

Next, we apply the transformation $x_l \rightarrow x_l^* + \varepsilon$. Given the normalization constraint, it follows that $x_o \rightarrow x_o^* - \varepsilon$, expanding to first order in ε .

$$\begin{aligned} J_{11} &= -s_l(x_l^* + \varepsilon)^a + as_o(x_o^* - \varepsilon)^{a-1}(x_l^* + \varepsilon) \\ &= -s_l(x_l^*)^a \left(1 + \frac{\varepsilon}{x_l^*}\right)^a + as_o(x_o^*)^{a-1} \left(1 - \frac{\varepsilon}{x_o^*}\right)^{a-1} x_l^* \left(1 + \frac{\varepsilon}{x_l^*}\right) \\ &\approx -s_l(x_l^*)^a \left(1 + \frac{a}{x_l^*} \varepsilon\right) + as_o(x_o^*)^{a-1} x_l^* \left(1 - \frac{a-1}{x_o^*} \varepsilon\right) \left(1 + \frac{\varepsilon}{x_l^*}\right) \end{aligned}$$

As the next step, we will collect the terms corresponding to the orders ε^0 and ε^1 :

$$\begin{aligned}
J_{11} &\approx -s_l(x_l^*)^a \left(1 + \frac{a}{x_l^*} \varepsilon\right) + as_o(x_o^*)^{a-1} x_l^* \left(1 - \frac{a-1}{x_o^*} \varepsilon\right) \left(1 + \frac{\varepsilon}{x_l^*}\right) \\
&= -s_l(x_l^*)^a - as_l(x_l^*)^{a-1} \varepsilon + as_o(x_o^*)^{a-1} x_l^* \left[1 + \left(\frac{1}{x_l^*} - \frac{a-1}{x_o^*}\right) \varepsilon\right] \\
&= \underbrace{-s_l(x_l^*)^a + as_o(x_o^*)^{a-1} x_l^*}_{\text{order } \varepsilon^0} + \underbrace{\left[-as_l(x_l^*)^{a-1} + as_o(x_o^*)^{a-1} x_l^* \left(\frac{1}{x_l^*} - \frac{a-1}{x_o^*}\right)\right]}_{\text{order } \varepsilon^1} \varepsilon
\end{aligned}$$

Finally, we set:

$J_{11} = J_{11}^{(0)} + J_{11}^{(1)}$, where $J_{11}^{(0)} = -s_l(x_l^*)^a + as_o(x_o^*)^{a-1} x_l^*$. For $J_{11}^{(1)}$ we proceed in the following manner:

$$\begin{aligned}
J_{11}^{(1)} &= -as_l(x_l^*)^{a-1} + as_o(x_o^*)^{a-1} x_l^* \left(\frac{1}{x_l^*} - \frac{a-1}{x_o^*}\right) = \\
&= -as_l(x_l^*)^{a-1} + as_o(x_o^*)^{a-1} - (a-1) as_o(x_o^*)^{a-2} x_l^* = \\
&= \underbrace{-a(s_l(x_l^*)^{a-1} - s_o(x_o^*)^{a-1})}_{=0 \text{ in } p^*(x_l^*, x_o^*)} - (a-1) as_o(x_o^*)^{a-2} x_l^*
\end{aligned}$$

In conclusion:

$$J_{11} = -s_l(x_l^*)^a + as_o(x_o^*)^{a-1} x_l^* - [(a-1) as_o(x_o^*)^{a-2} x_l^*] \varepsilon$$

Finally, we calculate the trace of the Jacobian from the previous results:

$$Tr(J) = J_{00} + J_{11} = Tr(J^{(0)}) + Tr(J^{(1)}) = J_{00}^{(0)} + J_{11}^{(0)} + J_{00}^{(1)} + J_{11}^{(1)}$$

The first step is to sum $J_{00}^{(0)}$ and $J_{11}^{(0)}$, where we put $x_i^* \rightarrow x_i$ for clarity.

$$\begin{aligned}
J_{00}^{(0)} &= as_l x_l^{a-1} x_o - s_o x_o^a \\
J_{11}^{(0)} &= -s_l x_l^a + as_o x_o^{a-1} x_l \\
J_{00}^{(0)} + J_{11}^{(0)} &= as_l x_l^{a-1} x_o - s_o x_o^a - s_l x_l^a + as_o x_o^{a-1} x_l = \\
&= ax_l x_o (s_l x_l^{a-2} + s_o x_o^{a-2}) - (s_l x_l^a + s_o x_o^a) \\
&= ax_l x_o^{a-1} \left(s_l \left(\frac{x_l}{x_o}\right)^{a-2} + s_o\right) - x_o^a \left(s_l \left(\frac{x_l}{x_o}\right)^a + s_o\right)
\end{aligned}$$

Following the definition of β, r, x_l , and x_o :

$$\begin{aligned}
J_{00}^{(0)} + J_{11}^{(0)} &= ax_l x_o^{a-1} \left(s_l \left(\frac{x_l}{x_o}\right)^{a-2} + s_o\right) - x_o^a \left(s_l \left(\frac{x_l}{x_o}\right)^a + s_o\right) \\
&= a \frac{\beta}{1+\beta} \left(\frac{1}{1+\beta}\right)^{a-1} \left(s_l \frac{r}{\beta} + s_o\right) - \left(\frac{1}{1+\beta}\right)^a (s_l r \beta + s_o) \\
&= as_o \frac{\beta}{1+\beta} \left(\frac{1}{1+\beta}\right)^{a-1} \left(\frac{1}{\beta} + 1\right) - s_o \left(\frac{1}{1+\beta}\right)^a (\beta + 1) \\
&= (a-1) s_o \left(\frac{1}{1+\beta}\right)^{a-1}
\end{aligned}$$

Similarly, we proceed with the first order in ε , $J_{00}^{(1)}$ and $J_{11}^{(1)}$

$$\begin{aligned}
J_{00}^{(1)} &= (a-1)as_lx_l^{a-2}x_o \\
J_{11}^{(1)} &= -(a-1)as_ox_o^{a-2}x_l \\
&= a(a-1)x_lx_o(s_lx_l^{a-3} - s_o^{a-3}) \\
&= a(a-1)x_lx_o^{a-2}\left(s_l\left(\frac{x_l}{x_o}\right)^{a-3} - s_o\right) = a(a-1)x_lx_o^{a-2}\left(s_l\left(\frac{r}{\beta^2}\right) - s_o\right) \\
&= a(a-1)s_ox_lx_o^{a-2}\left(\frac{1}{\beta^2} - 1\right) \\
&= a(a-1)s_o\frac{\beta}{1+\beta}\left(\frac{1}{1+\beta}\right)^{a-2}\left(\frac{1}{\beta^2} - 1\right) = a(a-1)s_o\left(\frac{1}{1+\beta}\right)^{a-1}\left(\frac{1}{\beta} - \beta\right)
\end{aligned}$$

In conclusion:

$$Tr(J) = (a-1)s_o\left(\frac{1}{1+\beta}\right)^{a-1}\left[1 + a\varepsilon\left(\frac{1}{\beta} - \beta\right)\right]$$

Since $\beta = \left(\frac{s_o}{s_l}\right)^{\frac{1}{a-1}}$

$$Tr(J) = (a-1)s_o\left(\frac{1}{1 + \left(\frac{s_o}{s_l}\right)^{\frac{1}{a-1}}}\right)^{a-1}\left[1 + a\varepsilon\left(\frac{1 - \left(\frac{s_o}{s_l}\right)^{\frac{2}{a-1}}}{\left(\frac{s_o}{s_l}\right)^{\frac{1}{a-1}}}\right)\right]$$

A.2.2 Reduced Ordinary Differential Equation (ODE)

We begin with the simplified differential equation for the Abrams-Strogatz model:

$$\frac{dx}{dt} = s_lx^a(1-x) - s_ox(1-x)^a = (x-x^2)(s_lx^{a-1} - s_o(1-x)^{a-1}) \quad (\text{A.14})$$

The next step is to calculate the derivative, which includes two terms:

$$\begin{aligned}
T_1 &= (x-x^2) \\
T_2 &= (s_lx^{a-1} - s_o(1-x)^{a-1}) \text{ so that:} \\
\frac{d}{dx}\left(\frac{dx}{dt}\right) &= \frac{dT_1}{dx}T_2 + \frac{dT_2}{dx}T_1 \\
\frac{dT_1}{dx} &= 1 - 2x \\
\frac{dT_2}{dx} &= (a-1)(s_lx^{a-2} + s_o(1-x)^{a-2})
\end{aligned}$$

The first fixed point we investigate is p_o with the coordinates $(\varepsilon, 1-\varepsilon)$. We then evaluate the ODE in $x = \varepsilon$.

$$\begin{aligned}
\frac{d}{dx}\left(\frac{dx}{dt}\right) &= \frac{dT_1}{dx}T_2 + \frac{dT_2}{dx}T_1 \\
&= (1-2\varepsilon)(s_lx^{a-1} - s_o(1-x)^{a-1}) + (a-1)(x-\varepsilon^2)(s_lx^{a-2} + s_o(1-x)^{a-2}) \\
&= (1-2\varepsilon)(s_l\varepsilon^{a-1} - s_o(1-\varepsilon)^{a-1}) + (a-1)\varepsilon(1-\varepsilon)(s_l\varepsilon^{a-2} + s_o(1-\varepsilon)^{a-2}) \\
&\approx (1-2\varepsilon)(s_l\varepsilon^{a-1} - s_o) + (a-1)\varepsilon(1-\varepsilon)(s_l\varepsilon^{a-2} + s_o) \\
&= s_l\varepsilon^{a-1} - s_o - \underbrace{s_l\varepsilon^a + 2s_o\varepsilon}_{\rightarrow 0} + (a-1)\left(s_l\varepsilon^{a-1} + \underbrace{s_o\varepsilon - s_l\varepsilon^a - s_o\varepsilon^2}_{\rightarrow 0}\right) \\
&= -s_o + as_l\varepsilon^{a-1}
\end{aligned}$$

The second fixed point is p_1 with the coordinates $(1 - \varepsilon, \varepsilon)$. We then evaluate the ODE in $x = 1 - \varepsilon$.

$$\begin{aligned}
\frac{d}{dx} \left(\frac{dx}{dt} \right) &= \frac{dT_1}{dx} T_2 + \frac{dT_2}{dx} T_1 \\
&= (1 - 2x) (s_l x^{a-1} - s_o (1 - x)^{a-1}) + (a - 1)(x - x^2) (s_l x^{a-2} + s_o (1 - x)^{a-2}) \\
&= (1 - 2(1 - \varepsilon)) (s_l (1 - \varepsilon)^{a-1} - s_o (1 - (1 - \varepsilon))^{a-1}) + \\
&\quad + (a - 1)((1 - \varepsilon) - (1 - \varepsilon)^2) (s_l (1 - \varepsilon)^{a-2} + s_o (1 - (1 - \varepsilon))^{a-2}) \\
&\quad (2\varepsilon - 1) (s_l (1 - \varepsilon)^{a-1} - s_o \varepsilon^{a-1}) + (a - 1)(1 - \varepsilon - 1 + 2\varepsilon - \varepsilon^2) (s_l (1 - \varepsilon)^{a-2} + s_o \varepsilon^{a-2}) \\
&\approx (2\varepsilon - 1) (s_l - s_o \varepsilon^{a-1}) + (a - 1) (\varepsilon - \varepsilon^2) (s_l + s_o \varepsilon^{a-2}) \\
&= -s_l + s_o \varepsilon^{a-1} + \underbrace{2\varepsilon s_l - s_o \varepsilon^a}_{\rightarrow 0} + (a - 1) \left(\underbrace{s_l \varepsilon - s_l \varepsilon^2 - s_o \varepsilon^a}_{\rightarrow 0} + s_o \varepsilon^{a-1} \right) \\
&= -s_l + a s_o \varepsilon^{a-1}
\end{aligned}$$

Finally, we calculate $p^*(x^*, 1 - x^*)$. We evaluate the the ODE in $(x + \varepsilon, 1 - x - \varepsilon)$. The first term is as follows:

$$\begin{aligned}
\frac{dT_1}{dx} T_2 &= (1 - 2x) (s_l x^{a-1} - s_o (1 - x)^{a-1}) \\
&= (1 - 2(x + \varepsilon)) \left[s_l (x + \varepsilon)^{a-1} - s_o (1 - x - \varepsilon)^{a-1} \right] \\
&= (1 - 2x - 2\varepsilon) \left[s_l x^{a-1} \left(1 + \frac{\varepsilon}{x} \right)^{a-1} - s_o (1 - x)^{a-1} \left(1 - \frac{\varepsilon}{1-x} \right)^{a-1} \right] \\
&\approx (1 - 2x - 2\varepsilon) \left[s_l x^{a-1} \left(1 + \frac{a-1}{x} \right) \varepsilon - s_o (1 - x)^{a-1} \left(1 - \frac{a-1}{1-x} \right) \varepsilon \right] \\
&= (1 - 2x - 2\varepsilon) \left[\underbrace{s_l x^{a-1} - s_o (1 - x)^{a-1}}_{=0} + (a - 1) (s_l x^{a-2} + s_o (1 - x)^{a-2}) \varepsilon \right] \\
&\approx (1 - 2x) \left[(a - 1) (s_l x^{a-2} + s_o (1 - x)^{a-2}) \varepsilon \right] + \mathcal{O}(\varepsilon^2)
\end{aligned}$$

Similarly, the second term:

$$\begin{aligned}
\frac{dT_2}{dx} T_1 &= (a - 1)(x - x^2) (s_l x^{a-2} + s_o (1 - x)^{a-2}) \\
&= (a - 1) (x + \varepsilon - (x^2 - 2x\varepsilon + \varepsilon^2)) \left[s_l (x + \varepsilon)^{a-2} + s_o (1 - x - \varepsilon)^{a-2} \right] \\
&= (a - 1) (x - x^2 + \varepsilon (1 - 2x) + \varepsilon^2) \left[s_l x^{a-2} \left(1 + \frac{\varepsilon}{x} \right)^{a-2} + s_o (1 - x)^{a-2} \left(1 - \frac{\varepsilon}{1-x} \right)^{a-2} \right] \\
&\approx (a - 1) (x - x^2 + \varepsilon (1 - 2x)) \left[s_l x^{a-2} \left(1 + \frac{a-2}{x} \right) \varepsilon + s_o (1 - x)^{a-2} \left(1 - \frac{a-2}{1-x} \right) \varepsilon \right] \\
&= (a - 1) \left\{ \underbrace{(x - x^2) (s_l x^{a-2} + s_o (1 - x)^{a-2})}_{\varepsilon^0} + \right. \\
&\quad \left. \underbrace{[(1 - 2x) (s_l x^{a-2} + s_o (1 - x)^{a-2} + (a - 2) (x - x^2) (s_l x^{a-3} - s_o (1 - x)^{a-3}))]}_{\varepsilon^1} \varepsilon \right\} + \mathcal{O}(\varepsilon^2)
\end{aligned}$$

At the zero order in ε , recalling the definition of $x = \beta/(1 - \beta)$ and $r = s_o/s_l$, the ODE becomes:

$$\begin{aligned} \frac{d}{dx} \left(\frac{dx}{dt} \right)^{(0)} &= (a-1) (x - x^2) \left(s_l x^{a-2} + s_o (1-x)^{a-2} \right) \\ &= (a-1) x (1-x) (1-x)^{a-2} \left(s_l \left(\frac{x}{1-x} \right)^{a-2} + s_o \right) \\ &= (a-1) \frac{\beta}{1+\beta} \left(1 - \frac{\beta}{1+\beta} \right)^{a-1} \left(s_l \frac{r}{\beta} + s_o \right) \\ &= (a-1) \frac{\beta}{1+\beta} \left(\frac{1}{1+\beta} \right)^{a-1} s_o \left(\frac{1}{\beta} + 1 \right) = (a-1) s_o \left(\frac{1}{1+\beta} \right)^{a-1} \end{aligned}$$

By focusing on the terms in ε^1 , we derive the following results:

$$\begin{aligned} \frac{d}{dx} \left(\frac{dx}{dt} \right)^{(1)} &= (a-1) \left\{ 2(1-2x) \left(s_l x^{a-2} + s_o (1-x)^{a-2} \right) + (a-2) (x-x^2) \left(s_l x^{a-3} - s_o (1-x)^{a-3} \right) \right\} \\ &= (a-1) \left[2 \frac{1-\beta}{1+\beta} (1-x)^{a-2} \left(s_l \left(\frac{x}{1-x} \right)^{a-2} + s_o \right) + \right. \\ &\quad \left. + (a-2) x (1-x)^{a-2} \left(s_l \left(\frac{x}{1-x} \right)^{a-3} - s_o \right) \right] \\ &= (a-1) \left[2 \frac{1-\beta}{1+\beta} \left(\frac{1}{1+\beta} \right)^{a-2} \left(s_l \frac{r}{\beta} + s_o \right) + (a-2) \frac{\beta}{1+\beta} \left(\frac{1}{1+\beta} \right)^{a-2} \left(s_l \frac{r}{\beta^2} - s_o \right) \right] \\ &= (a-1) s_o \left[2 \frac{1-\beta}{1+\beta} \left(\frac{1}{1+\beta} \right)^{a-2} \left(\frac{1+\beta}{\beta} \right) + (a-2) \frac{\beta}{1+\beta} \left(\frac{1}{1+\beta} \right)^{a-2} \left(\frac{1}{\beta^2} - 1 \right) \right] \\ &= (a-1) s_o \left[2 \frac{1-\beta}{\beta} \left(\frac{1}{1+\beta} \right)^{a-2} + (a-2) \left(\frac{1-\beta}{\beta} \right) \left(\frac{1}{1+\beta} \right)^{a-2} \right] \\ &= a s_o \left(\frac{1}{1+\beta} \right)^{a-2} \left[2 \frac{1-\beta}{\beta} + (a-2) \frac{1-\beta}{\beta} \right] = s_o \left(\frac{1}{1+\beta} \right)^{a-2} \left(a \frac{1-\beta}{\beta} \right) \\ &= a s_o \left(\frac{1}{1+\beta} \right)^{a-1} \left(\frac{1}{\beta} - \beta \right) \end{aligned}$$

In conclusion, our calculations result in:

$$\frac{d}{dx} \left(\frac{dx}{dt} \right) \Big|_{p^*} = (a-1) \left(\frac{1}{1 + \left(\frac{s_o}{s_l} \right)^{\frac{1}{a-1}}} \right)^{a-1} \left[1 + a\varepsilon \frac{1 - \left(\frac{s_o}{s_l} \right)^{\frac{2}{a-1}}}{\left(\frac{s_o}{s_l} \right)^{\frac{1}{a-1}}} \right] + \mathcal{O}(\varepsilon^2)$$

A.2.3 An alternative approach to stability Analysis of fixed points

In this section, we propose an alternative approach to the stability analysis of fixed points following the strategy outlined in [7], [10].

We introduce a “magnetization” $m = x_o - x_l = 2x_o - 1$. Equations (A.1) simplify to:

$$\frac{dm}{dt} = 2(s_o x_o^a x_l - s_l x_l^a x_o) \tag{A.15}$$

From the definition of m , it follows:

$$x_o = \frac{1+m}{2}, x_l = \frac{1-m}{2}$$

Equation (A.15) reduces to:

$$\frac{dm}{dt} = 2^{-a}(1-m^2)(s_o(1+m)^{a-1} - s_l(1-m)^{a-1}) \quad (\text{A.16})$$

Equation (A.15) has three fixed points:

$$m = \begin{cases} m_+ = +1 & \text{corresponding to } x_l = 0 \text{ and } x_o = 1 \\ m_- = -1 & \text{corresponding to } x_l = 1 \text{ and } x_o = 0 \\ m^* = \frac{1-\beta}{1+\beta} \end{cases} \quad (\text{A.17})$$

Calculation of m^*

$$\begin{aligned} s_o(1+m)^{a-1} - s_l(1-m)^{a-1} &= 0 \\ s_o(1+m)^{a-1} &= s_l(1-m)^{a-1} \\ \frac{s_o}{s_l} &= \left(\frac{1-m}{1+m}\right)^{a-1} = r \\ \frac{1-m}{1+m} &= r^{\frac{1}{a-1}} = \beta \\ 1-m &= (1+m)\beta = \beta + m\beta \rightarrow m = \frac{1-\beta}{1+\beta} \end{aligned}$$

There is correspondence between the fixed points p in (A.4) and the values of m : m_+ corresponds to p_0 , m_- to p_1 , and m^* is related to p^* . We approach the stability at the fixed points as follows. The values of magnetization m are between -1 and $+1$:

$$-1 \leq m \leq +1, -1 \leq \nu \leq 1$$

where $\nu = -1$ corresponds to $s'_o = 1 (s'_l = 0)$ and the opposite when ν assumes the value of 1.

The relevant limits for m approaching ± 1 are:

$$m_+(\epsilon) = 1 - \epsilon, m_-(\epsilon) = -1 + \epsilon, \epsilon > 0$$

We calculate the derivative of (A.16) for m and then substitute $m_+(\epsilon)$ in the resulting derivative to evaluate the sign.

$$\begin{aligned} \frac{d}{dm} \left(\frac{dm}{dt} \right) &= 2^{-a}(-2m)[s_o(1+m)^{a-1} - s_l(1-m)^{a-1}] \\ &\quad + 2^{-a}(1-m^2)[s_o(a-1)(1+m)^{a-2} + s_l(a-1)(1-m)^{a-2}] \end{aligned} \quad (\text{A.18})$$

The first term of (A.18) becomes:

$$\begin{aligned} \frac{d}{dm} \left(\frac{dm}{dt} \right) \Big|_{m=1-\varepsilon}^{1st} &= 2^{-a}(-2(1-\varepsilon))[s_o(2-\varepsilon)^{a-1} - s_l(\varepsilon)^{a-1}] \\ &\approx -2^{-(a-1)}[s_o 2^{a-1} - s_l(\varepsilon)^{a-1}] \\ &= -s_o + s_l \left(\frac{\varepsilon}{2}\right)^{a-1} \end{aligned} \quad (\text{A.19})$$

The second term reduces to:

$$\begin{aligned} \frac{d}{dm} \left(\frac{dm}{dt} \right) \Big|_{m=1-\varepsilon}^{2nd} &= 2^{-a}(\varepsilon)(2-\varepsilon)[s_o(a-1)(2-\varepsilon)^{a-2} + s_l(a-1)(\varepsilon)^{a-2}] \\ &\approx 2^{-(a-1)}\varepsilon[s_o(a-1)2^{a-2} + s_l(a-1)\varepsilon^{a-2}] \\ &\approx s_l(a-1) \left(\frac{\varepsilon}{2}\right)^{a-1} \end{aligned} \quad (\text{A.20})$$

Finally, we combine the two terms into (A.21).

$$\frac{d}{dm} \left(\frac{dm}{dt} \right) \Big|_{m=1-\varepsilon}^{1^{st}} \approx -s_o + as_l \left(\frac{\varepsilon}{2} \right)^{a-1} \quad (\text{A.21})$$

Similarly, we substitute the value $m_- = -1 + \varepsilon$ in (A.18) and the first term of (A.18) now becomes:

$$\begin{aligned} \frac{d}{dm} \left(\frac{dm}{dt} \right) \Big|_{m=-1+\varepsilon}^{1^{st}} &= 2^{-a}(-2(-1 + \varepsilon)) [s_o(1 + (-1 + \varepsilon))^{a-1} - s_l(2 - \varepsilon)^{a-1}] \\ &\approx 2^{-(a-1)} [s_o\varepsilon^{a-1} - s_l2^{a-1}] \\ &= -s_l + s_o \left(\frac{\varepsilon}{2} \right)^{a-1} \end{aligned} \quad (\text{A.22})$$

The second term reduces to:

$$\begin{aligned} \frac{d}{dm} \left(\frac{dm}{dt} \right) \Big|_{m=-1+\varepsilon}^{2^{nd}} &= 2^{-a}(\varepsilon)(2 - \varepsilon) [s_o(a - 1)\varepsilon^{a-2} + s_l(a - 1)(2 - \varepsilon)^{a-2}] \\ &\approx 2^{-(a-1)}\varepsilon [s_o(a - 1)\varepsilon^{a-2} + s_l(a - 1)2^{a-2}] \\ &\approx s_o(a - 1) \left(\frac{\varepsilon}{2} \right)^{a-1} \end{aligned} \quad (\text{A.23})$$

Finally, we combine the two terms into (A.24):

$$\frac{d}{dm} \left(\frac{dm}{dt} \right) \Big|_{m=-1+\varepsilon} \approx -s_l + as_o \left(\frac{\varepsilon}{2} \right)^{a-1} \quad (\text{A.24})$$

The last fixed point to evaluate is m^* . We substitute $m(\varepsilon) \rightarrow m + \varepsilon$ in (A.18) to evaluate the sign of the derivative.

The first term is the following:

$$\begin{aligned} \frac{d}{dm} \left(\frac{dm}{dt} \right) \Big|_{m=m^*}^{1^{st}} &= 2^{-a}(-2(m + \varepsilon)) [s_o(1 + m + \varepsilon)^{a-1} - s_l(1 - m - \varepsilon)^{a-1}] \\ &= 2^{-a}(-2(m + \varepsilon)) \left[s_o(1 + m)^{a-1} \left(1 + \frac{\varepsilon}{1 + m} \right)^{a-1} + \right. \\ &\quad \left. - s_l(1 - m)^{a-1} \left(1 - \frac{\varepsilon}{1 - m} \right)^{a-1} \right] \\ &\approx 2^{-a}(-2(m + \varepsilon)) \left[s_o(1 + m)^{a-1} \left(1 + \frac{(a - 1)\varepsilon}{1 + m} \right) + \right. \\ &\quad \left. - s_l(1 - m)^{a-1} \left(1 - \frac{(a - 1)\varepsilon}{1 - m} \right) \right] \\ &= 2^{-a}(-2(m + \varepsilon)) \left\{ s_o(1 + m)^{a-1} - s_l(1 - m)^{a-1} + \right. \\ &\quad \left. + (a - 1)\varepsilon [s_o(1 + m)^{a-2} + s_l(1 - m)^{a-2}] \right\} \\ &= 2^{-a}(-2(m + \varepsilon)) \left\{ (1 + m)^{a-1} \left[s_o - s_l \left(\frac{1 - m}{1 + m} \right)^{a-1} \right] + \right. \\ &\quad \left. + (a - 1)\varepsilon (1 + m)^{a-2} \left[s_o + s_l \left(\frac{1 - m}{1 + m} \right)^{a-2} \right] \right\} \end{aligned} \quad (\text{A.25})$$

Using the definitions of β and r to explicit the relation between s_l and s_o :

$$\beta = \frac{1-m}{1+m}, \beta^{a-1} = \left(\frac{1-m}{1+m}\right)^{a-1} = r = \frac{s_o}{s_l}$$

Equation (A.25), to the first order in ε , becomes:

$$\text{terms in } \varepsilon^0 = -2^{-(a-1)}(m) \left\{ (1+m)^{a-1} \underbrace{\left[s_o - s_l \left(\frac{1-m}{1+m} \right)^{a-1} \right]}_{s_l - s_l \frac{s_o}{s_l} = 0} + \right\} = 0 \quad (\text{A.26a})$$

$$\begin{aligned} \text{terms in } \varepsilon^1 &= -2^{-(a-1)}(\varepsilon) \left\{ (1+m)^{a-1} \underbrace{\left[s_o - s_l \left(\frac{1-m}{1+m} \right)^{a-1} \right]}_{s_l - s_l \frac{s_o}{s_l} = 0} + \right\} + \\ &\quad - 2^{-(a-1)} m(a-1) s_o \varepsilon \left[(1+m)^{a-2} \frac{\beta+1}{\beta} \right] \\ &= -2^{-(a-1)} m(a-1) s_o \varepsilon (1+m)^{a-2} \frac{\beta+1}{\beta} \end{aligned} \quad (\text{A.26b})$$

Similarly, we compute the second term of (A.18):

$$\begin{aligned} \frac{d}{dm} \left(\frac{dm}{dt} \right) \Big|_{m=m^*}^{2^{nd}} &= 2^{-a} (1-m^2) [s_o(a-1)(1+m)^{a-2} + s_l(a-1)(1-m)^{a-2}] \\ &= 2^{-a} (1+m+\varepsilon)(1-m-\varepsilon) [s_o(a-1)(1+m+\varepsilon)^{a-2} + s_l(a-1)(1-m-\varepsilon)^{a-2}] \\ &= 2^{-a} \underbrace{(1-m-\varepsilon+m-m^2-m\varepsilon+\varepsilon-m\varepsilon-\varepsilon^2)}_{1-m^2-2m\varepsilon+O(\varepsilon^2)} \cdot \\ &\quad [s_o(a-1)(1+m+\varepsilon)^{a-2} + s_l(a-1)(1-m-\varepsilon)^{a-2}] \\ &\approx 2^{-a} (1-m^2-2m\varepsilon) (a-1) \left[s_o (1+m)^{a-2} \left(1 + \frac{\varepsilon}{1+m} \right)^{a-2} + \right. \\ &\quad \left. + s_l (1-m)^{a-2} \left(1 - \frac{\varepsilon}{1-m} \right)^{a-2} \right] \\ &\approx 2^{-a} (1-m^2-2m\varepsilon) (a-1) \left[s_o (1+m)^{a-2} \left(1 + (a-2) \frac{\varepsilon}{1+m} \right) + \right. \\ &\quad \left. + s_l (1-m)^{a-2} \left(1 - (a-2) \frac{\varepsilon}{1-m} \right) \right] \\ &\approx 2^{-a} (1-m^2-2m\varepsilon) (a-1) \left\{ (1+m)^{a-2} \left[s_o + s_l \left(\frac{1-m}{1+m} \right)^{a-2} \right] + \right. \\ &\quad \left. + (a-2)\varepsilon \left[s_o (1+m)^{a-3} - s_l (1-m)^{a-3} \right] \right\} \\ &\approx 2^{-a} (1-m^2-2m\varepsilon) (a-1) \left\{ (1+m)^{a-2} \left[s_o + s_l \left(\frac{1-m}{1+m} \right)^{a-2} \right] + \right. \\ &\quad \left. + (a-2)\varepsilon (1+m)^{a-3} \left[s_o - s_l \left(\frac{1-m}{1+m} \right)^{a-3} \right] \right\} \end{aligned} \quad (\text{A.27})$$

Equation (A.27), to the first order in ε , becomes:

$$\text{terms in } \varepsilon^0 = 2^{-a}(1-m^2)(a-1) [s_o(1+m)^{a-2} + s_l(1-m)^{a-2}] \quad (\text{A.28a})$$

$$\begin{aligned} \text{terms in } \varepsilon^1 &+ -2^{-(a-1)}m\varepsilon(a-1) [s_o(1+m)^{a-2} + s_l(1-m)^{a-2}] + \\ &+ 2^{-a}(1-m^2)(a-1)(a-2)\varepsilon [s_o(1+m)^{a-3} - s_l(1-m)^{a-3}] \end{aligned} \quad (\text{A.28b})$$

Substituting β and r into (A.28):

$$\begin{aligned} \text{terms in } \varepsilon^0 &= 2^{-a}(1-m^2)(a-1)(1+m)^{a-2} \left[s_o + s_l \left(\frac{1-m}{1+m} \right)^{a-2} \right] \\ &= 2^{-a}(1-m^2)(a-1)(1+m)^{a-2}s_o \left[\frac{\beta+1}{\beta} \right] \\ &= 2^{-a}(1-m)(a-1)(1+m)^{a-1}s_o \left[\frac{\beta+1}{\beta} \right] \end{aligned} \quad (\text{A.29a})$$

$$\begin{aligned} \text{terms in } \varepsilon^1 &= -2^{-(a-1)}m\varepsilon(a-1) [s_o(1+m)^{a-2} + s_l(1-m)^{a-2}] + \\ &+ 2^{-a}(1-m^2)(a-1)(a-2)\varepsilon [s_o(1+m)^{a-3} - s_l(1-m)^{a-3}] \end{aligned} \quad (\text{A.29b})$$

At this stage, we can combine the terms in ε^0 and ε^1 from (A.26) and (A.29).

$$\begin{aligned} \text{terms in } \varepsilon^0 &= 2^{-a}(1-m)(a-1)(1+m)^{a-1}s_o \left[\frac{\beta+1}{\beta} \right] \\ &= 2^{-a}(a-1)s_o \left(\frac{\beta+1}{\beta} \right) \left(1 - \frac{1-\beta}{1+\beta} \right) \left(1 + \frac{1-\beta}{1+\beta} \right)^{a-1} \\ &= 2^{-a}(a-1)s_o \left(\frac{\beta+1}{\beta} \right) \left(\frac{2\beta}{1+\beta} \right) \left(\frac{2}{1+\beta} \right)^{a-1} \\ &= (a-1)s_o \left(\frac{1}{1+\beta} \right)^{a-1} \end{aligned} \quad (\text{A.30a})$$

$$\begin{aligned} \text{terms in } \varepsilon^1 &= -2^{-(a-1)}m\varepsilon(a-1) [s_o(1+m)^{a-2} + s_l(1-m)^{a-2}] \\ &+ 2^{-a}(1-m^2)(a-1)(a-2)\varepsilon [s_o(1+m)^{a-3} - s_l(1-m)^{a-3}] \end{aligned} \quad (\text{A.30b})$$

The terms in ε^1 can be further factorized:

$$\begin{aligned} \text{terms in } \varepsilon^1 &= -2^{-(a-1)}m\varepsilon(a-1) [s_o(1+m)^{a-2} + s_l(1-m)^{a-2}] + (a) \\ &+ 2^{-a}(1-m^2)(a-1)(a-2)\varepsilon [s_o(1+m)^{a-3} - s_l(1-m)^{a-3}] + (b) \\ &- 2^{-(a-1)}m(a-1)s_o\varepsilon(1+m)^{a-2} \frac{\beta+1}{\beta} (c) \end{aligned}$$

$$\begin{aligned} \text{term a} &= -2^{-(a-1)}m\varepsilon(a-1) [s_o(1+m)^{a-2} + s_l(1-m)^{a-2}] + (a) \\ &= -2^{-(a-1)} \frac{1-\beta}{1+\beta} \varepsilon(a-1)(1+m)^{a-2} \left[s_o + s_l \left(\frac{1-m}{1+m} \right)^{a-2} \right] \\ &= -2^{-(a-1)} \frac{1-\beta}{1+\beta} \varepsilon(a-1) \left(\frac{2}{1+\beta} \right)^{a-2} s_o \left(1 + \frac{1}{\beta} \right) \end{aligned} \quad (\text{A.31})$$

$$\begin{aligned}
\text{term b} &= 2^{-a}(1-m^2)(a-1)(a-2)\varepsilon [s_o(1+m)^{a-3} - s_l(1-m)^{a-3}] \\
&= 2^{-a}(1-m)(a-1)(a-2)(1+m)(1+m)^{a-3}\varepsilon \left[1 - \frac{1}{\beta^2}\right] \\
&= 2^{-(a-1)}(a-1)(a-2)\varepsilon \left(\frac{\beta}{1+\beta}\right) \left(\frac{2}{1+\beta}\right)^{a-2} \left(1 - \frac{1}{\beta^2}\right)
\end{aligned} \tag{A.32}$$

$$\begin{aligned}
\text{term c} &- 2^{-(a-1)}m(a-1)s_o\varepsilon(1+m)^{a-2} \frac{\beta+1}{\beta} \\
&= -2^{-(a-1)}\frac{1-\beta}{1+\beta}(a-1)s_o\varepsilon \left(\frac{2}{1+\beta}\right)^{a-2} \frac{\beta+1}{\beta}
\end{aligned} \tag{A.33}$$

After combining the terms (a), (b), and (c) in ε^1 , we get the final result:

$$\begin{aligned}
&= -2^{-(a-1)}\frac{1-\beta}{1+\beta}\varepsilon(a-1) \left(\frac{2}{1+\beta}\right)^{a-2} s_o \left(1 + \frac{1}{\beta}\right) + \\
&+ 2^{-(a-1)}(a-1)(a-2)\varepsilon \left(\frac{\beta}{1+\beta}\right) \left(\frac{2}{1+\beta}\right)^{a-2} \left(1 - \frac{1}{\beta^2}\right) + \\
&- 2^{-(a-1)}\frac{1-\beta}{1+\beta}(a-1)s_o\varepsilon \left(\frac{2}{1+\beta}\right)^{a-2} \frac{\beta+1}{\beta} + \\
&= -2^{-(a-1)}(a-1) \left(\frac{2}{1+\beta}\right)^{a-2} \left[2\frac{1-\beta}{1+\beta}\frac{\beta+1}{\beta} - (a-2)\frac{\beta}{1+\beta} \left(\frac{(\beta-1)(\beta+1)}{\beta^2}\right)\right] + \\
&= -2^{a-1}(a-1) \left(\frac{2}{1+\beta}\right)^{a-2} \left[2\frac{1-\beta}{\beta} - (a-2)\left(\frac{\beta-1}{\beta}\right)\right] \\
&= -2^{-(a-1)}(a-1) \left(\frac{2}{1+\beta}\right)^{a-2} \frac{\beta-1}{\beta} a
\end{aligned} \tag{A.34}$$

Recalling (A.30), we write the value of $\frac{dm}{dt} \Big|_{m=m^*}$:

$$\begin{aligned}
\frac{dm}{dt} \Big|_{m=m^*} &= (a-1)s_o \left(\frac{1}{1+\beta}\right)^{a-1} + 2^{-(a-1)}(a-1) \left(\frac{2}{1+\beta}\right)^{a-2} \frac{\beta-1}{\beta} a s_o \varepsilon \\
&= (a-1)s_o \left(\frac{1}{1+\beta}\right)^{a-1} \left[1 + a\frac{\varepsilon}{2} \left(\frac{1}{\beta} - \beta\right)\right]
\end{aligned} \tag{A.35}$$

Below, we report the stability of the three fixed points: The stability of m_+ primary depends on a , then on ε .

$$\begin{cases} m_+ = -s_o + s_l \left(\frac{\varepsilon}{2}\right)^{a-1} \text{ stable if } a > 1 \\ m_- = -s_l + s_o \left(\frac{\varepsilon}{2}\right)^{a-1} \text{ stable if } a > 1 \\ m^* = (a-1)s_o \left(\frac{1}{1+\beta}\right)^{a-1} \left[1 + a\frac{\varepsilon}{2} \left(\frac{1-\beta^2}{\beta}\right)\right] \text{ stable if } a < 1 \end{cases} \tag{A.36}$$

Please note that when we use $m, \varepsilon \rightarrow \varepsilon/2$, which makes (A.36) and (A.6) equivalent.

A.2.4 Evaluation of stability

The equation in reference A.3 has three fixed points, which are as follows:

$$x_l = 0, \quad x_l = 1, \quad \text{and} \quad x_l^* = \frac{\beta}{\beta+1}$$

The stability of these fixed points is influenced not only by the three model parameters— a , s_l , and s_o —but also by the parameter that governs the expansion of the equation around the fixed point, ε . $\lambda^a st$ is stable for $a < 1$ and the component that depends on ε goes to 0 as ε approaches 0. On the contrary, for $a < 1$, there are combinations of s_o and s_l which leads to a stability change in λ_o and λ_l . From

$$\lambda_o \approx -s_o + s_l a \varepsilon^{a-1}$$

follows that for $\varepsilon < \varepsilon_o^*$, where $\varepsilon_o^* = \left(\frac{s_o}{a s_l}\right)^{\frac{1}{a-1}}$, λ_o is positive. Similarly, from

$$\lambda_l \approx -s_l + s_o a \varepsilon^{a-1}$$

follows that for $\varepsilon < \varepsilon_l^*$, where $\varepsilon_l^* = \left(\frac{s_l}{a s_o}\right)^{\frac{1}{a-1}}$, λ_l is positive.

$$\lambda_o(s_l, s_o, a) = \begin{cases} \lambda_o > 0 & \text{if } \varepsilon < \varepsilon_o^* \\ \lambda_o < 0 & \text{if } \varepsilon > \varepsilon_o^* \end{cases} \quad (\text{A.37})$$

$$\lambda_l(s_l, s_o, a) = \begin{cases} \lambda_l > 0 & \text{if } \varepsilon < \varepsilon_l^* \\ \lambda_l < 0 & \text{if } \varepsilon > \varepsilon_l^* \end{cases} \quad (\text{A.38})$$

A.3 Special cases

A.3.1 The Abrams-Strogatz model in the linear case: $a=1$

In the linear case, the couple of differential equations become the logistic curve, [10]:

$$\frac{dx}{dt} = s_l x(1-x) - s_o x(1-x) = (s_l - s_o)x(1-x) \quad (\text{A.39})$$

Equation (A.39) has two fixed points: $x = 0$ and $x = 1$. Their stability depends on the value of $s_l - s_o$.

- $s_l - s_o < 0$ $x = 0$ is stable and $x = 1$ is unstable;
- $s_l - s_o > 0$ $x = 0$ is unstable and $x = 1$ is stable.

Study of $x = 0, x = 1$ stability

$$\frac{d}{dx} \left(\frac{dx}{dt} \right) = (s_l - s_o)(1 - 2x)$$

When $x = 0$, the derivative of $\frac{dx}{dt}$ is negative if $s_l - s_o < 0$ and positive if $s_l - s_o > 0$. The stability changes in $x = 1$.

If $s_l = s_o$, the languages have the same prestige values, which leads to a zero derivative.

$$\frac{dx}{dt} = 0, \forall x$$

This dynamic indicates a degeneracy of fixed points⁵ and it aligns with the voter dynamic model, which has been examined in various complex networks research papers, [9], [11], [14], [16].

⁵In this case, any initial condition is a fixed point.

A.3.2 The Abrams-Strogatz model in the quasi-linear case: $a = 1 + \mu$

In the quasi-linear case, the equation of the Abrams-Strogatz model (A.14), becomes the following, where we have set $\mu = a - 1$:

$$\frac{dx_l}{dt} = x_l (1 - x_l) (s_l x_l^\mu - s_o (1 - x_l)^\mu)$$

Since μ is small, we can use the approximation:

$$x^\mu = e^{\mu \log(x)} \approx 1 + \mu \log(x)$$

As the next step, we proceed to calculate the derivative⁶:

$$\begin{aligned} \frac{d}{dx} \left(\frac{dx}{dt} \right) &= (1 - 2x) (s_l (1 + \mu \log(x)) - s_o (1 + \mu \log(1 - x))) + \\ &+ (x - x^2) \mu \left(\frac{s_l}{x} + \frac{s_o}{1 - x} \right) \end{aligned}$$

Then we set $x = \varepsilon$ and consider $\varepsilon \rightarrow 0$. The first term is the following:

$$\begin{aligned} T_1 &= (1 - 2x) (s_l (1 + \mu \log(x)) - s_o (1 + \mu \log(1 - x))) \\ &= (1 - 2\varepsilon) (s_l (1 + \mu \log(\varepsilon)) - s_o (1 + \mu \log(1 - \varepsilon))) \\ &= (1 - 2\varepsilon) [s_l - s_o + \mu (\log(\varepsilon) - s_o \log(1 - \varepsilon))] \\ &= \underbrace{s_l - s_o}_{\text{order } \varepsilon^0} + \mu \underbrace{(s_l \log(\varepsilon) - s_o \log(1 - \varepsilon))}_{\text{order } \log(\varepsilon)} - 2\mu \varepsilon \underbrace{[\log(\varepsilon) - s_o \log(1 - \varepsilon)]}_{\text{order } \varepsilon \log \varepsilon} \\ &\approx \underbrace{s_l - s_o}_{\text{order } \varepsilon^0} + \mu \underbrace{(s_l \log(\varepsilon))}_{\text{order } \log(\varepsilon)} - \underbrace{s_o \mu \log(1)}_{=0} - 2\mu \varepsilon \underbrace{\log(\varepsilon)}_{=0} + 2\mu \varepsilon s_o \underbrace{\log(1)}_{=0} \\ &= s_l - s_o + \mu s_l \log \varepsilon \end{aligned}$$

Similarly, the second term becomes:

$$\begin{aligned} T_2 &= (x - x^2) \mu \left(\frac{s_l}{x} + \frac{s_o}{1 - x} \right) = \mu x (1 - x) \left(\frac{s_l}{x} + \frac{s_o}{1 - x} \right) \\ &= \mu (s_l (1 - x) + s_o (x)) \\ &\approx \mu (s_l (1 - \varepsilon) + s_o \varepsilon) = \mu \left(s_l + \underbrace{s_o \varepsilon}_{=0} \right) = \mu s_l \end{aligned}$$

We gather terms of the same order:

$$\approx s_l + \mu s_l - s_o + \mu \log \varepsilon = -s_o + s_l + \mu s_l (1 + \log \varepsilon)$$

We proceed as before to calculate λ_l :

$$\approx s_o + \mu s_o - s_l + \mu \log \varepsilon = -s_l + s_o + \mu s_o (1 + \log \varepsilon)$$

On the contrary, to manage the non-trivial fixed point of (A.14) is less straightforward. Indeed, we look for a x^* that solves:

$$s_l x^\mu - s_o (1 - x)^\mu = 0, \text{ for small } \mu \quad (\text{A.40})$$

⁶ x_l written as x for clarity.

We proceed by expanding $x^\mu \approx 1 + \mu \log(x)$ and by substituting the expansion in (A.40):

$$\begin{aligned}
s_l x^\mu - s_o (1-x)^\mu &= 0 \approx s_l (1 + \mu \log(x)) - s_o (1 + \mu \log(1-x)) = 0 \\
s_l \left[1 + \mu \log(x) - \underbrace{\frac{s_o}{s_l}}_r (1 + \mu \log(1-x)) \right] &= 0 \text{ and forgetting } s_l \\
1 + \mu \log(x) - r (1 + \mu \log(1-x)) &= 0 \rightarrow 1 - r + \mu (\log(x) - r \log(1-x)) = 0 \\
\mu \log \left[\frac{x}{(1-x)^r} \right] &= r - 1
\end{aligned} \tag{A.41}$$

The solution of (A.41) is reported in (A.42):

$$x^* = e^{\frac{r-1}{\mu}} (1 - x^*)^r \tag{A.42}$$

This is an implicit equation for x^* , which generally requires numerical methods to solve. We will rewrite the derivative here for clarity.

$$\begin{aligned}
\frac{d}{dx} \left(\frac{dx}{dt} \right) &= (1 - 2x) (s_l (1 + \mu \log(x)) - s_o (1 + \mu \log(1-x))) + \\
&\quad + (x - x^2) \mu \left(\frac{s_l}{x} + \frac{s_o}{1-x} \right)
\end{aligned}$$

We set $x = x + \varepsilon$, where x solves (A.42), and consider $\varepsilon \rightarrow 0$. The first term is as follows:

$$\begin{aligned}
T_1 &= (1 - 2x) (s_l (1 + \mu \log(x)) - s_o (1 + \mu \log(1-x))) \\
&= (1 - 2x) s_l [1 + \mu \log(x) - r (1 + \mu \log(1-x))] \\
&\approx (1 - 2x - 2\varepsilon) s_l [1 + \mu \log(x + \varepsilon) - r - r\mu \log(1 - x - \varepsilon)] \\
&= (1 - 2x - 2\varepsilon) s_l \left[1 - r + \mu \log \left(x \left(1 + \frac{\varepsilon}{x} \right) \right) - r\mu \left(\log(1-x) \left(1 - \frac{\varepsilon}{1-x} \right) \right) \right] \\
&= (1 - 2x - 2\varepsilon) s_l \left[1 - r + \mu \left(\log(x) + \log \left(1 + \frac{\varepsilon}{x} \right) \right) - r\mu \left(\log(1-x) + \log \left(1 - \frac{\varepsilon}{1-x} \right) \right) \right] \\
&\approx (1 - 2x - 2\varepsilon) s_l \left(1 - r + \mu \log(x) - r\mu \log(1-x) + \frac{\mu\varepsilon}{x} + \frac{r\mu\varepsilon}{1-x} \right) \\
&= (1 - 2x - 2\varepsilon) s_l \left(\underbrace{1 - r + \mu \log(x) - r\mu \log(1-x)}_{=0} + \mu\varepsilon \left(\frac{1}{x} + \frac{r}{1-x} \right) \right) \\
&\approx (1 - 2x) s_l \mu \left(\frac{1}{x} + \frac{r}{1-x} \right) \varepsilon + \mathcal{O}(\varepsilon^\epsilon)
\end{aligned}$$

Similarly, the second term becomes:

$$\begin{aligned}
T_2 &= (x - x^2) \mu \left(\frac{s_l}{x} + \frac{s_o}{1-x} \right) \\
&\approx ((x + \varepsilon) - (x + \varepsilon)^2) \mu \left(\frac{s_l}{x + \varepsilon} + \frac{s_o}{1-x-\varepsilon} \right) \\
&\approx (x - x^2 + \varepsilon(1-2x)) \mu s_l \left(\frac{1}{(x)(1+\frac{\varepsilon}{x})} + \frac{r}{(1-x)(1-\frac{\varepsilon}{1-x})} \right) \\
&\approx (x - x^2 + \varepsilon(1-2x)) \mu s_l \left(\frac{1}{x} \left(1 - \frac{\varepsilon}{x}\right) + \frac{r}{(1-x)} \left(1 + \frac{\varepsilon}{1-x}\right) \right) \\
&= (x - x^2 + \varepsilon(1-2x)) \mu s_l \left(\frac{1}{x} - \frac{\varepsilon}{x^2} + \frac{r}{1-x} + \frac{r\varepsilon}{(1-x)^2} \right) \\
&= (x - x^2 + \varepsilon(1-2x)) \mu s_l \left[\frac{1}{x} + \frac{r}{1-x} - \varepsilon \left(\frac{1}{x^2} - \frac{r}{(1-x)^2} \right) \right] \\
&= \underbrace{(1+x(r-1)) \mu s_l}_{\text{order } \varepsilon^0} + \\
&\quad + \mu s_l \underbrace{\left[(1-2x) \left(\frac{1}{x} + \frac{r}{1-x} \right) + \frac{1-x(r+1)}{x(1-x)} \right]}_{\text{order } \varepsilon^1} \varepsilon + \mathcal{O}(\varepsilon^\infty)
\end{aligned}$$

We now gather the various terms and obtain:

$$\lambda_* = (1+x(r-1)) \mu s_l + \mu s_l \left[2(1-2x) \left(\frac{1}{x} + \frac{r}{1-x} \right) + \frac{1-x(r+1)}{x(1-x)} \right] \varepsilon$$

Finally, we write the three eigenvalues for the quasi-linear Abrams-Strogatz model:

$$\lambda(s_l, r, \mu) = \begin{cases} \lambda_o \approx -s_o + s_l + \mu s_l (1 + \log \varepsilon) = s_l [-r + 1 + \mu (1 + \log \varepsilon)] & \text{corresponding to } p_0 \\ \lambda_l \approx -s_l + s_o + \mu s_o (1 + \log \varepsilon) = s_l [-1 + r + \mu r (1 + \log \varepsilon)] & \text{corresponding to } p_1 \\ \lambda_* \approx (1+x^*(r-1)) \mu s_l + \mu s_l \left[2(1-2x^*) \left(\frac{1}{x^*} + \frac{r}{1-x^*} \right) + \frac{1-x^*(r+1)}{x^*(1-x^*)} \right] \varepsilon & \text{corresponding to } p^* \end{cases} \quad (\text{A.43})$$

As we did previously, we now analyze the stability of fixed points reported in (A.43). Each eigenvalue is scaled by s_l .

$$\frac{\lambda_o}{s_l} = 1 - r + \mu (1 + \log \varepsilon)$$

From the model parameters, we know that $r > 1$. Meanwhile, $1 + \log \varepsilon < 0$ for $\varepsilon < e^{-1} \approx 0.37$. Since $\varepsilon \rightarrow 0$, $1 + \log \varepsilon < 0$. If μ is positive, λ_o is always stable. On the contrary, when μ is negative, the stability of λ_o changes when $\varepsilon > \varepsilon^*$, where $\varepsilon^* = e^{\frac{r-1}{\mu}-1}$.

Next, we will assess the stability of λ_l .

$$\frac{\lambda_l}{s_l} = r - 1 + r\mu (1 + \log \varepsilon)$$

In this case, the stability of λ_l depends on a combination of various factors. ε^* is the following:

$$\varepsilon^* = e^{\frac{1-r}{r\mu}-1}$$

If μ is positive the exponent $\frac{1-r}{r\mu} - 1$ is negative, so there exists a small ε^* so that:

$$\lambda_l(s_l, r, \mu) = \begin{cases} \lambda_l > 0 & \text{if } \varepsilon > \varepsilon^* \\ \lambda_l < 0 & \text{if } \varepsilon < \varepsilon^* \end{cases}$$

For negative μ , the exponent in ε^* may become large, and λ_l is always positive.

Finally, we examine the stability of λ^* . As ε approaches 0, the component that depends on ε approaches 0 as well. The portion of the eigenvalues that is independent of ε is positive for positive values of μ and negative for negative values of μ . The following equation summarizes the results:

$$\lambda_o(s_l, r, \mu) = \begin{cases} \lambda_o < 0 & \text{if } \varepsilon > \varepsilon^* \text{ and } \mu < 0 \\ \lambda_o > 0 & \text{if } \varepsilon < \varepsilon^* \text{ and } \mu < 0 \\ \lambda_o < 0 & \text{if } \mu > 0 \end{cases} \quad (\text{A.44})$$

$$\lambda_l(s_l, r, \mu) = \begin{cases} \lambda_l > 0 & \text{if } \varepsilon > \varepsilon^* \text{ and } \mu > 0 \\ \lambda_l < 0 & \text{if } \varepsilon < \varepsilon^* \text{ and } \mu > 0 \\ \lambda_l > 0 & \text{if } \mu < 0 \end{cases} \quad (\text{A.45})$$

$$\lambda^*(s_l, r, \mu) = \begin{cases} \lambda^* > 0 & \text{if } \mu > 0 \\ \lambda^* < 0 & \text{if } \mu < 0 \end{cases} \quad (\text{A.46})$$

A.4 Numerical simulations

This section describes the authors' strategies to provide numerical support and simulation of the theoretical results. We propose the following areas of numerical simulation.

- Evaluation of stability. We examine how the eigenvalues of (A.3) depend on the model's parameters, considering divergences as ε approaches 0;
- Numerical calculation of the derivative at the fixed points. The numerical values derived by numerical calculation compare with the theoretical result. The numerical values are also compared with the stability behavior;
- Perturbative Theory at stable fixed points. For stable fixed points, we apply a perturbative theory to demonstrate that the fixed point acts as a basin of attraction;
- Critical dynamics around $a \sim 1 + \mu$;
- Data extrapolation.

A.4.1 Evaluation of stability

Equation A.3 has three fixed points:

$$x_l = 0, x_l = 1(x_o = 0), x_l^* = \frac{\beta}{\beta + 1}(x_o = \frac{1}{\beta + 1})$$

Stability depends on three parameters: a , s_l , and s_o , as well as the expansion parameter around the fixed point, ε . We define a normalization parameter $k = s_o + s_l$ and scale the prestige values, $s_l \rightarrow s'_l = s_l/k$, and $s_o \rightarrow s'_o = s_o/k$. We also introduce a "tuning" parameter, ν , and rewrite s'_l and s'_o in terms of ν :

$$s'_l = \frac{1 - \nu}{2}, s'_o = \frac{1 + \nu}{2}$$

The eigenvalues of (A.6), rescaled by k become:

$$\lambda(\nu, a)' = \begin{cases} \lambda_o \approx -s_o + s_l a \varepsilon^{a-1} \rightarrow k \left(-\frac{s_o}{k} + a \frac{s_l}{k} \varepsilon^{a-1} \right) \rightarrow \lambda'_o = \left(-\frac{1+\nu}{2} + a \frac{1-\nu}{2} \varepsilon^{a-1} \right) \\ \lambda_l \approx -s_l + s_o a \varepsilon^{a-1} \rightarrow k \left(-\frac{s_l}{k} + a \frac{s_o}{k} \varepsilon^{a-1} \right) \rightarrow \lambda'_l = \left(-\frac{1-\nu}{2} + a \frac{1+\nu}{2} \varepsilon^{a-1} \right) \\ \lambda_* \approx (a-1)s_o \rightarrow k(a-1)\frac{1+\nu}{2} \rightarrow \lambda'_* = (a-1)\frac{1+\nu}{2} \end{cases} \quad (\text{A.47})$$

where λ_* and λ'_* do not contain the term $\left(\frac{1}{1+\beta}\right)^{a-1}$ which is always positive.

The volatility parameter a defines the stability of λ' . When $0 < a < 1$, $\lambda'_* < 0$, thanks to the factor $(a - 1)$, but the discussion on the stability of p_0 i.e. λ_o and p_1 i.e. λ_l is more intricate, since the tuning parameter $0 < \nu < 1$ is also involved. In fact,

$$\lambda'_o \geq 0 \implies \left(-\frac{1+\nu}{2} + a\frac{1-\nu}{2}\varepsilon^{a-1}\right) \geq 0; \implies \varepsilon \leq \varepsilon_o$$

where $\varepsilon_o = \left(\frac{1+\nu}{a(1-\nu)}\right)^{\frac{1}{a-1}}$

$$\lambda'_l \geq 0 \implies \left(-\frac{1-\nu}{2} + a\frac{1+\nu}{2}\varepsilon^{a-1}\right) \geq 0; \implies \varepsilon \leq \varepsilon_l$$

where $\varepsilon_l = \left(\frac{1-\nu}{a(1+\nu)}\right)^{\frac{1}{a-1}}$

Figure A.1 illustrates the trend of λ'_o for $0 < a < 1$, $0 < \nu < 1$, with ε values ranging from 10^{-1} to 10^{-7} . Each subplot corresponds to a specific ε value and uses a color gradient to represent positive values in red and negative values in blue. Further, each subplot features a black contour line that indicates where the equation equals zero, signifying the boundary between positive and negative regions. The Figure is generated from the parameters $a = 0.75$, $s_o = 0.75$, and $s_l = 0.25$; $\varepsilon_o^* \approx 0.004$.

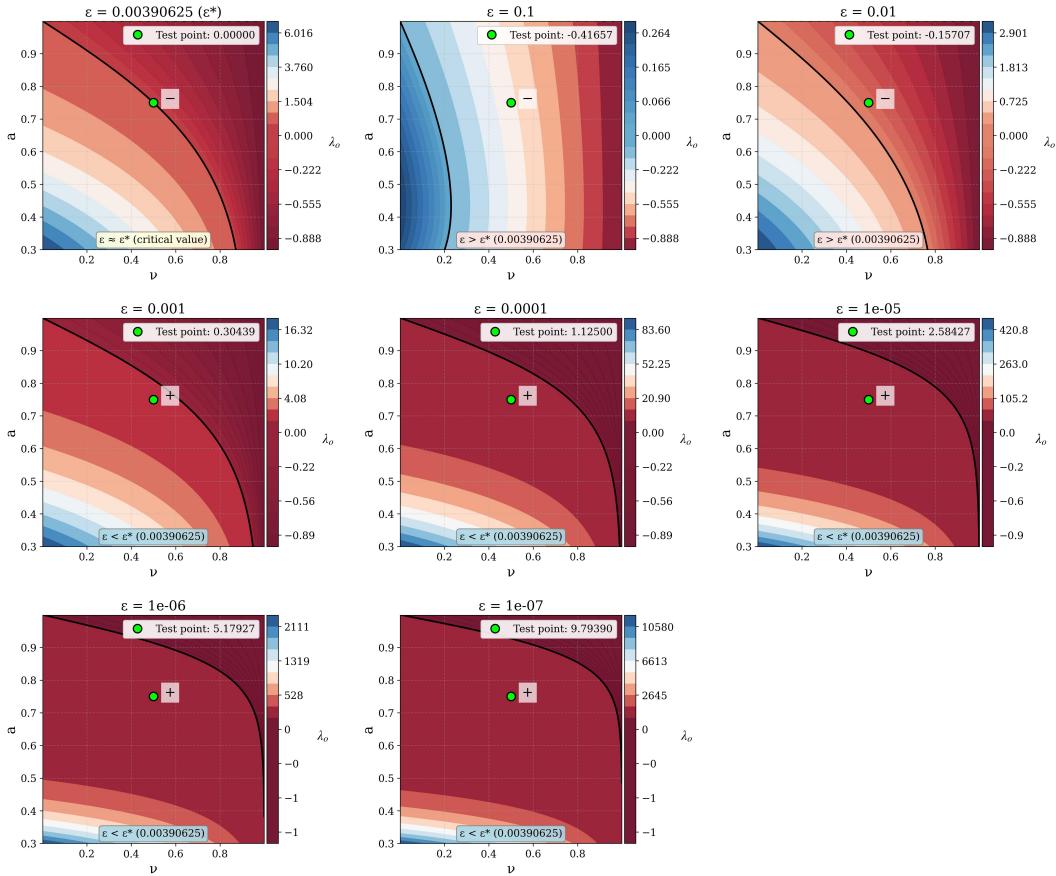


Figure A.1: As a function of a , ν , and ε , λ'_o shows that as ε approaches zero, the region where the function is positive tends to encompass the entire plot. When $\varepsilon < \varepsilon_o^*$, λ_o is positive, denoting a change of stability.

A.4.2 Perturbation theory

Perturbation theory is applied around the stable fixed point $f_p = x^*$:

$$\frac{d(x^* + \delta x)}{dt} = f(x^* + \delta x) = f(x^*) + f'(x^*)\delta x \rightarrow \frac{d\delta x}{dt} = f'(x^*)\delta x$$

where $f'(x^*) < 0$.

The procedure for numerical simulation in perturbation theory is outlined below. The fixed point x_{th} is calculated from the AS Model and from the fixed point, we extrapolated the time t_p at which $x_l = 2x_{th}$. We then selected five x_0 between $\frac{x_{th}}{2}$ and $2x_{th}$ and applied the perturbation theory from t_p :

$$x_l(t) = x_{th} + (x_o(t_p) - x_{th}) e^{-\lambda^*(t-t_p)}$$

where λ^* is the value of the derivative at the fixed point.

A.4.3 Evaluation of stability in critical dynamics

This section evaluates stability within the context of critical dynamics. In the Abrams-Strogatz scenario, critical dynamics occur when the volatility parameter a is approximately 1.

Figure A.2 illustrates the trend of λ_o for the range $1 < r < 2$ and $-0.1 < \mu < 0.1$, with ε values varying from 10^{-1} to 10^{-7} . Each subplot corresponds to a specific ε value and utilizes a color gradient to represent positive values in red and negative values in blue. Additionally, each subplot includes a black contour line indicating where the equation equals zero, marking the boundary between positive and negative regions. This figure is generated with the parameters $\mu = -0.03$, $s_o = 0.75$, and $s_l = 0.25$, resulting in $\varepsilon_o^* \approx 0.023$.

In contrast, according to equation (A.44), when $\mu > 0$, λ_o is always negative. Figure A.3 is generated using $\mu = 0.007$, $s_o = 0.75$, and $s_l = 0.25$, resulting in $\varepsilon_o^* \approx 5.4 \times 10^5$, which is a non-indicative value.

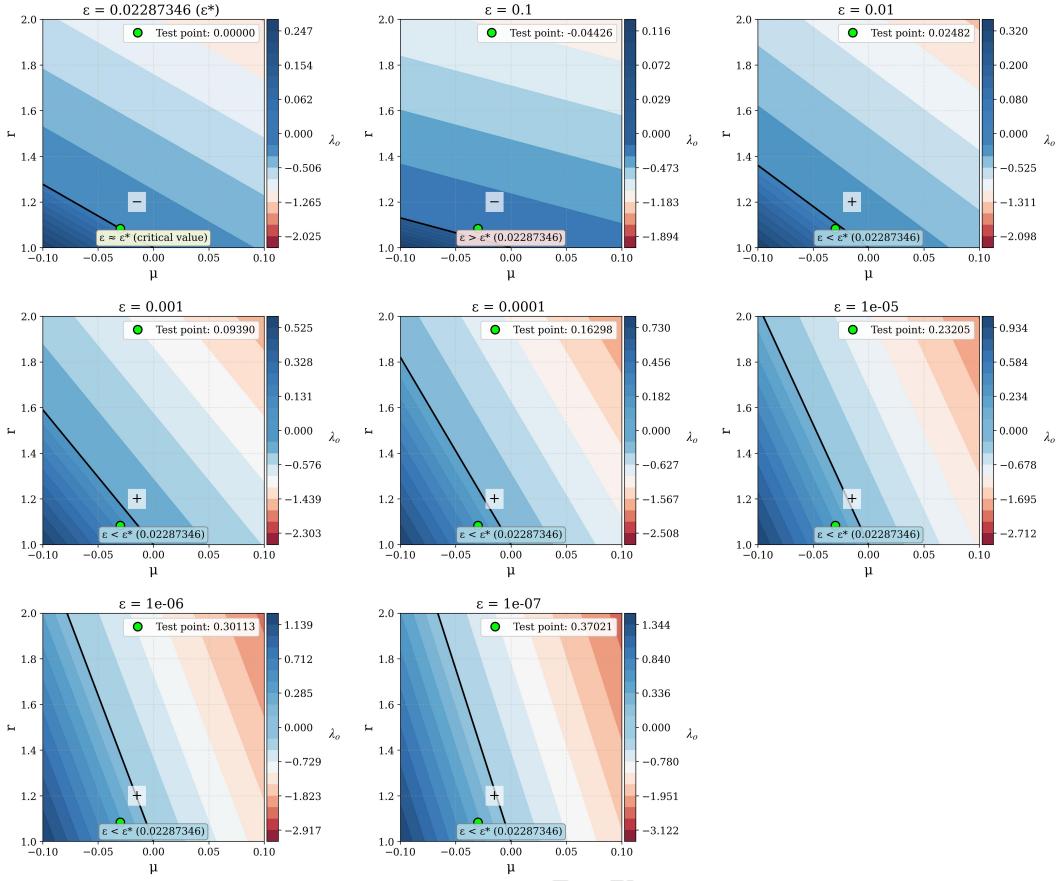


Figure A.2: As a function of r , μ , and ε , λ_o shows that as ε approaches zero, the region where the function is positive gradually increases. When $\varepsilon < \varepsilon_o^*$, λ_o is positive, denoting a change of stability.

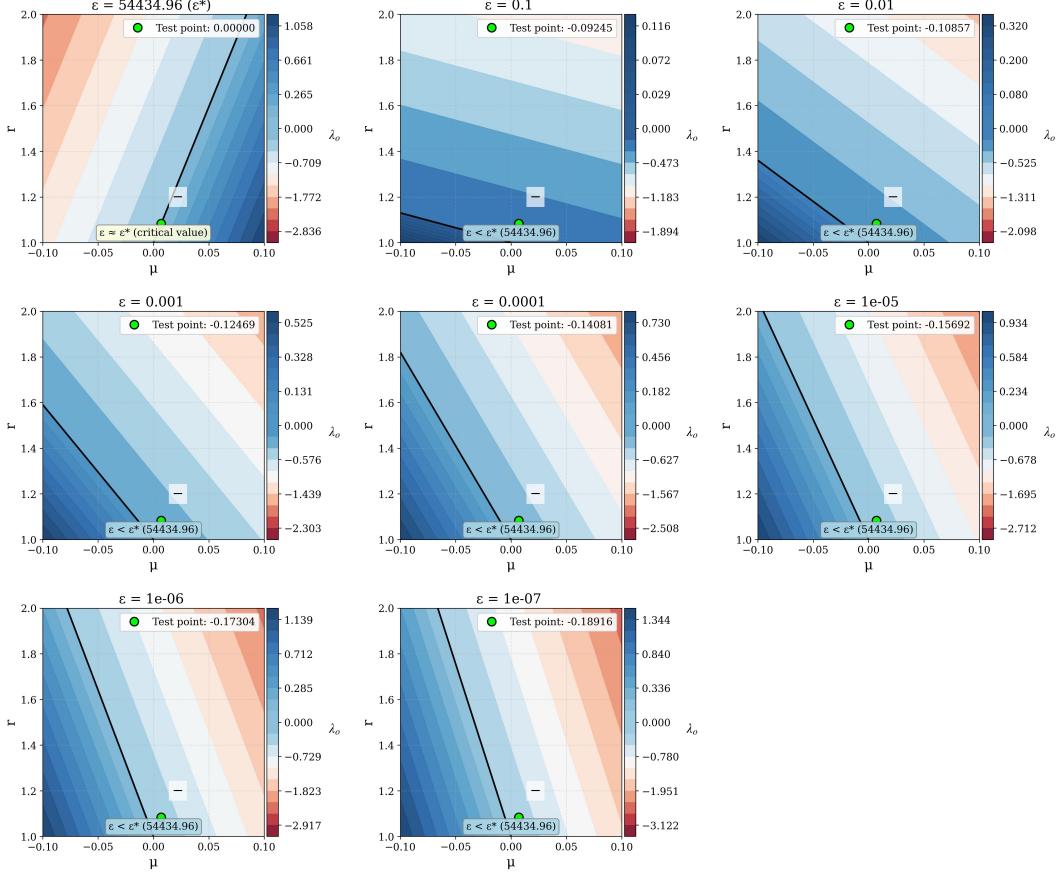


Figure A.3: As a function of a , ν , and ε , λ'_o shows that as ε approaches zero, the region where the function is positive tends to encompass the entire plot. When $\varepsilon < \varepsilon^*$, λ'_o is positive, denoting a change of stability.

We conducted further analyses related to the dependence of ε^* and the fixed point (x^*), obtained when the volatility a is lesser than 1, on volatility and the ratio of the prestige value of the majority language compared to that of the minority language, $r = s_o/s_l$.

Figure A.4 illustrates the trend of ε^* , calculated from the equation $\lambda'_o = -s_o + as_l(\varepsilon)^{1/(a-1)}$ in the standard model, and as $\varepsilon^* = e^{\frac{r-1}{a-1}-1}$ for the quasi-linear Abrams-Strogatz models with a single prestige value ($s_l = 1 - s_o$). The values of ε^* shown in the legend represent the first value at which λ'_o becomes positive.

For both models, this occurs when $\nu \approx 0$, which corresponds to a $r \approx 1$ ratio. This observation aligns with the single prestige model, where the minimum value, s_o^{min} , for s_o is slightly greater than 0.5. The difference in values is attributed to the mathematical laws: $\varepsilon^* \sim (1/a)^{1/(a-1)}$ for the standard model, in contrast to $\varepsilon^* \sim \frac{1}{e}$ for the quasi-linear model.

As ν comes to 1, the prestige value for s_o increases, leading to a drop in the critical value of ε^* . This suggests that a slight change in the fraction of minority speakers can significantly affect the stability of the fixed point at $x = 0$. Additionally, when the volatility a is close to 1, ε^* becomes smaller, which reinforces the model's behavior in a near-linear fashion.

A notable difference between the left (standard) and right (quasi-linear) models is the steepness of their curves. The quasi-linear model displays steeper curves than the standard model, meaning that the transition from stable to unstable occurs for lower values of ν . However, when the volatility approaches 1, the values between the two models differ by about 2%.

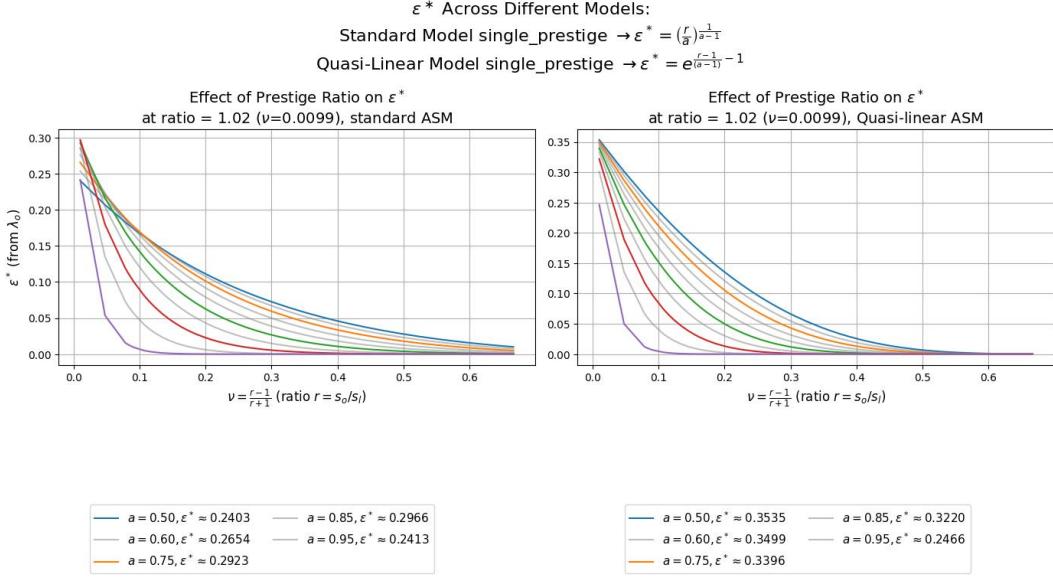


Figure A.4: As a function of a and ν , ε^* (from λ_o) decreases when ν approaches 1. The quasi-linear model has steeper curves, but the values tend to converge near $a = 1$.

The values of ε^* reported in Figure A.5 are generated using $s_o = 0.75$ and $s_l = 0.25$, i.e. $r = 3$ ($\nu = 0.5$). The values are different for a far from 1, but tend to converge as a approaches 1.

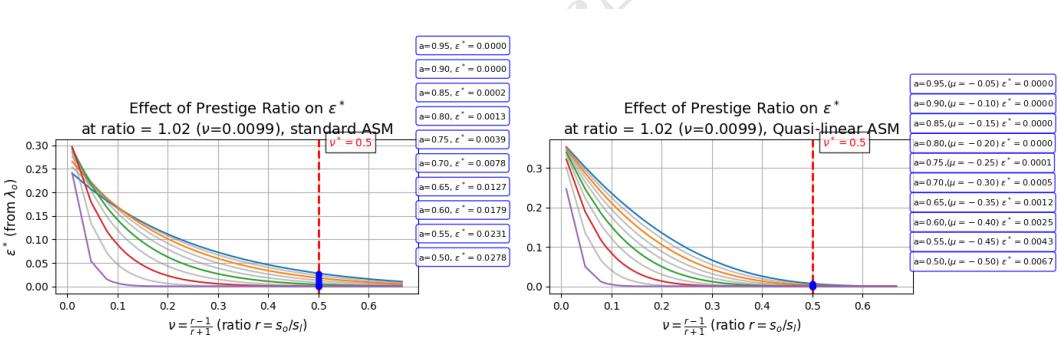


Figure A.5: At $\nu = 0.5$, $r = 3$, The quasi-linear and the standard models have similar values when a tends to 1.

Figure A.6 illustrates the trends of the fixed point x^* as the parameters a and ν vary. The values of x^* are derived from a double prestige model with $\nu \approx 0.53$, which corresponds to a ratio of $r \approx 3.25$. Although the curves exhibit a similar steepness, the quasi-linear model reaches zero more rapidly than the standard model. As a result, the fixed point x^* predicted by the quasi-linear model is slightly smaller than that predicted by the standard model. In the case of the quasi-linear model, when $a < 1$, the stable fixed point x^* is quite close to the unstable fixed point $x = 0$.

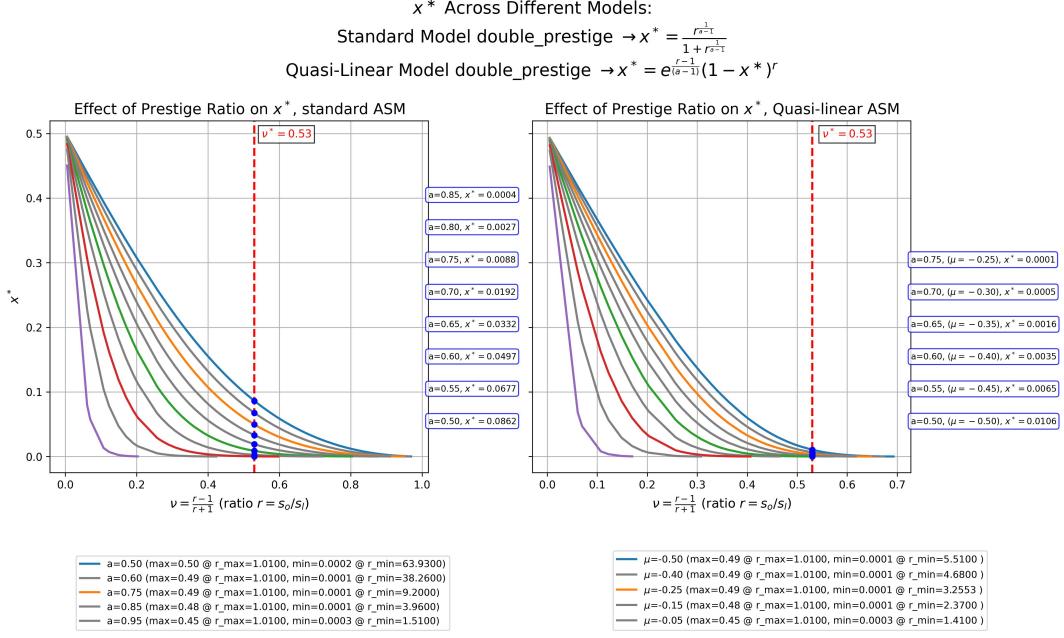


Figure A.6: As a function of a and ν , x^* decreases similarly in both models, but the quasi-linear model reaches zero for smaller μ , compared with the standard model.

A.4.4 Numerical calculation of the derivative at the fixed points

The parameters used for numerical simulation are the optimal parameters (including uncertainties) provided by the (Alg.A.3) algorithm. The algorithm uses the bisection method to calculate non-trivial fixed points (i.e., those different from 0 and 1) of (A.3). The bisection method highlights the areas where the function (A.3) changes sign in the open interval $(0, 1)$ and returns the list of non-trivial fixed points with a tolerance of 10^{-8} .

For each fixed point, f_p , including 0 and 1, we calculate the derivative at that point using the central difference method to truncate the error to the second order in h , $\mathcal{O}(\langle \epsilon \rangle)$:

$$\left. \frac{df}{dt} \right|_{f_p} \approx \frac{f(f_p + h) - f(f_p - h)}{2h}$$

This formula is beneficial, as we can choose the value of h to be one of the ϵ values used in Section A.4.1 to verify both the value and sign of the derivative itself. For instance, the eigenvalues $\lambda_o - \lambda_l$ for the standard model can be calculated accordingly to (A.49)

$$\left. \frac{dx_l}{dt} \right|_{p_0} = D_{p_0} = -s_o + as_l (h)^{a-1} \quad (\text{A.49a})$$

$$\left. \frac{dx_l}{dt} \right|_{p_1} = D_{p_1} = -s_l + as_o (h)^{a-1} \quad (\text{A.49b})$$

Similarly, (A.50) calculates the same eigenvalues for the quasi-linear model:

$$\left. \frac{dx_l}{dt} \right|_{p_0} = D_{p_0}^{ql} = -s_o + s_l + \mu s_l (1 + \log h) \quad (\text{A.50a})$$

$$\left. \frac{dx_l}{dt} \right|_{p_1} = D_{p_1}^{ql} = -s_l + s_o + \mu s_o (1 + \log h) \quad (\text{A.50b})$$

A.5 Abrams-Strogatz Model with a varying population

This section describes an extension of the AS model that accounts for a varying population. As detailed in Section A, the basic differential equations for x_o and x_l are given by:

$$\begin{aligned}\frac{dx_l}{dt} &= s_l x_o x_l^a - s_o x_o^a x_l \\ \frac{dx_o}{dt} &= s_o x_o^a x_l - s_l x_o x_l^a \text{ with } x_o + x_l = 1\end{aligned}\quad (\text{A.51})$$

Here, we relax the normalization constraint and focus on the following relation:

$$X_o(t) + X_l(t) = I(t), x_o(t) = X_o(t)/I(t)$$

We suppose that the x_o quantity has two types of increments. The first one derives from the differential equation (A.51) and concerns the percentage of x_l that changes to x_o . The second increment depends on the probability that out of n new individuals from I , k are already x_o . The increments of population X_o are described below:

$$X_o(t + dt) = X_o(t) + \Delta X_o \Big|_{ASM} + \Delta X_o \Big|_{growth} \quad (\text{A.52})$$

The first term is directly connected to the ASM dynamics:

$$\Delta X_o \Big|_{ASM} = \Delta(Ix_o) = I(t)\Delta x_o + x_o \underbrace{\Delta I}_{=0}$$

The second term measures the X_o natural increment:

$$\Delta X_o \Big|_{growth} = \Delta(pI) = p\Delta + I\Delta p$$

The probability p is modeled as a binomial likelihood. This approach assumes that each new individual in the population I independently faces a probability of becoming X_o , influenced by whether they are born into a X_o group. If we denote by p the probability of an individual joining the X_o group, and by $1 - p$ the probability of remaining in the X_l , the number of new X_o individuals, k , among n new individuals follows a binomial distribution:

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

However, our interest lies in the expected proportion of individuals joining X_o , which is proportional to p :

$$\mathbb{E}(k) = np.$$

For a short time interval Δt , the number of new individuals in the group I can be approximated by the formula $n = \frac{dI}{dt}\Delta t$. Therefore, the expected number of new Bilinguals per unit of time is defined by this relationship.

$$\frac{\mathbb{E}(k)}{\Delta t} = p \frac{dI}{dt}.$$

This formulation expresses that the overall growth of I and the probability p of transitioning to X_o influence the growth of the X_o group. We may also suppose that $p = p(I, \dots)$.

In the limit as $dt \rightarrow 0$, the contribution of ASM to the global dynamics becomes:

$$\Delta X_o \Big|_{ASM} \rightarrow \frac{d}{dt} (Ix_o) = I(t) \frac{dx_o}{dt} = I(t) [s_o x_o^a (1 - x_o) - s_l (1 - x_o)^a x_o] + x_o \underbrace{\frac{dI}{dt}}_{=0}$$

while the growth term is:

$$\Delta X_o \Big|_{growth} \rightarrow \frac{d}{dt} (pI) = p \frac{dI}{dt} + I \frac{dp}{dt}$$

The differential equation for the absolute number of X_o is reported in the following equation:

$$\frac{dX_o}{dt} \Big|_{tot} = I \frac{dx_o}{dt} \Big|_{ASM} + p_o \frac{dI}{dt} + I \frac{dp_o}{dt} \quad (\text{A.53})$$

Using $x_o = X_o/I$, it follows:

$$\frac{dx_o}{dt} = \frac{d}{dt} \left(\frac{X_o}{I} \right) = \frac{1}{I} \frac{dX_o}{dt} - x_o \frac{1}{I} \frac{dI}{dt}$$

Using (A.53), we obtain:

$$\begin{aligned} \frac{dx_o}{dt} \Big|_{tot} &= \frac{1}{I} \left[I \frac{dx_o}{dt} \Big|_{ASM} + p_o \frac{dI}{dt} + I \frac{dp_o}{dt} \right] \\ &\quad - \frac{1}{I} x_o \frac{dI}{dt} \end{aligned} \quad (\text{A.54})$$

With the explicit dependence of p on I , expressed as $dp/dt = (dp/dI) \cdot (dI/dt)$, we obtain the following formulation for the complete dynamics, where we have used (A.51):

$$\begin{aligned} \frac{dx_o}{dt} \Big|_{tot} &= s_o x_o^a (1 - x_o) - s_l (1 - x_o)^a x_o \\ &\quad + \frac{1}{I} \left[p_o - x_o + I \frac{dp_o}{dI} \right] \frac{dI}{dt} \end{aligned} \quad (\text{A.55})$$

B Data Pre-processing

Census data, particularly from the late 1800s and early 1900s, lacks some of the information necessary for this study. Since it is essential for all categories we examined to have complete data, some preprocessing activities were required. The following sections describe these activities in detail.

B.1 Data Pre-fitting

We use a logistic model, commonly applied in population growth studies (see, for example, [4], [5]), to estimate missing data for the categories: Population, Indigenous, Spanish, and Bilingual.

The adopted methodology is broken down as follows. We used the available data to calculate the growth rate (r), the initial population (N_0), and the carrying capacity (K) to obtain the best-fitting logistic curve. Then, we calculated the means and standard deviations of the residuals we inserted in a normal distribution to generate the missing values. We executed the code ten times and took the mean of the resulting values to obtain data close to the logistic curve but not on it. The logistic curve used is reported in (B.1), while Figure B.1 shows the picture obtained for the total missing (men and women) values.

$$L(t, K, r, N_0) = \frac{K}{(1 + ((K - N_0)/N_0)e^{(-rt)})} \quad (\text{B.1})$$

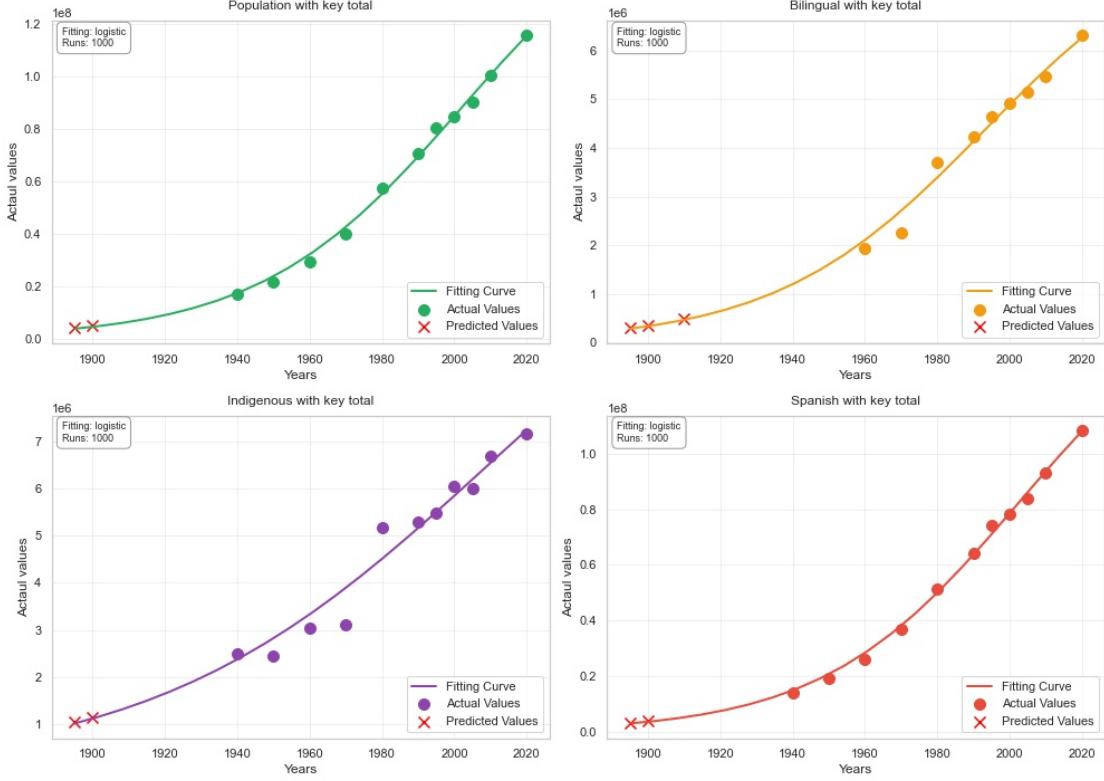


Figure B.1: Missing data (red x) and logistic fitting for Population, Bilingual, Indigenous, and Spanish.

Table B.1 reports the missing values obtained with the methodology described above.

Table B.1: Missing values for different categories of data.

Missing Year(s)	Predicted Values	Category
1895,1900	4146290,4926065	Population
1895,1900,1910	297977,337528,478232	Bilingual
1895,1900	1030670,1144766	Indigenous
1895,1900	3394259,4054539	Spanish

B.2 Managing “I do not know (DK)” categories

As outlined in Section 3, there are two similar types of “I do not know” (DK) responses. The first type includes individuals who speak neither Spanish nor Indigenous languages but rather foreign languages, such as German and Italian. The second type is more significant for this study and involves Indigenous individuals who did not respond to the question about their Bilingualism.

We removed such values from the total Population category for the first typology. The following factors support this decision: (i) the study refers to Spanish/Indigenous Bilingualism, and (ii) the average is approximately 0.5 percent of the total population.

In the second typology, the DK values within the Indigenous category average 2 percent. This figure falls below the 5 percent threshold, regarded as the maximum value for the proportional redistribution of DKs among other categories, [4], [8]. As a result, we have updated the Bilingual and Indigenous_Monolingual values based on the proportions derived from the DKs. Figures B.2 and B.3 exhibit the results of proportional redistribution.

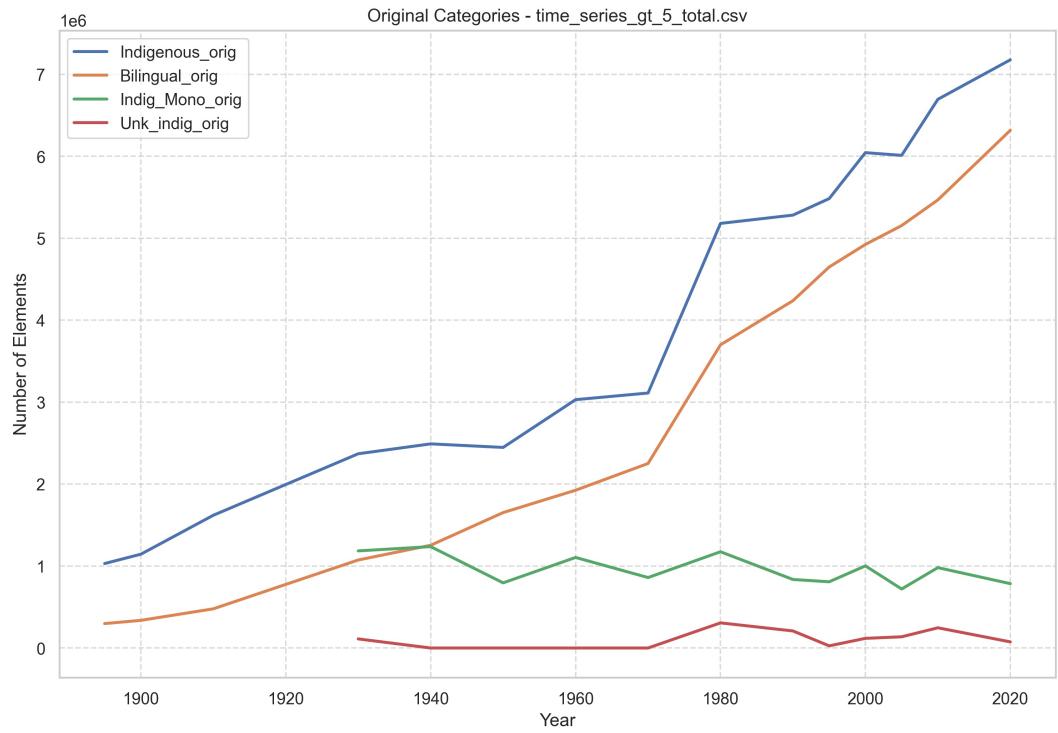


Figure B.2: Trends in bilingual, monolingual, and “don’t know” (DK) responses within the Indigenous category.

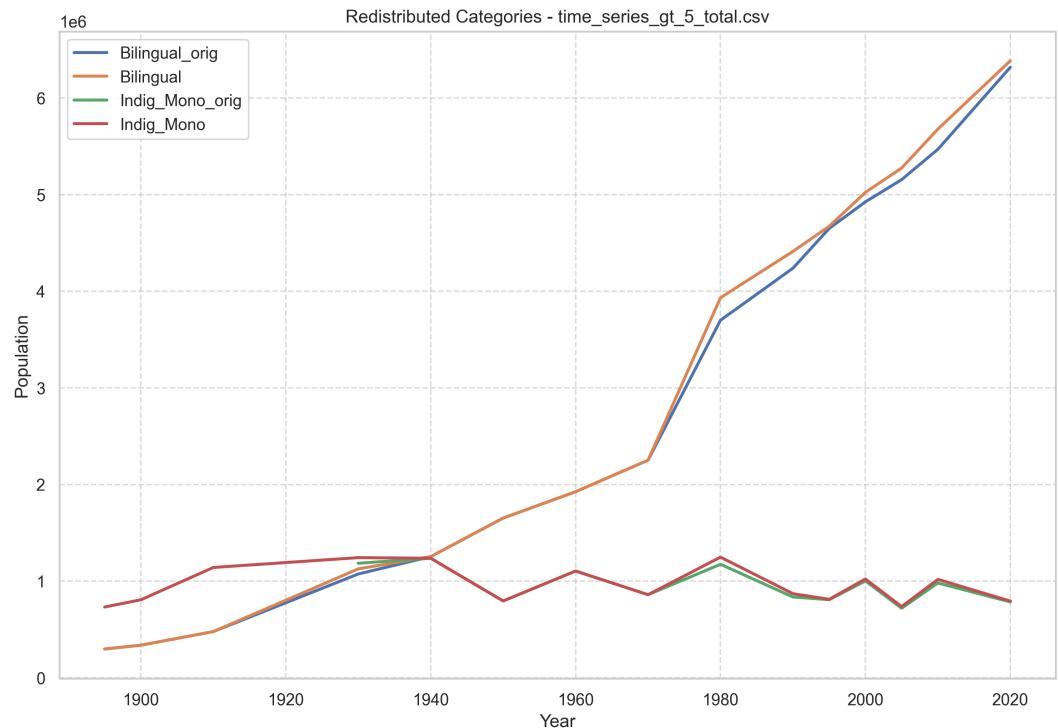


Figure B.3: Comparison of Original and Redistributed Trends for Bilingual and Monolingual Speakers within the Indigenous category.

C Data Processing

We used the logistic fit of equation (B.1) to calculate the growth rate (r), the initial value (N_0), and the carrying capacity (K) for the categories under investigation. We also figured out the R-squared (R^2) measure to verify the behavior of the fitting process. Table C.1 reports the results.

Table C.1: Logistic curve parameters and R^2 .

N ₀	K	r	(R ²)	Category
3939935	183391900	0.035	0.996	Population
272030	8615939	0.036	0.990	Bilingual
1012848	12918933	0.022	0.974	Indigenous
3117878	165995301	0.037	0.996	Spanish

We expected the model to correctly explain the variance of the Population and Spanish categories (in fact, R^2 is over 99 percent in these cases). However, we are surprised by the R^2 value of 97 percent for the Indigenous category, given the objective difficulties in obtaining this information in a country as large and ethnologically complex as Mexico.

Below are the current values and their corresponding logistic curves⁷, both for the years studied (Figure C.1) and for a more ample version (1500 to 2300), to see how the current data fit a curve with a more extended temporal perspective, Figure C.2.

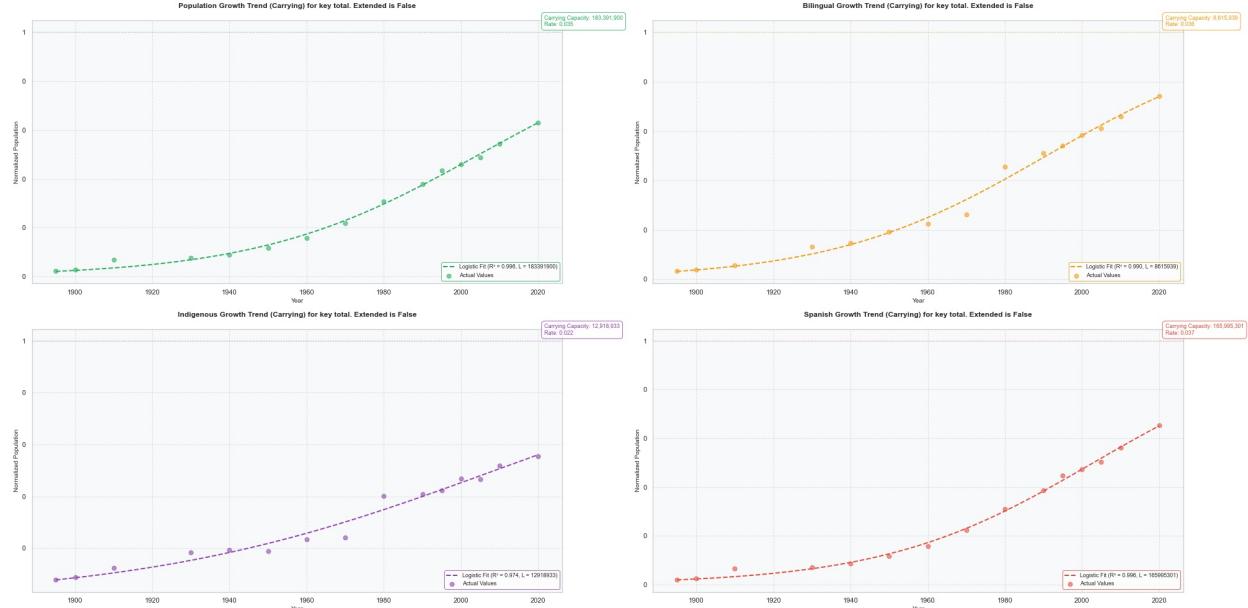


Figure C.1: Logistic curve and actual values for Population, Spanish, Bilingual, and Indigenous from 1895 to 2020.

⁷The values are divided by their carrying capacity (K).

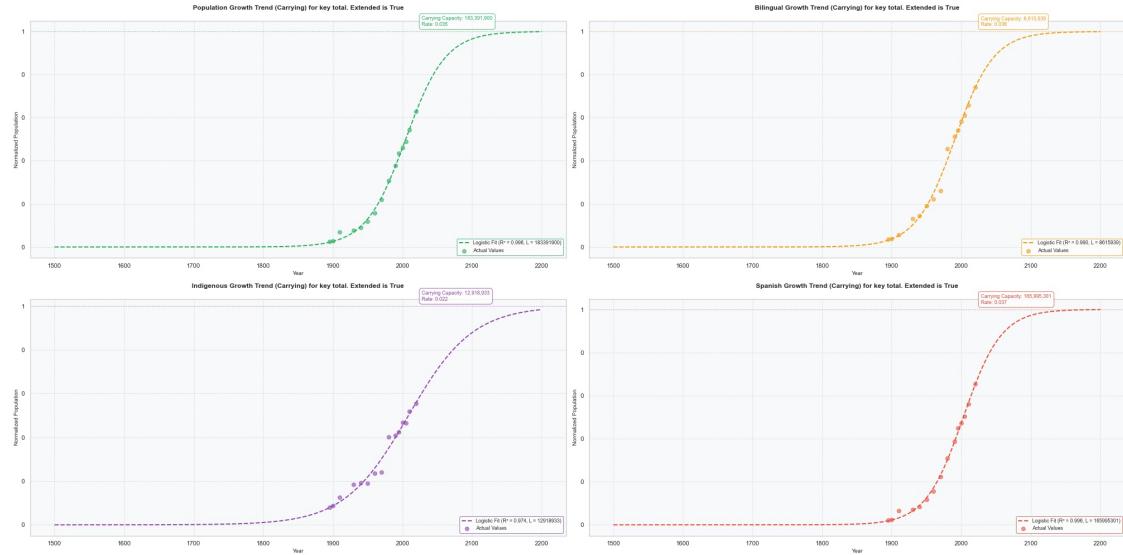


Figure C.2: Logistic curve and actual values for Population, Spanish, Bilingual, and Indigenous from 1500 to 2300.

The R^2 value obtained for Indigenous indicates that the model is a good fit; however, we further evaluated the model's performance by calculating the absolute residuals for each category across the common years. To achieve this, we applied the fitting algorithm exclusively to the initial dataset without filling in the missing values described in Section B. We derived the characteristic parameters r , N_0 , and K from this. We then plotted the logistic curves of the original data alongside the corresponding residuals and calculated various statistical measures. Notably, the maximum percentage of residuals for the category "Population" reached 7.8% in 1939, while for the "Indigenous" category, it was 2.9% in 1932.

Figures C.3 and C.4 below report the obtained plots along with some statistical measures.

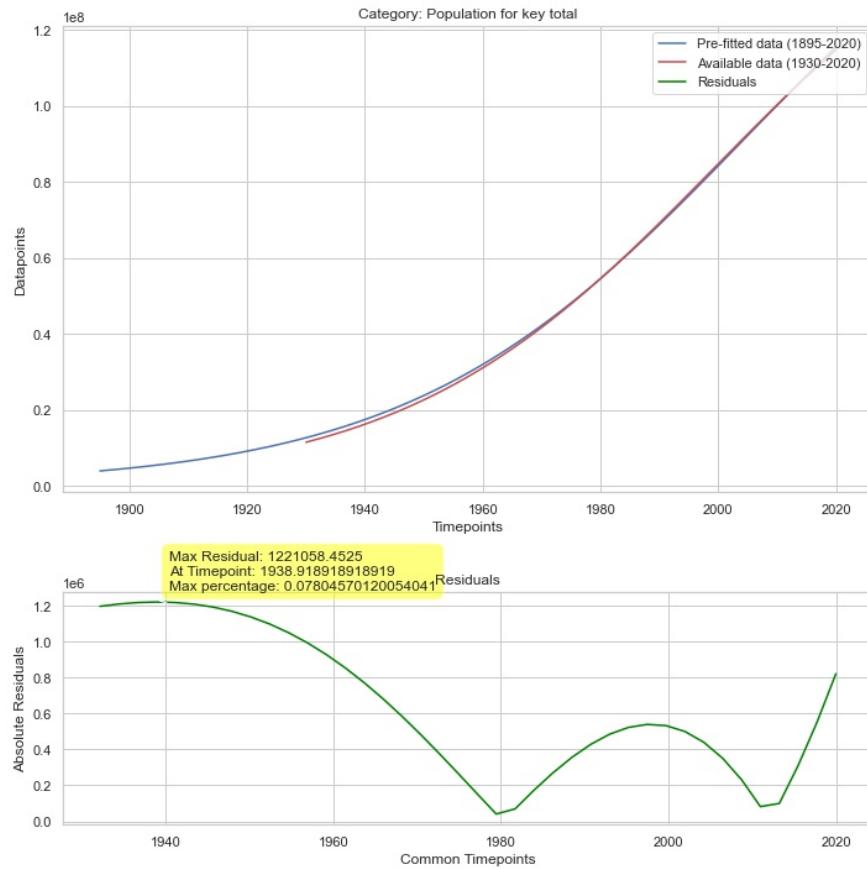


Figure C.3: Pre-fitted vs available data for Population. In 1939, they differ by no more than 8 percent.

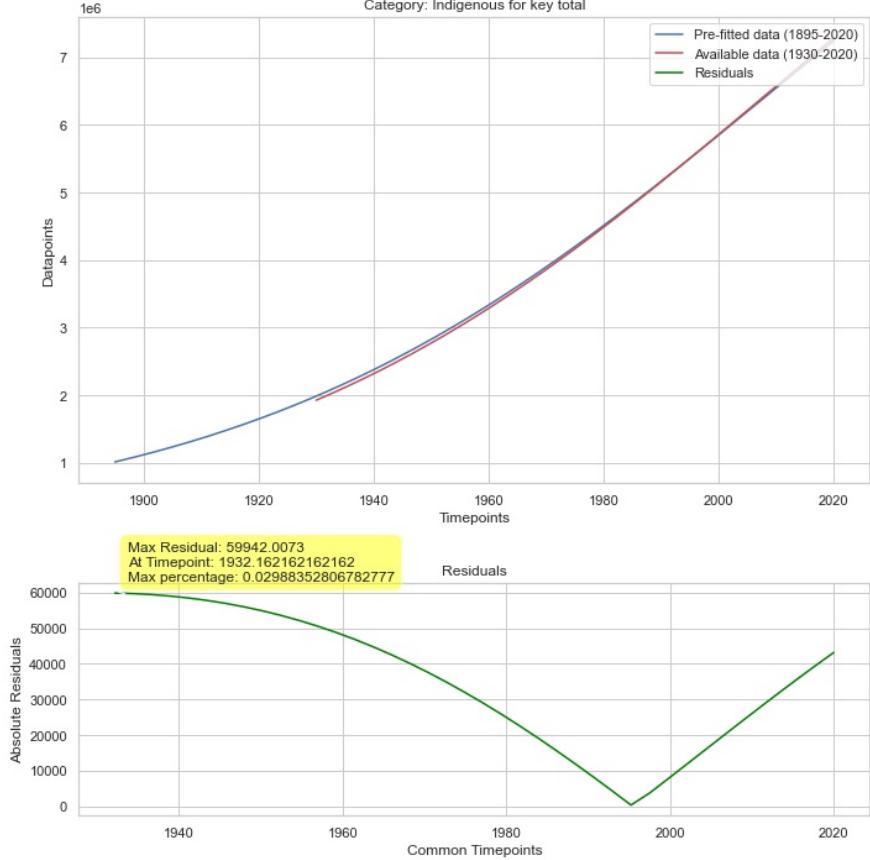


Figure C.4: Pre-fitted vs available data for Population. In 1932, they differ by no more than 3 percent.

D Data Analysis

As stated in Section 1, the primary purpose of this study is to model transitions to and from Bilingual speakers as a dynamic three-state system.

We start supposing a dynamic for the Bilingual Population, $B(t)$, which accounts for both the loss of bilingual individuals, $L(B(t))$, and the gain from language adoption transitions, $(G(I, S, t))$. The dynamic for $B(t)$ is given by:

$$B(t + \Delta t) = B(t) + L(B(t)) + G(I, S, B, t), \quad (\text{D.1})$$

Where the loss L and gain G terms include the factors λ_{ij} and μ_{ji} in Figure 1.

According to the Figure, the categories S (representing Spanish speakers) and I (representing indigenous language speakers) are independent, and each has its growth rate, denoted as r_s and r_i respectively (refer to Table C.1). Meanwhile, B changes not only due to its intrinsic growth rate r_b (also see Table C.1) but also from the transitions of Indigenous people who start speaking Spanish and bilingual speakers who turn back to their original language. Additionally, Section 3 clarifies that the Bilingual category is a subset of the Indigenous category, consisting only of indigenous speakers who are also Spanish speakers.

In this Section, we examine the following aspects of the data

- trends of Bilingualism compared to those of Spanish and Indigenous;
- correlations between Bilingualism and the ratio of Spanish and Indigenous.

Figure D.1 shows three possible trends⁸ for bilingual speakers within the Indigenous category. Bilingual

⁸The plot in Figure D.1 serves as an example.

speakers may trend similarly to the Indigenous category (red line) and then drift up (orange dots), down (blue dots), or take the same values, like the green dots. From our data, we need to figure out which situation to interpret.

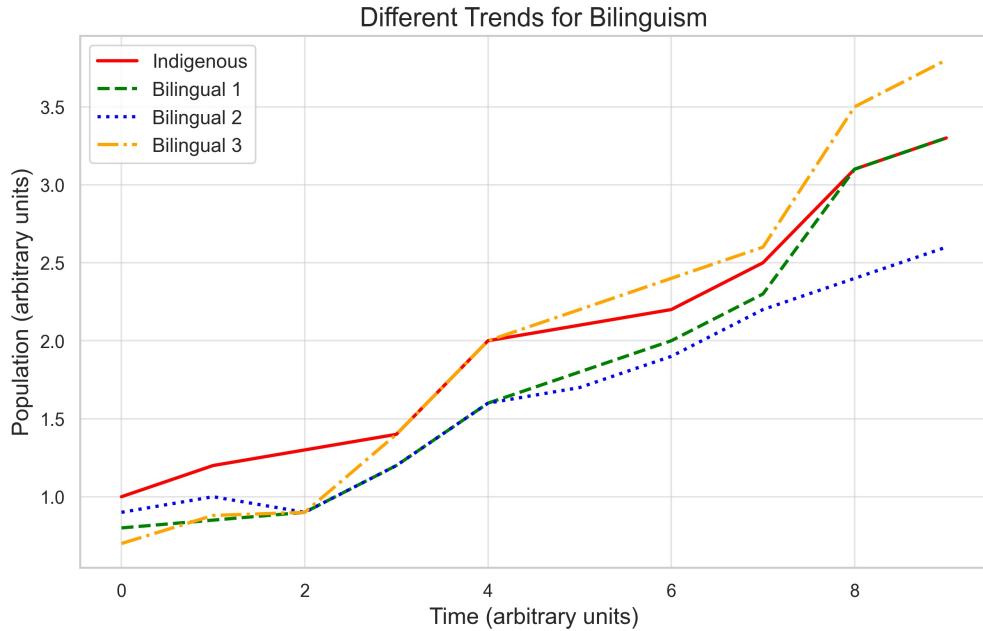


Figure D.1: Three possible trends for bilingual speakers within the Indigenous category.

D.1 Plots interpretations and insights

Figure D.2 compares Bilingual with Indigenous, with the help of four distinct but connected plots. The plot in the upper left of Figure D.2 compares the Bilingual group to the Indigenous logistic curve and shows that Bilingual and Indigenous have followed a similar trajectory over the years, indicating they are closely related. We are in the situation of the green dots in Figure D.1. One possible explanation is that as soon as social dynamics become relevant, many Indigenous start adopting Spanish as a second language. However, the converging trends in recent years indicate that many families are already Bilingual.

The Q-Q plots for residuals also support the strong correlation between Bilingual and Indigenous. The residuals are calculated as the difference between the original Bilingual values and the fitted values from the logistic models for Indigenous, with a mean of 0.034 for their absolute values. This value indicates no significant deviations from the logistic model.

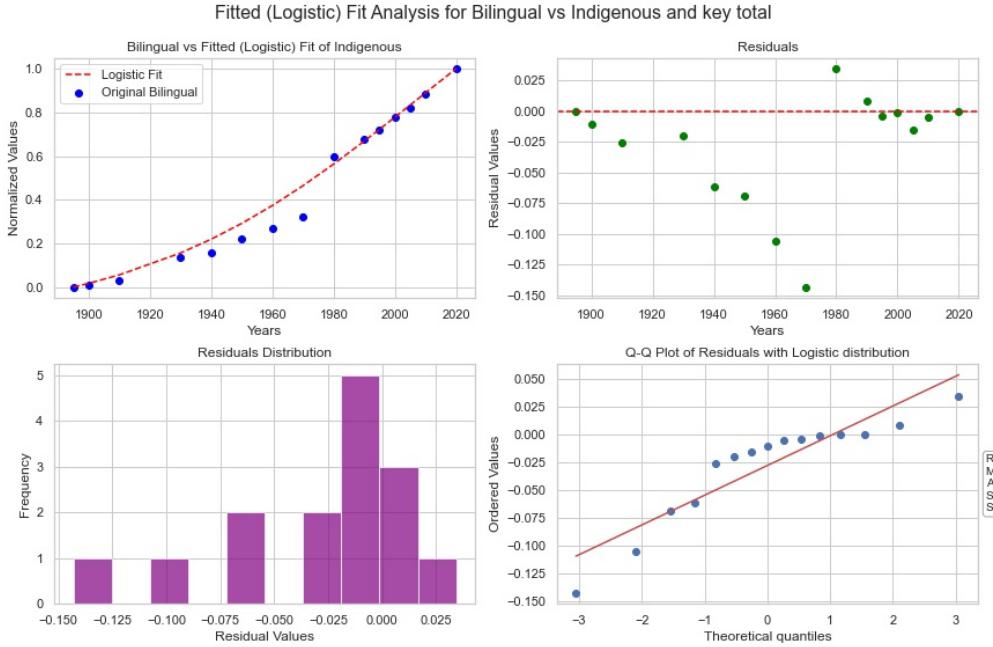


Figure D.2: Bilingual vs Indigenous. The upper left plot (normalized to 1) shows how Bilingual values fit the logistic curve of Spanish. The lower right plot shows the Q-Q plot.

For completeness of analysis, we calculated correlations between Bilingual and Spanish to verify if they share the same behavior with Indigenous. The plot at the upper left of Figure D.3 compares the Bilingual group to the Spanish logistic curve. Only a few bilingual (blue) points are close enough to the curve and share the same trend later than Indigenous. In addition, the mean absolute residual value is 0.0414 for Bilingual vs. Spanish, indicating that the original Bilingual data has more significant deviations from the logistic model fit for Spanish.

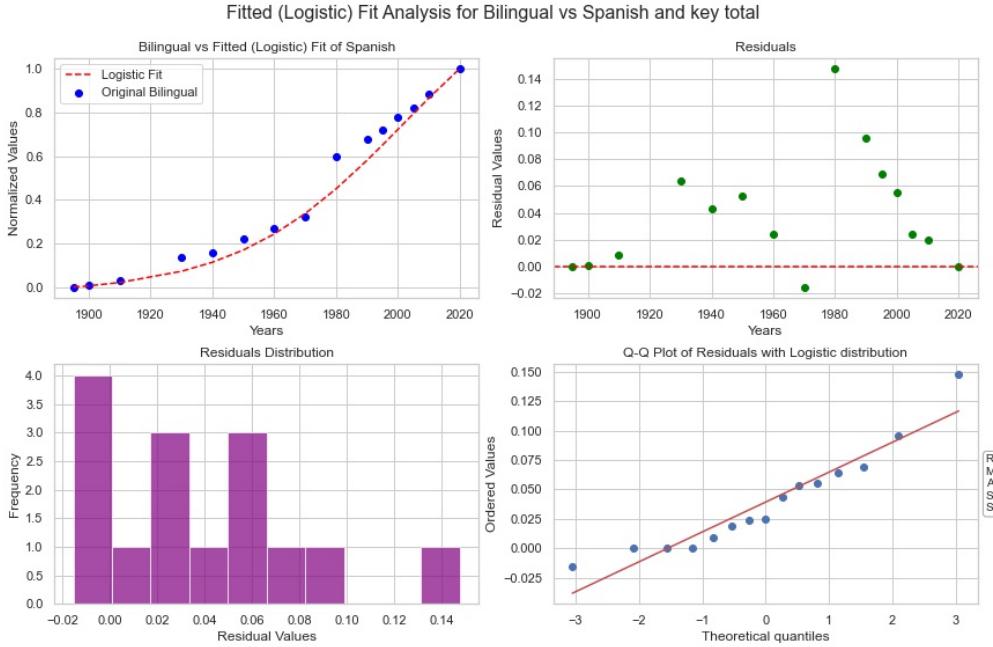


Figure D.3: Bilingual vs Spanish. The upper left plot (normalized to 1) shows how Bilingual values fit the logistic curve of Spanish. The lower right plot shows the Q-Q plot.

D.2 Trends analysis

An additional line of inquiry in our data concerns trends in the percentages of bilingual and monolingual speakers, calculated within the Indigenous category and across the Population. The upper plot in Figure D.4 shows the trend of the percentages of bilingual and monolingual speakers limited to the Indigenous category. In contrast, the lower plot displays their percentages across the total Population.

The pattern of the percentages in the upper plot resembles a typical pattern for a transition adoption process, where the increase of one pattern pairs with the decrease of the other.⁹ On the other hand, the lower plot indicates a stabilization trend for both bilingual and monolingual speakers. A linear regression analysis resulted in a near-zero slope for the last five censuses: the values of -0.0143 and -0.0141 for bilingual and monolingual speakers (respectively) indicate low variation in recent years and an overall tendency to decrease. These values are coherent with a negative slope (about -0.0330) for the Indigenous Population, while over the same period, monolingual Spanish speakers grow linearly with a positive slope of 0.0330.

⁹From the plot, it emerges that the bilingual percentage reveals a more classic adoption pattern when compared to the slightly more complex curve for monolinguals.

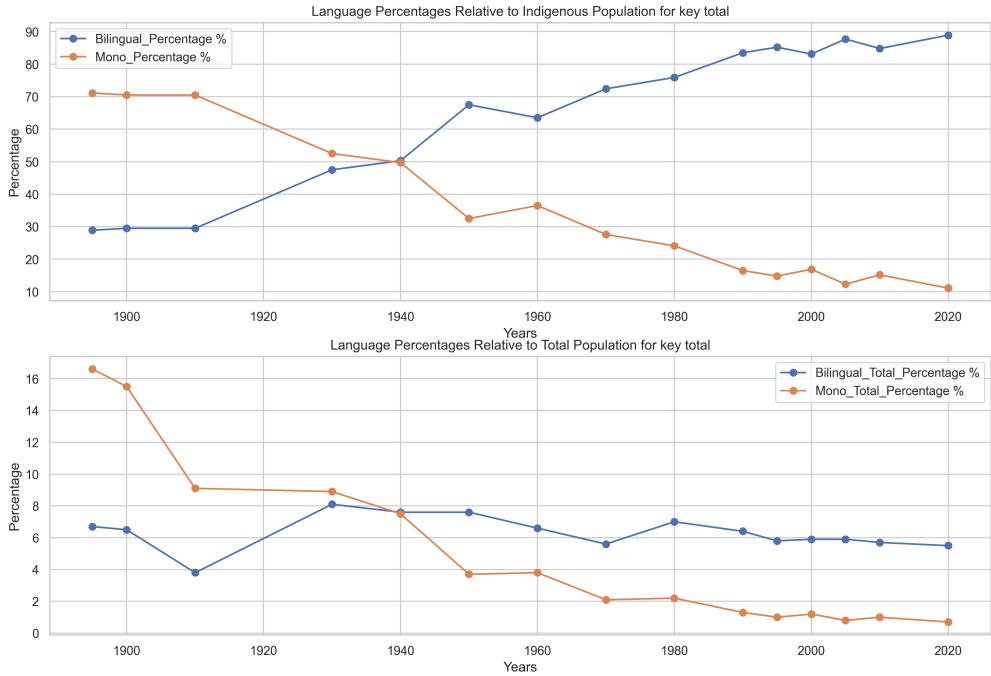


Figure D.4: Bilingual and monolingual percentages calculated on Indigenous (upper plot) and Population (lower plot) categories.

D.3 Metrics and correlations

In a scenario where a dominant language, such as Spanish, is present, It is important to analyze how the ratio of Spanish speakers to Indigenous speakers affects the growth rate of the bilingual Population within the Indigenous community. Additionally, we should consider whether the proportion of the Indigenous Population compared to the total Population is related to trends over time in the overall percentages of bilingual and monolingual speakers.

For this purpose, we define the following metrics:

$$M_{si} = \frac{N_S}{N_I} \quad (\text{D.2a})$$

$$M_{pi} = \frac{N_I}{N_P} \quad (\text{D.2b})$$

where M_{si} is the ratio between the number of Spanish, N_S , and Indigenous, N_I , speakers, while M_{pi} measures the ratio between the number of Indigenous speakers w.r.t the overall population N_P .

D.3.1 M_{si} : Spanish to Indigenous metric

In this Section, we analyze the correlation between the metric M_{si} and the percentage of bilingual individuals for the entire census period from 1895 to 2020, Figure D.5, as well as for the last five censuses, Figure D.6. We will also compare these correlations with the trends in Figure D.4.

Figures D.5 and D.6 present the same information over different time scales. Each Figure includes five connected plots with statistical annotations. The plot in the upper left corner of both figures illustrates the correlation between the metric (D.2a) and the bilingual percentage for the entire period 1895 – 2020, as well as for the last five surveys. The result shows a strong and statistically significant correlation: the correlation coefficient is $r = 0.939$ (with $p < 0.001$) for the total period and $r = 0.890$ (with $p = 0.043$) for the last five surveys, demonstrating a clear upward trend.

The time series comparison in Figure D.5 shows a good alignment between predicted and actual values, capturing long-term trends and short-term variations. In contrast, the time series in Figure D.6 exhibits more fluctuations. Finally, both Q-Q plots indicate that the points closely follow the diagonal line, demonstrating an acceptable adherence to normality.

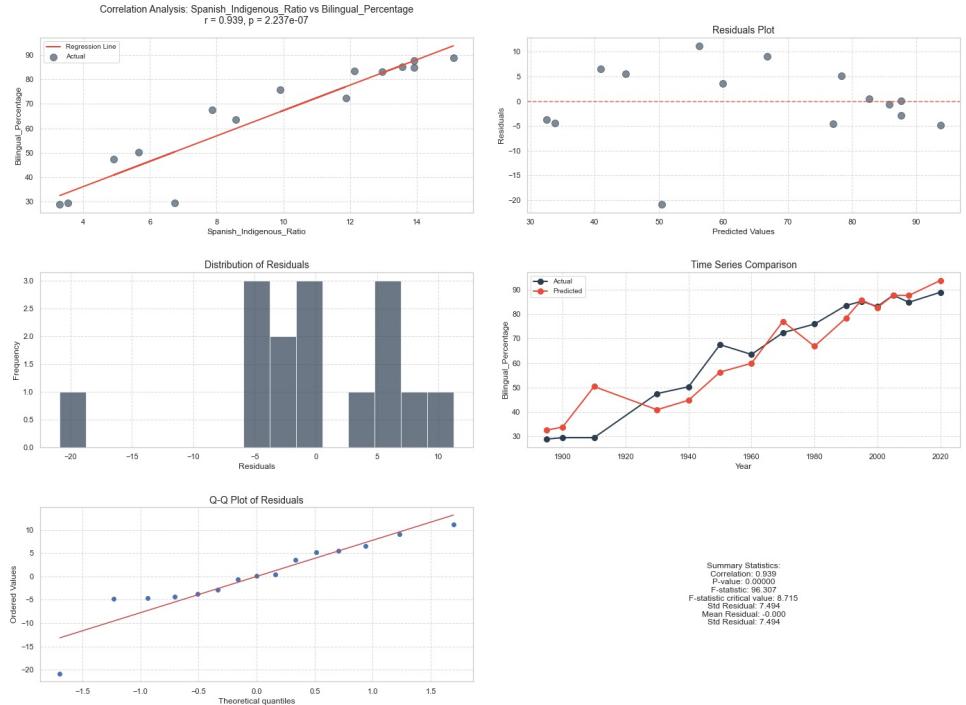


Figure D.5: Correlation between M_{si} and bilingual percentage, 1895 – – 2020.

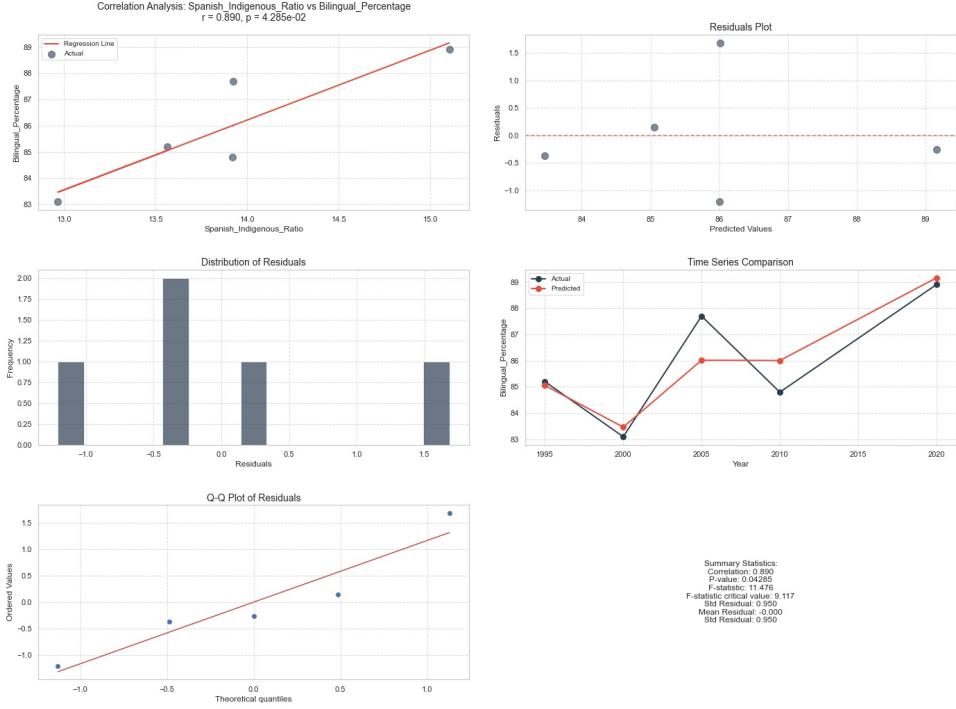


Figure D.6: Correlation between M_{si} and bilingual percentage, last five censuses.

It is interesting to compare Figure D.5 with D.4 to check the consistency of a language adoption process. The upper plot in Figure D.4 displays the shift from predominantly monolingual (orange line, $\sim 75\%$ in 1900) to largely bilingual (blue line, $\sim 90\%$ by 2020). The plot matches the time series of the growing bilingual percentage as shown in Figure D.5.

The correlation analyses shown in Figures D.5 and D.6 (even for a narrower time frame) are more informative when viewed alongside Figure D.4 rather than in isolation. This combined visualization clarifies the strong correlations observed: there is a historical trend of increasing Bilingualism among Indigenous populations, while monolingualism has steadily declined during the same period. The lower graph in Figure D.4 illustrates similar trends as a percentage of the total Population, showing a general decline over time. The rate of bilingual individuals remains around 6 percent, while the percentage of monolingual individuals is less than 2 percent.

D.3.2 M_{pi} : Population to Indigenous metric

In this Section, we analyze the correlation between M_{pi} and the percentage of bilingual and monolingual individuals w.r.t the overall Population for the entire census period from 1895 to 2020.

Figures D.7, for bilingual, and D.8, for monolingual, present the same information. Each Figure includes five connected plots with statistical annotations. The plot in the upper left corner of both figures illustrates the correlation between the metric (D.2b) and the bilingual (monolingual) percentage for the entire census period. Figure D.8 shows a strong and statistically significant negative correlation with coefficients of $r = -0.912$ ($p < 0.001$) between the same metric and the monolingual percentage. As the regression line shows, the linear relationship indicates an inverse relation. The Q-Q plot suggests that the points closely follow the diagonal line, demonstrating a satisfactory adherence to normality. The time series comparison shows a reasonable alignment between predicted and actual values, even though the last censuses diverge.

On the contrary, Figure D.7 shows a weak negative, not statistically significant, correlation with coefficients of $r = -0.43$ ($p = 0.11$) between the metric and the bilingual percentage. This correlation is not linear and more scattered, as reinforced by the Q-Q plot. Noticeably, the time series comparison shows a good alignment between predicted and actual values for the last censuses.

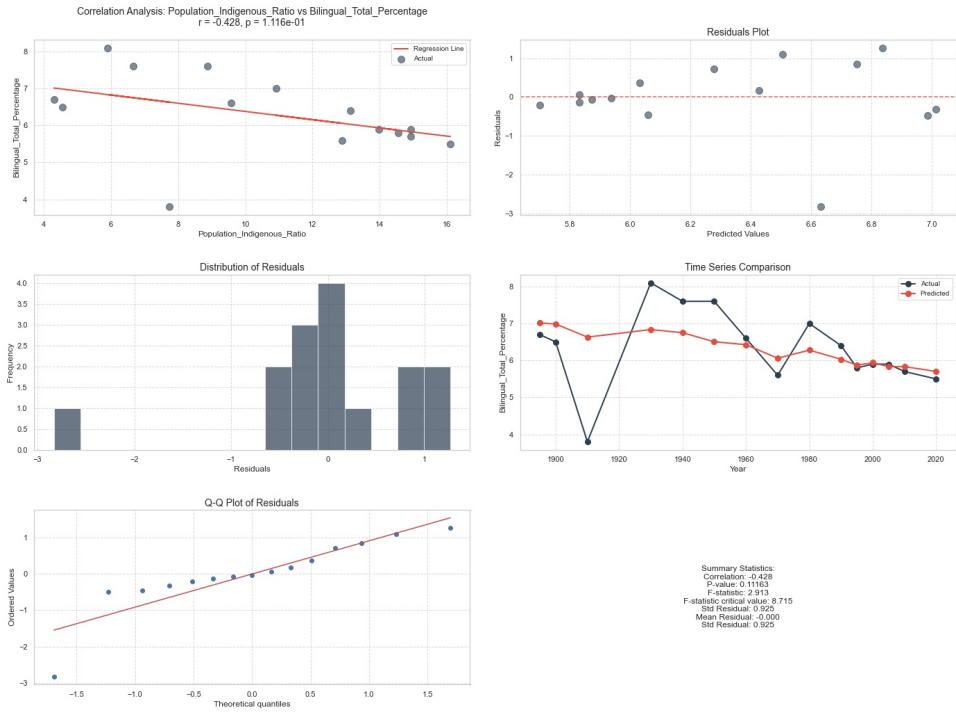


Figure D.7: Correlation between M_{pi} and bilingual percentage w.r.t the overall population, 1895 – –2020.

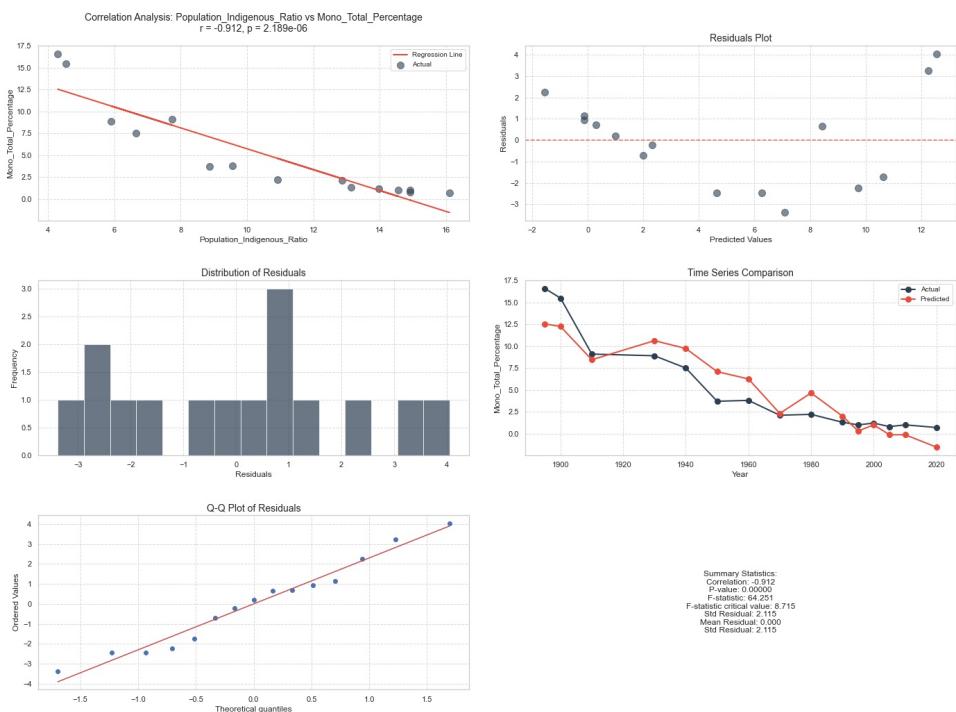


Figure D.8: Correlation between M_{pi} and monolingual percentage w.r.t the overall population, 1895 – –2020.

For complete information, we report the values of correlation coefficients along with p values for the last five censuses. The coefficient r for the percentage of monolingual speakers w.r.t the overall Population is $r = -0.911$, with $p \sim 0.04$, thus in line with the previously obtained value of $r = -0.912$. In contrast, the coefficient for the percentage of bilingual speakers is $r = -0.87$ with $p \sim 0.05$. This last result reinforces the inverse correlation between percentages and the chosen metric. It also explains the alignment of the latest censuses for the percentage of bilingual speakers.

D.4 Data Analysis: conclusions

As a result of the data analyses in the previous sections, we conclude that the dynamics of adopting Spanish as a second language are highly complex. A pattern of language adoption is visible within the Indigenous Population, where the increase of one pattern pairs with the decrease of the other. However, the stabilization of the percentage of bilinguals in recent censuses suggests that the number of already bilingual families is growing. The growing number of bilingual speakers within the Indigenous Population depends both on the phenomenon of language adoption and on the fact that many Indigenous people are already born bilingual. The decline in the percentage of monolingual Indigenous people is due much more to the decrease in the total number of Indigenous people than to a phenomenon of language adoption.

Figure D.9 compares trends in the percentages of bilingual and monolingual speakers between the historical data and data extracted from the logistic fit. Using a threshold equal to the average of the percentage increases in both series, we see that the percentage of bilingual speakers alternates in increasing and decreasing periods and can be considered stable from 1995. Meanwhile, the rate of monolingual speakers constantly decreased and became stable at around 2010.

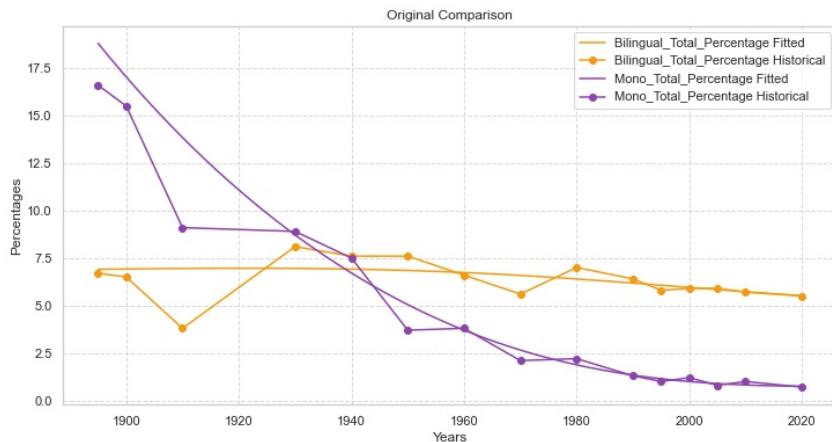


Figure D.9: Comparison of bilingual and monolingual percentages: historical vs fitted time series

The resulting dynamic model must consider several temporal regions. In light of the analyses, it is not easy to propose a single all-inclusive model of logistic population growth and the phenomenon of language adoption for the entire period.

E Parameters estimation and optimization strategies

Our dataset comes from thirteen censuses, typically conducted every ten years. Additionally, we have included data from two intercensal surveys. While the available data is limited, it provides an opportunity to enhance our understanding of how it may align with theoretical models. It is essential to acknowledge that, due to the limited nature of the data, the initial parameters chosen and any constraints applied to

these parameters (as used by specific optimization algorithms) may affect our results. In this section, we aim to establish a constructive preliminary estimate for both the parameters and the constraints, which will improve our fitting processes. We can estimate the bounds for each parameter (a , s_l , and s_o) in the Abrams-Strogatz model (reported in (E.1)) based on the observed properties of minority speakers x_l .

$$\frac{dx_l}{dt} = s_l(1 - x_l)x_l^a - s_o(1 - x_l)^a x_l \quad (\text{E.1})$$

Figure E.1 illustrates the trends in the percentage of monolingual speakers from 1895 to 2020. The figure highlights four important regions that reveal the uneven trends in speaker percentages.

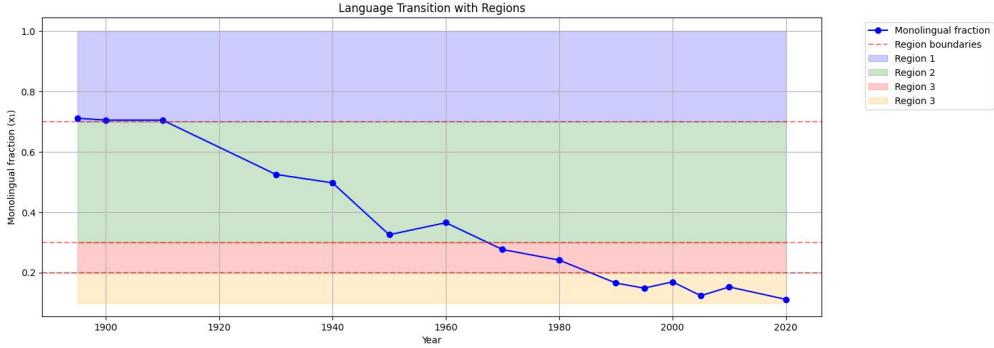


Figure E.1: Monolingual percentages trends over the entire 1895 – –2020 period.

E.1 Volatility estimation

This observation suggests that setting the parameter a close to 1 may enhance our ability to effectively capture both linear and moderately nonlinear transitions. We defined the bounds for a as $a \in [0.5, 2]$. Figure E.2 illustrates the polynomial fit x_l^a of the percentage of monolingual speakers for various volatility parameters a . The plot indicates that the linear and quadratic models provide the best fit for capturing the final trends in the data. Although the cubic model achieves the highest R^2 value of approximately 0.98, it fails to accurately represent the final trend of the monolingual percentage.

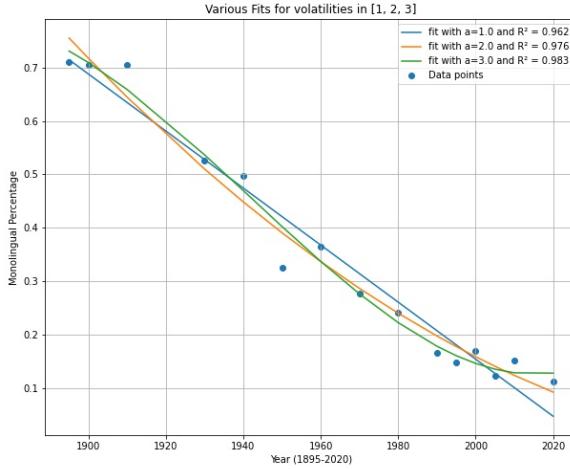


Figure E.2: Polynomial fit of actual data.

The lower bound $a = 0.5$, which allows for sublinear interactions and enables minority groups to influence the dynamics as the population grows, is not shown in the plot, as it appears as a horizontal line. Conversely, the upper bound $a = 2.0$ accounts for potential superlinear effects while restricting extreme nonlinearity, particularly associated with $a = 3.0$, which the data trends do not indicate.

E.2 Prestige estimation

The prestige factors s_l and s_o in (E.1) measure how quickly each population responds to changes in their respective proportions. Specifically, s_l represents the response of the monolingual speaker population to changes caused by the influence of bilingual speakers. Similarly, s_o reflects the response of bilingual speakers to changes influenced by monolingual speakers:

$$s_l = \frac{k}{x_l^a(1-x_l)} \quad (\text{E.2a})$$

$$s_o = \frac{k}{x_l(1-x_l)^a} \quad (\text{E.2b})$$

The constant k is related to the value of $|dx_l/dt|$ which we calculate from the dataset as:

$$\frac{dx_l}{dt} \approx \frac{x_l(t_{i+1}) - x_l(t_i)}{t_{i+1} - t_i}$$

The values of $\frac{dx_l}{dt}$ represent the observed dynamics of the data, and it is bounded by the value of

$$k_{max} \approx 1.72 \cdot 10^{-3}$$

E.2.1 s_x estimation algorithm

In the Abram-Strogatz model, see (E.1), s_l and s_o represent the prestige or status of the minority and majority languages, respectively. One of the fundamental assumptions of the model is that the minority language should have lower prestige than the majority language. The algorithm manages the order of the prestige values by swapping the values of s_o with the values of s_l if the former is less than the latter.

Algorithm 1: Estimation of s_l and s_o

```

1 foreach For each volatility a do
2   foreach For each time step  $t_i$  do
3     Calculate  $s_l$  according to (E.2a) ;
4     Calculate  $s_o$  according to (E.2b);
5     if  $s_l > s_o$  then
6       Swap( $s_l, s_o$ );
7     Calculate the mean, median, lower bound, and upper bound of  $s_l$  and  $s_o$  using the range 10%
    - 90% percentiles;
8 return datasets for initial parameters and bounds ds_param, ds_param_by_a, ds_bounds,
  ds_bounds_by_a;
```

Where for the initial parameters:

ds_param For each volatility level a , this dataset includes the absolute minimum and maximum values of the median of s_l and s_o . In this set of initial parameters, the only variable parameter is the volatility level a ;

ds_param_by_a For each volatility level a , this dataset contains the minimum and maximum values of the median of s_l and s_o that correspond to a . In this set of initial parameters, all the parameters a , s_l , and s_o are variable.

While for bounds:

ds_bounds This dataset contains fixed values for a (refer to Section E.1). For the variables s_l and s_o , the lower and upper bounds are defined by setting them equal to the absolute minimum lower bound and the maximum upper bound of s_l and s_o , respectively;

ds_bounds_by_a For each volatility value a , this dataset fixed values for a (refer to Section E.1). For the variables s_l and s_o , the lower and upper bounds are defined by setting them equal to the absolute minimum lower bound and the maximum upper bound of s_l and s_o .

Below are two useful figures to visualize the results of the algorithm. Figure E.3 shows the mean, the median, and the Kernel Density Estimation (KDE) of s_l and s_o , along with the means, medians, and percentile bounds.

The distributions are constrained, showing a more precise separation between the prestige values of s_l and s_o . The s_l distribution, at the top right of the Figure, has lower density peaks and is flat. On the contrary, the s_o distribution (bottom right) is sharper and has higher KDE peaks. Both distributions are right-skewed, with the mean higher than the median. This type of skewness justifies the median of the prestige values as initial parameters.

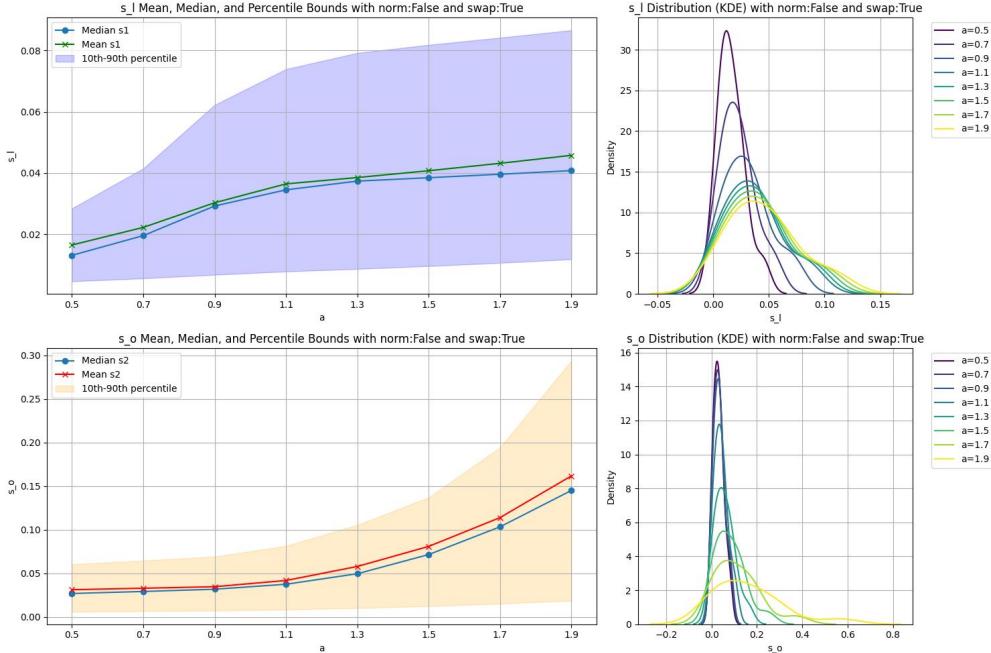


Figure E.3: Plots of s_l and s_o (scaled by k_{max}) for $a \in [0.5, 2]$.

The left plot of Figure E.3 shows the prestige values' means, medians, and percentile bounds. The forced swap between s_l and s_o generates a broader range of possible values for s_l compared to the range of s_o in the same percentile.

Figure E.4 below illustrates a distinct asymmetry between the prestige values of majority and minority languages, which is coherent with the KDE distributions and the forced swap between s_l and s_o .

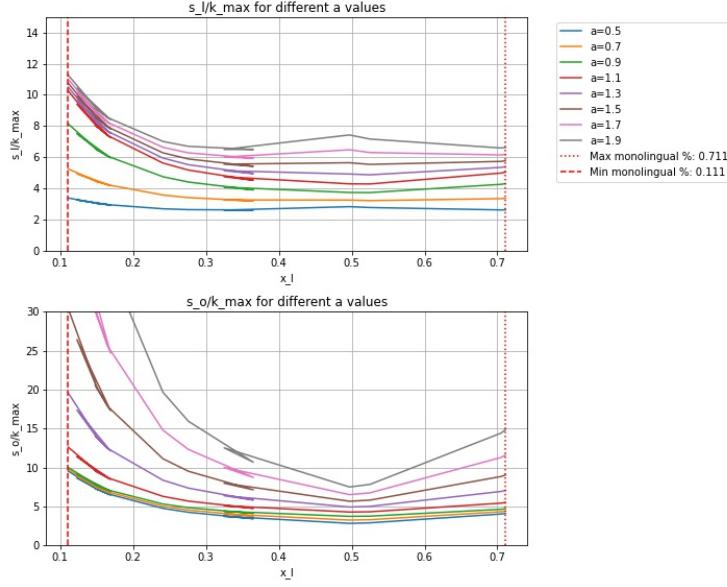


Figure E.4: Plots of s_l and s_o (scaled by k_{max}) for $a \in [0.5, 2]$.

E.3 Normalized prestige estimation

In literature, the Abrams-Strogatz model has the following structure:

$$\frac{dx_l}{dt} = s(1 - x_l)x_l^a - (1 - s)(1 - x_l)^a x_l \quad (\text{E.3})$$

$s \leq 0.5$ is the prestige of the minority language, and $1 - s \geq 0.5$ is the prestige of the majority language. In this case, we calculate the prestige factors as follows:

$$term_1 = x_l^a(1 - x_l) \quad (\text{E.4a})$$

$$term_2 = x_l(1 - x_l)^a \quad (\text{E.4b})$$

$$s_l = \frac{\frac{dx_l}{dt} + term_2}{term_1 + term_2} \quad (\text{E.4c})$$

$$s_o = 1 - s_l \quad (\text{E.4d})$$

Derivation of (E.4c)

From (E.3), (E.4a), and (E.4b):

$$\begin{aligned} \frac{dx_l}{dt} &= s(1 - x_l)x_l^a - (1 - s)(1 - x_l)^a x_l = s \cdot term_1 - (1 - s) \cdot term_2 \\ &= s(term_1 + term_2) - term_2 \\ s &= \frac{\frac{dx_l}{dt} + term_2}{term_1 + term_2} \end{aligned}$$

We used the following algorithm, similar to 1 of Section E.2:

We present two pictures that visualize the results of the algorithm. These images convey the same information as Figures E.3 and E.4, but they demonstrate an evident symmetry.

Algorithm 2: Estimation of s and $1 - s$

```

1 foreach For each volatility  $a$  do
2   foreach For each time step  $t_i$  do
3     Calculate  $s_l$  according to (E.4c) ;
4     Calculate  $s_o$  according to (E.4d);
5     if  $s_l > s_o$  then
6       Swap( $s_l, s_o$ );
7     Set  $s = s_l$ ;
8     Set  $1 - s = s_o$ ;
9     Calculate the mean, median, lower bound, and upper bound of  $s$  and  $1 - s$  using the range
10    10% - 90% percentiles;
11
12 return datasets for initial parameters and bounds ds_param, ds_param_by_a, ds_bounds,
  ds_bounds_by_a;

```

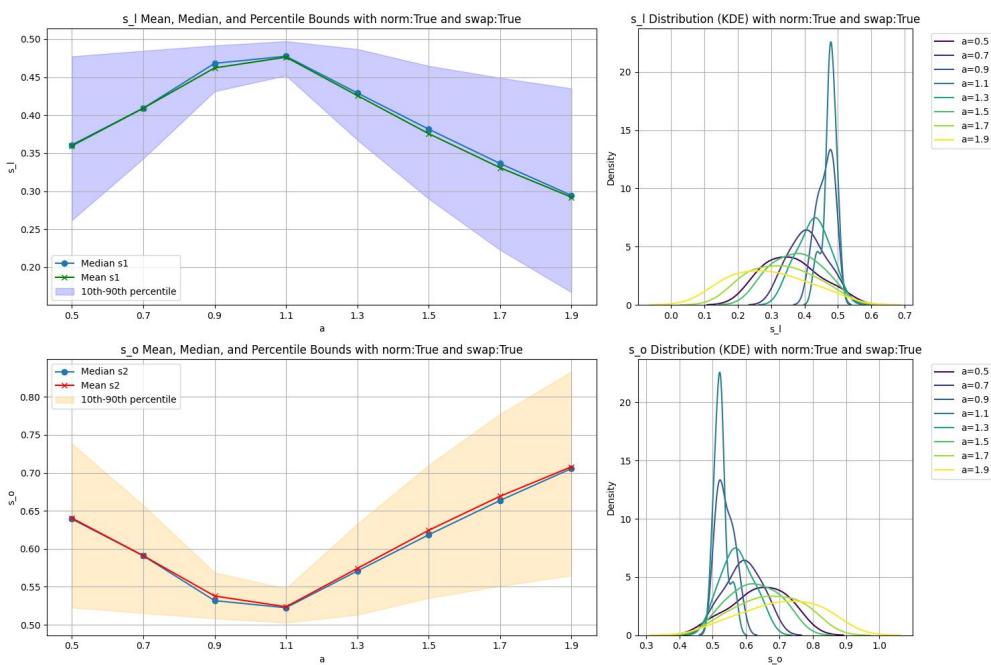


Figure E.5: Plots of s and $1 - s$ for $a \in [0.5, 2]$.

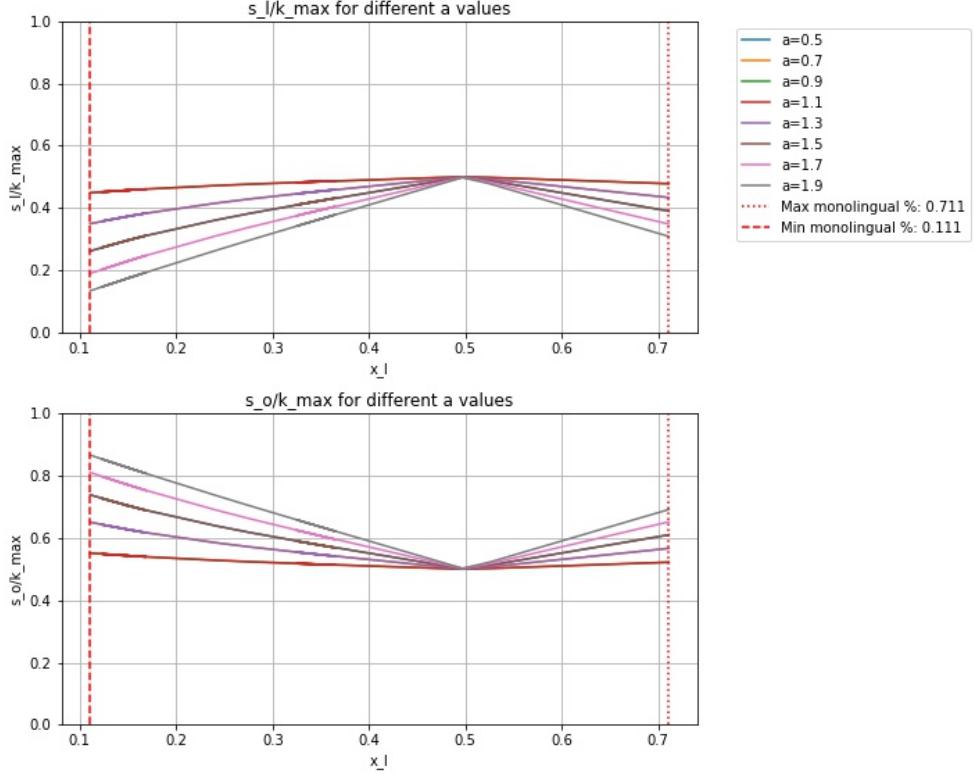


Figure E.6: Plots of s and $1 - s$ (scaled by k_{max}) for $a \in [0.5, 2]$.

E.4 Loss functions and execution strategies

In this section, we outline the strategies for evaluating the proposed models. The data fitting process employs two loss functions and three algorithms, L-BFGS-B [15], Powell [18], and Nelder-Mead [17], to find their minima. The loss functions include a standard one that minimizes the sum of squared deviations and a weighted function that adjusts the squared deviations based on the gradient of the actual value calculated at the specific point x_l of interest. The following algorithm sketches the minimization process.

Algorithm 3: Minimization process

```

1 foreach For each initial parameter param do
2   Set  $a = param[0]$ ;
3   Set  $s_l = param[1]$ ;
4   Set  $s_o = param[2](*)$ ;
5   Set bounds corresponding to  $a$ : bounds=bounds[a] (**);
6   foreach For each minimization algorithms  $m_a$  do
7     foreach For each loss function fun do
8       Fit the model parameters vs. actual data;
9       Create a list of minimization  $L_m$ ;
10    Calculate the best minimization parameters from  $L_m$ ;
11    return initial parameters, initial bounds, optimal parameters, loss function, loss value;

```

(*)

In the case of normalized prestige values $s_o = 1 - s_l$, the third value of the initial parameters is not used.

(**)

When the bounds are not influenced by the volatility parameter a , they maintain consistent values throughout the loop. However, if the bounds are affected by the volatility parameter, each specific bound corresponds with its respective volatility.

It is important to notice that the minimization algorithm (see (Alg.E.3)) works with three different strategies used to verify to which extent the optimization process is sensible to the initial parameters and to their bounds. The first approach uses the initial parameters and their bounds as they are calculated from (Alg.E.1) and (Alg.E.2); the second uses an *unbounded* space; the last approach implements the SciPy¹⁰ *differential_evolution* which does not use the gradient to find the minima of the loss function.

Unbounded space

Given a parameter p and its lower and upper bounds, L and U , the following p_x is sent to the minimization process:

$$p_x = \log \left(\frac{U - s}{s - L} \right)$$

During the minimization process, the p_x is inverted:

$$p = L + \frac{U - L}{1 - e^{-p_x}}$$

before p is sent to the model for evaluation.

This strategy is useful to prevent p from being neither U nor L .

E.5 Error estimation for relevant model parameters

In this section, we describe how to compute an estimate of the errors on the values of the volatility a and the parameters of the prestige s_l and s_o for the Abrams-Strogatz Model.

The strategy described in this section utilizes a multi-method approach to estimate parameter errors. The method integrates three primary techniques: Hessian matrix computation, surface likelihood profiling, and bootstrapped resampling.

The estimation process starts with a Hessian matrix computation to get a local quadratic approximation of the loss function's curvature around the optimal parameter. This computation calculates the inverse of the Hessian matrix and computes the standard errors by taking the square root of the diagonal elements.

The Hessian-based estimate may fail or produce inconsistent results; for example, the non-quadratic shape of the loss function near the minimum causes the errors to be either negative or greater than the optimal parameter. In this case, the estimation process executes a profile likelihood method. This approach analyzes the local minimum variations in the loss function when the optimal parameters are narrowly modified. The lower and upper bounds of the parameters are determined by their range while keeping the variation of the local minimum within a specified threshold ($\sim 95\%$).

As in the Hessian strategy, the estimate process will use a bootstrap resampling technique if the errors are either negative or greater than the optimal parameter. This technique introduces a controlled noise to the

¹⁰https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.differential_evolution.html

data and re-estimates parameters multiple times to statistically derive parameter uncertainties. The process is sketched in (Alg.E.4).

Algorithm 4: Error estimation process. Optimal parameters, bounds, and loss function are calculated in 3.

```
1 Set params=optimal_params;
2 Set bounds=initial_bounds;
3 Set loss=best_loss_function;
4 Try Hessian(loss,params,bounds);
5 Set std_error_h=Hessian(loss,params,bounds);
6 if Hessian fails then
7   Try likelihood_profile(loss,params,bounds);
8   Set std_error_l=likelihood_profile(loss,params,bounds);
9   if likelihood_profile fails then
10    Try bootstrap(loss,params,bounds);
11    Set std_error_b=bootstrap(loss,params,bounds);
12    Set std_error=std_error_b;
13    Set process="bootstrap";
14  else
15    Set std_error=std_error_l;
16    Set process="likelihood_profil";
17 else
18  Set std_error=std_error_h;
19  Set process="hessian";
20 return std_error, process
```

References

- [1] N. Neev. "Preserving Voices: Indigenous Languages of Mexico and The Worlds Within Them." (2024), [Online]. Available: <https://psydeh.org>.
- [2] M. G. Policy. "Language and Multiculturalism Policies in Mexico." (2024), [Online]. Available: <https://www.indigenousmexico.org/articles/indigenous-mexico-and-the-spanish-language>.
- [3] J. P. Schmal. "Indigenous Mexico and the Spanish Language." (2024), [Online]. Available: <https://www.indigenousmexico.org/articles/indigenous-mexico-and-the-spanish-language>.
- [4] D. K. N. Rachim, A. Firdaus, and A. G. W. Saputro, "Analysis of the impact of population growth in DKI Jakarta using logistic model," *Jurnal Pendidikan Matematika (Kudus)*, vol. 5, no. 1, pp. 69–78, 2022.
- [5] N. Tkachenko, J. D. Weissmann, W. P. Petersen, G. Lake, C. P. E. Zollikofer, and S. Callegari, "Individual-based modelling of population growth and diffusion in discrete time," *PLOS ONE*, vol. 12, no. 4, e0176101, 2017. DOI: 10.1371/journal.pone.0176101. [Online]. Available: <https://doi.org/10.1371/journal.pone.0176101>.
- [6] X. Castelló, L. Loureiro-Porto, and M. San Miguel, "Agent-based models of language competition," *International journal of the sociology of language*, vol. 2013, no. 221, pp. 21–51, 2013.
- [7] M. Patriarca, X. Castelló, J. R. Uriarte, V. M. Eguíluz, and M. San Miguel, "Modeling two-language competition dynamics," *Advances in Complex Systems*, vol. 15, no. 03n04, p. 1250048, 2012.
- [8] R. M. Groves, F. J. Fowler, M. P. Couper, J. M. Lepkowski, E. Singer, and R. Tourangeau, *Survey Methodology*, 2nd. Hoboken, NJ: Wiley, 2009, ISBN: 978-0-470-46546-2.
- [9] F. Vazquez and V. M. Eguíluz, "Analytical solution of the voter model on uncorrelated networks," *New Journal of Physics*, vol. 10, no. 6, p. 063011, 2008.
- [10] D. Stauffer, X. Castelló, V. M. Eguíluz, and M. San Miguel, "Microscopic Abrams–Strogatz model of language competition," *Physica A: Statistical Mechanics and its Applications*, vol. 374, no. 2, pp. 835–842, 2007.
- [11] K. Sucheki, V. M. Eguíluz, and M. San Miguel, "Voter model dynamics in complex networks: Role of dimensionality, disorder, and degree distribution," *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, vol. 72, no. 3, p. 036132, 2005.
- [12] M. Patriarca and T. Leppänen, "Modeling language competition," *Physica A: Statistical Mechanics and its Applications*, vol. 338, no. 1-2, pp. 296–299, 2004. DOI: 10.1016/j.physa.2004.02.045.
- [13] D. M. Abrams and S. H. Strogatz, "Modeling the dynamics of language death," *Nature*, vol. 424, no. 6951, pp. 900–900, 2003. DOI: 10.1038/424900a.
- [14] C. Castellano, D. Vilone, and A. Vespignani, "Incomplete ordering of the voter model on small-world networks," *Europhysics Letters*, vol. 63, no. 1, p. 153, 2003.
- [15] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization," *ACM Transactions on Mathematical Software (TOMS)*, vol. 23, no. 4, pp. 550–560, 1997. DOI: 10.1145/279232.279236.
- [16] R. A. Holley and T. M. Liggett, "Ergodic theorems for weakly interacting infinite systems and the voter model," *The annals of probability*, pp. 643–663, 1975.
- [17] J. A. Nelder and R. Mead, "A Simplex Method for Function Minimization," *The Computer Journal*, vol. 7, no. 4, pp. 308–313, 1965. DOI: 10.1093/comjnl/7.4.308.
- [18] M. J. D. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," *The Computer Journal*, vol. 7, no. 2, pp. 155–162, 1964. DOI: 10.1093/comjnl/7.2.155.