

Language Projections: 2010-2020

Hyon B. Shin
Jennifer M. Ortman
U.S. Census Bureau

Presentation Prepared for the Federal Forecasters Conference,
Washington, DC, April 21, 2011

This paper is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed on statistical, methodological, technical, or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau.

Overview

- Background
- Data
- Methodology
- Results
- Future directions

Background

- Historically, language questions were asked primarily in the decennial censuses.
- From 1890-1970, various questions on languages spoken and English-speaking ability were asked.
- Since 1980, the same three questions were asked in the 1980, 1990, and 2000 censuses and are currently asked in The American Community Survey (ACS).
- Used to satisfy the Voting Rights Act of 1965.

Language Questions Asked

The three-part question that was asked in the 1980, 1990, and 2000 Censuses, and in the ACS from 2001 to the present.

Figure 1.
Reproduction of the Questions on
Language From the 2007 American
Community Survey

13

a. Does this person speak a language other than English at home?

Yes

No → SKIP to question 14

b. What is this language?

For example: Korean, Italian, Spanish, Vietnamese

c. How well does this person speak English?

Very well

Well

Not well

Not at all

Source: U.S. Census Bureau, 2007 American Community Survey.

Background (continued)

- Language write-ins are coded to about 380 detailed language categories.
- We use the classification list of 39 categories for detailed language groups.
- Languages are further classified into 4 major language groups: Spanish, Other Indo-European languages, Asian and Pacific Island languages, and all other languages.
- While Spanish is the largest non-English language spoken, there are nearly 20 million speakers of other languages.

Data

- American Community Survey (ACS)
- U.S. Census Bureau's 2008 and 2009 National Population Projections

The American Community Survey

- Data from 2006, 2007, 2008, and 2009
- We project languages with at least 500,000 speakers in 2009:

Spanish	Hindi
French	Chinese
Italian	Korean
Portuguese	Vietnamese
German	Tagalog
Russian	Arabic
Polish	

Table 1. Languages Spoken at Home: 2006-2009

Characteristic	2006		2007		2008		2009	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Population 5 Years and older	279,012,712	100.0	280,950,438	100.0	283,156,079	100.0	285,797,349	100.0
Spoke only English at home	224,154,288	80.3	225,505,953	80.3	227,295,534	80.3	228,699,523	80.0
Spoke a language other than English at home ¹	54,858,424	19.7	55,444,485	19.7	55,860,545	19.7	57,097,826	20.0
Spoke a language other than English at home	54,858,424	100.0	55,444,485	100.0	55,860,545	100.0	57,097,826	100.0
Spanish or Spanish Creole	34,044,945	62.1	34,547,077	62.3	34,615,394	62.0	35,468,501	62.1
French (includes Patois, Cajun, Creole)	1,997,618	3.6	1,984,824	3.6	1,973,531	3.5	1,964,556	3.4
Italian	828,524	1.5	798,801	1.4	782,173	1.4	753,992	1.3
Portuguese or Portuguese Creole	683,405	1.2	687,126	1.2	661,120	1.2	731,282	1.3
German	1,135,999	2.1	1,104,354	2.0	1,121,465	2.0	1,109,216	1.9
Russian	823,210	1.5	851,174	1.5	860,568	1.5	881,723	1.5
Polish	640,265	1.2	638,059	1.2	616,492	1.1	593,598	1.0
Hindi	504,607	0.9	532,911	1.0	562,587	1.0	560,983	1.0
Chinese	2,492,871	4.5	2,464,572	4.4	2,473,968	4.4	2,600,150	4.6
Korean	1,060,631	1.9	1,062,337	1.9	1,048,400	1.9	1,039,021	1.8
Vietnamese	1,207,721	2.2	1,207,004	2.2	1,236,419	2.2	1,251,468	2.2
Tagalog	1,415,599	2.6	1,480,429	2.7	1,496,208	2.7	1,513,734	2.7
Arabic	732,519	1.3	767,319	1.4	780,995	1.4	845,396	1.5
All other languages	7,290,510	13.3	7,318,498	13.2	7,631,225	13.7	7,784,206	13.6

¹ The languages listed in this table are the languages for which there were 500,000 or more speakers in 2009.

Source: U.S. Census Bureau, 2006, 2007, 2008, and 2009 American Community Survey.

Table 1. Languages Spoken at Home: 2006-2009

Characteristic	2006		2007		2008		2009	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Population 5 Years and older	279,012,712	100.0	280,950,438	100.0	283,156,079	100.0	285,797,349	100.0
Spoke only English at home	224,154,288	80.3	225,505,953	80.3	227,295,534	80.3	228,699,523	80.0
Spoke a language other than English at home ¹	54,858,424	19.7	55,444,485	19.7	55,860,545	19.7	57,097,826	20.0
Spoke a language other than English at home	54,858,424	100.0	55,444,485	100.0	55,860,545	100.0	57,097,826	100.0
Spanish or Spanish Creole	34,044,945	62.1	34,547,077	62.3	34,615,394	62.0	35,468,501	62.1
French (includes Patois, Cajun, Creole)	1,997,618	3.6	1,984,824	3.6	1,973,531	3.5	1,964,556	3.4
Italian	828,524	1.5	798,801	1.4	782,173	1.4	753,992	1.3
Portuguese or Portuguese Creole	683,405	1.2	687,126	1.2	661,120	1.2	731,282	1.3
German	1,135,999	2.1	1,104,354	2.0	1,121,465	2.0	1,109,216	1.9
Russian	823,210	1.5	851,174	1.5	860,568	1.5	881,723	1.5
Polish	640,265	1.2	638,059	1.2	616,492	1.1	593,598	1.0
Hindi	504,607	0.9	532,911	1.0	562,587	1.0	560,983	1.0
Chinese	2,492,871	4.5	2,464,572	4.4	2,473,968	4.4	2,600,150	4.6
Korean	1,060,631	1.9	1,062,337	1.9	1,048,400	1.9	1,039,021	1.8
Vietnamese	1,207,721	2.2	1,207,004	2.2	1,236,419	2.2	1,251,468	2.2
Tagalog	1,415,599	2.6	1,480,429	2.7	1,496,208	2.7	1,513,734	2.7
Arabic	732,519	1.3	767,319	1.4	780,995	1.4	845,396	1.5
All other languages	7,290,510	13.3	7,318,498	13.2	7,631,225	13.7	7,784,206	13.6

¹The languages listed in this table are the languages for which there were 500,000 or more speakers in 2009.

Source: U.S. Census Bureau, 2006, 2007, 2008, and 2009 American Community Survey.

Table 1. Languages Spoken at Home: 2006-2009

Characteristic	2006		2007		2008		2009	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Population 5 Years and older	279,012,712	100.0	280,950,438	100.0	283,156,079	100.0	285,797,349	100.0
Spoke only English at home	224,154,288	80.3	225,505,953	80.3	227,295,534	80.3	228,699,523	80.0
Spoke a language other than English at home ¹	54,858,424	19.7	55,444,485	19.7	55,860,545	19.7	57,097,826	20.0
Spoke a language other than English at home	54,858,424	100.0	55,444,485	100.0	55,860,545	100.0	57,097,826	100.0
Spanish or Spanish Creole	34,044,945	62.1	34,547,077	62.3	34,615,394	62.0	35,468,501	62.1
French (includes Patois, Cajun, Creole)	1,997,618	3.6	1,984,824	3.6	1,973,531	3.5	1,964,556	3.4
Italian	828,524	1.5	798,801	1.4	782,173	1.4	753,992	1.3
Portuguese or Portuguese Creole	683,405	1.2	687,126	1.2	661,120	1.2	731,282	1.3
German	1,135,999	2.1	1,104,354	2.0	1,121,465	2.0	1,109,216	1.9
Russian	823,210	1.5	851,174	1.5	860,568	1.5	881,723	1.5
Polish	640,265	1.2	638,059	1.2	616,492	1.1	593,598	1.0
Hindi	504,607	0.9	532,911	1.0	562,587	1.0	560,983	1.0
Chinese	2,492,871	4.5	2,464,572	4.4	2,473,968	4.4	2,600,150	4.6
Korean	1,060,631	1.9	1,062,337	1.9	1,048,400	1.9	1,039,021	1.8
Vietnamese	1,207,721	2.2	1,207,004	2.2	1,236,419	2.2	1,251,468	2.2
Tagalog	1,415,599	2.6	1,480,429	2.7	1,496,208	2.7	1,513,734	2.7
Arabic	732,519	1.3	767,319	1.4	780,995	1.4	845,396	1.5
All other languages	7,290,510	13.3	7,318,498	13.2	7,631,225	13.7	7,784,206	13.6

¹The languages listed in this table are the languages for which there were 500,000 or more speakers in 2009.

Source: U.S. Census Bureau, 2006, 2007, 2008, and 2009 American Community Survey.

National Projections

- The U.S. Census Bureau's official national population projections released in August 2008 and a series of supplemental projections released in December 2009.
- The projections produced in 2009 consist of four series based on alternative net international migration (NIM) assumptions and include:
 - (1) High Net International Migration
 - (2) Low Net International Migration
 - (3) Constant Net International Migration
 - (4) Zero Net International Migration

National Projections (continued)

- Based on Census 2000 and produced using a cohort-component method.
 - We do not use any data from the 2010 Census.
- Projections of the resident population and demographic components of change by age, sex, race, and Hispanic origin for the United States: 2000-2050.
 - We use the projected populations for years 2010-2020.
- Assumptions about future rates of mortality and fertility are the same in all five series.

Language Projection Methodology

- Projections of the total number of speakers of a language other than English (LOTE) and the number of speakers for individual languages with at least 500,000 speakers in 2009.
- Future LOTE use is projected for the years 2010 through 2020 and is based on trends in ACS data.
- The projected number of LOTE speakers is produced by applying the projected percentages of speakers to the projected population.

Methodology (continued)

- Projections of the percent speaking a LOTE are developed and applied by demographic characteristics:
 - LOTE and Spanish use are projected by age and Hispanic origin – 92 groups.
 - The other individual languages are projected by age – 46 groups.

Three Projection Models

Constant

Linear

Logistic

Constant Model

- LOTE use is held constant at the level reported in the 2009 ACS.
- Assumes no change in the distribution of LOTE speakers.

$$P_{2009} = P_{2010} = P_{2011} = \dots = P_{2020}$$

where:

P = Percent speaking a LOTE in a given year

Linear Model

- LOTE use is projected based on trends observed in the 2006 through 2009 ACS.
- Assumes that LOTE use in the future will change by the same amount as in the past.

$$P_t = a + b (t)$$

where:

P_t = Percent speaking LOTE at time t

a = estimated intercept

b = estimated slope

t = time (year)

Logistic Model

- LOTE use is projected based on trends observed in the 2006 through 2009 ACS.
- Assumes that changes in LOTE use are constrained.

$$P_t = \frac{a}{[1 + (b)(e^{-ct})]}$$

where:

P_t = Percent speaking LOTE at time t

a, b, c = estimated parameters

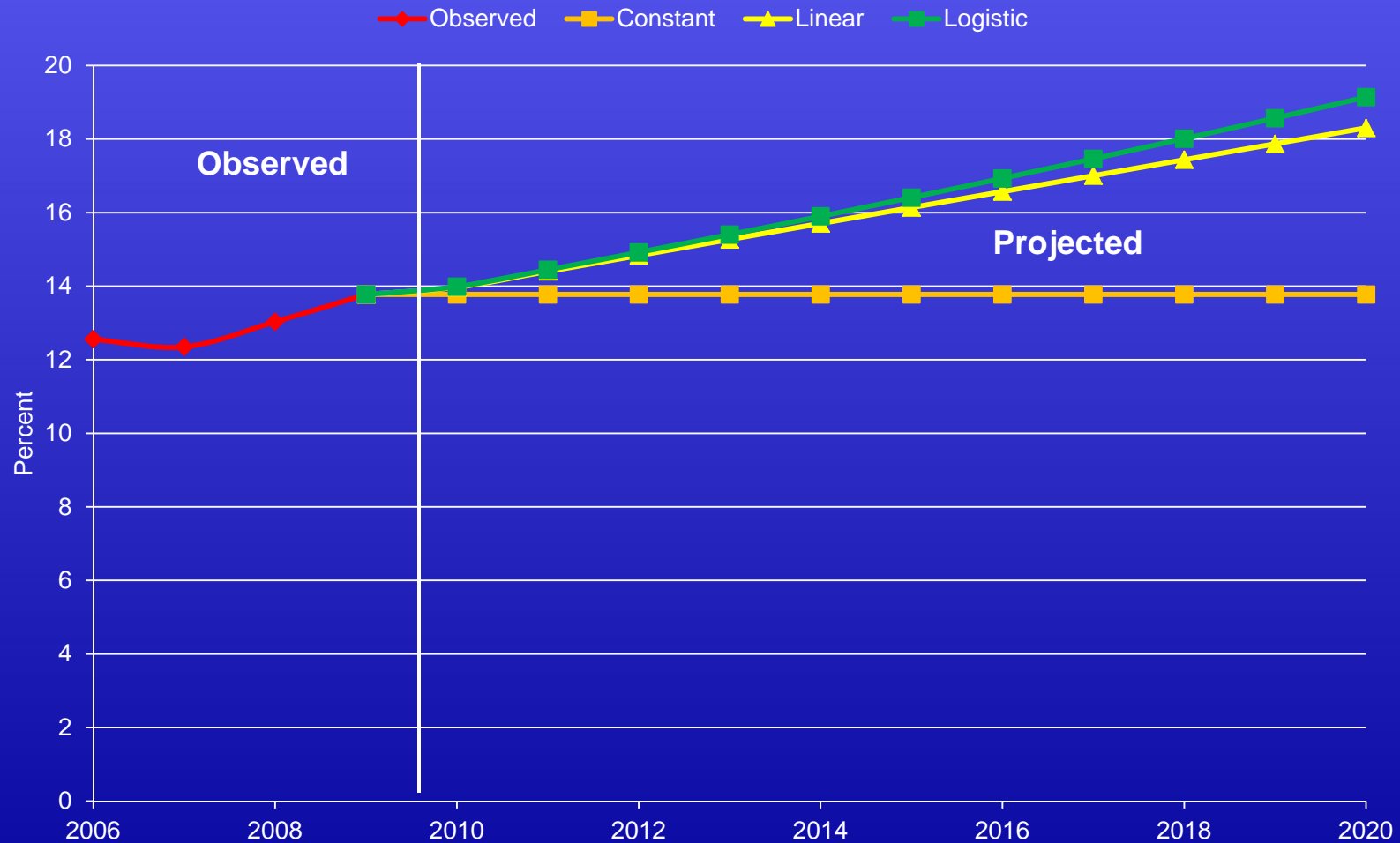
t = time (year)

Drivers of Change in LOTE Use

- Constant Model:
 - Changes in the number of speakers will be driven by the population projections.
- Linear and Logistic Model:
 - Changes in the number of speakers will be driven by changes in the projected distribution of LOTE speakers and by the population projections.

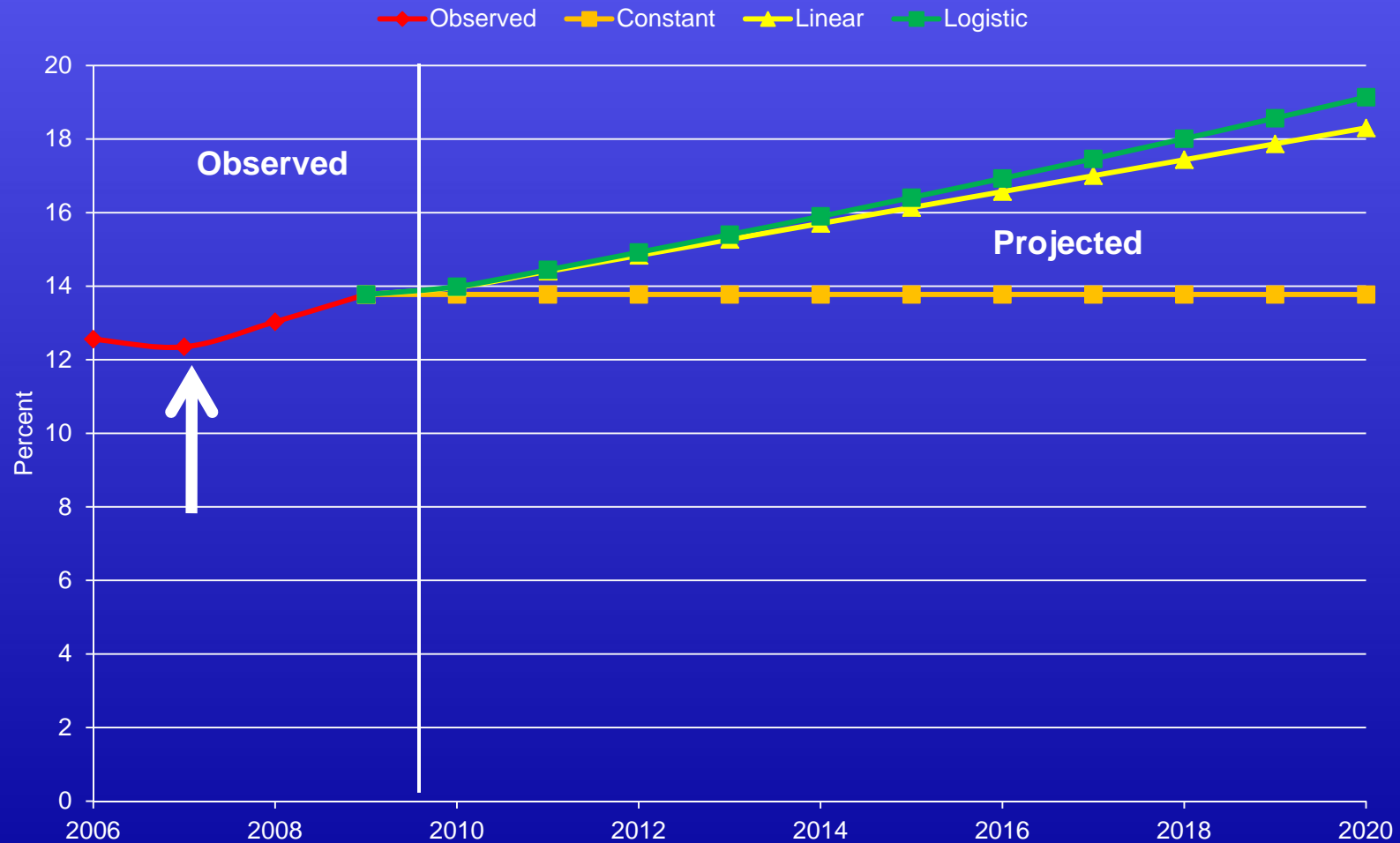
Comparison of Models: Example 1

Increase in LOTE use from 2006 to 2009



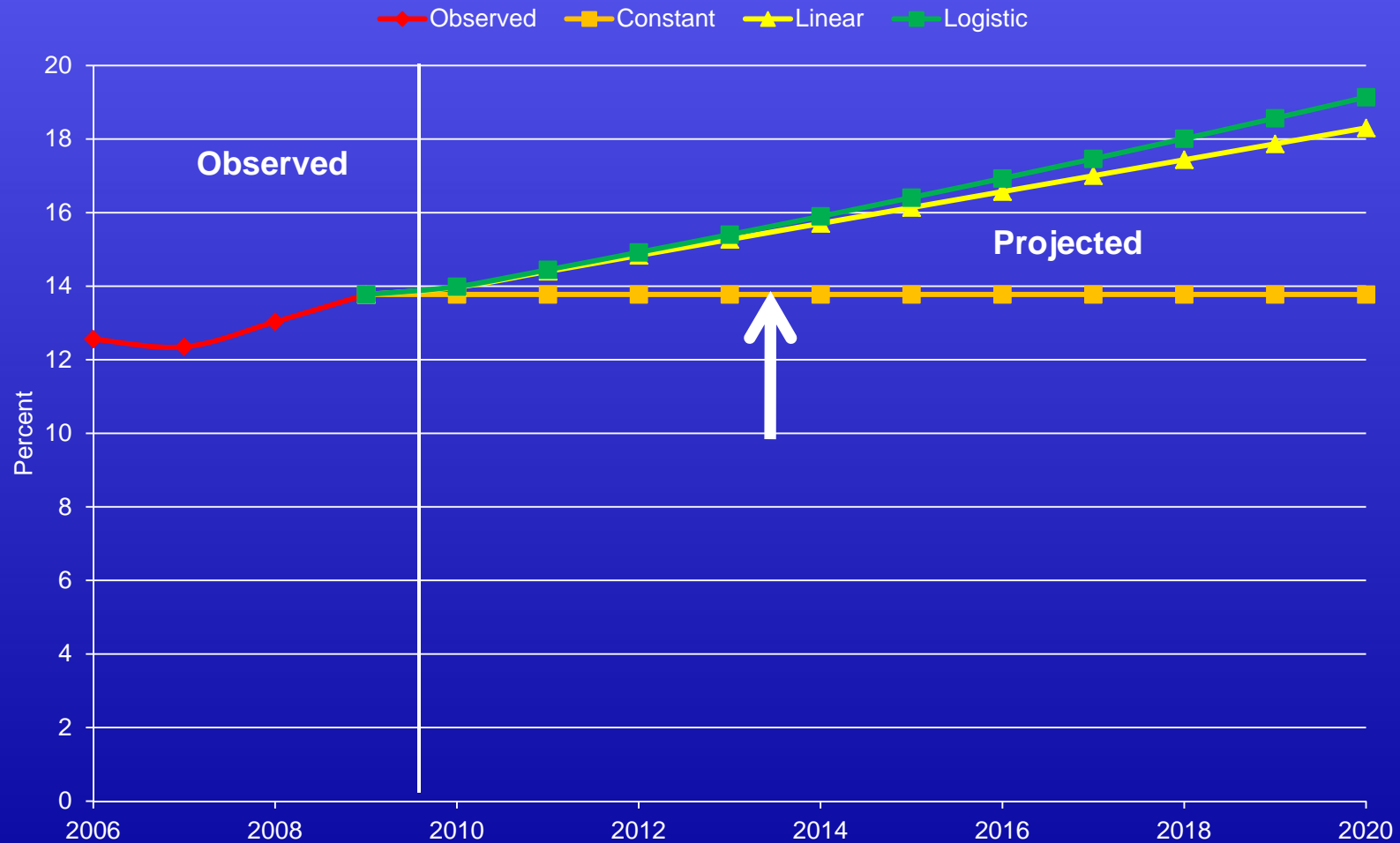
Comparison of Models: Example 1

Increase in LOTE use from 2006 to 2009



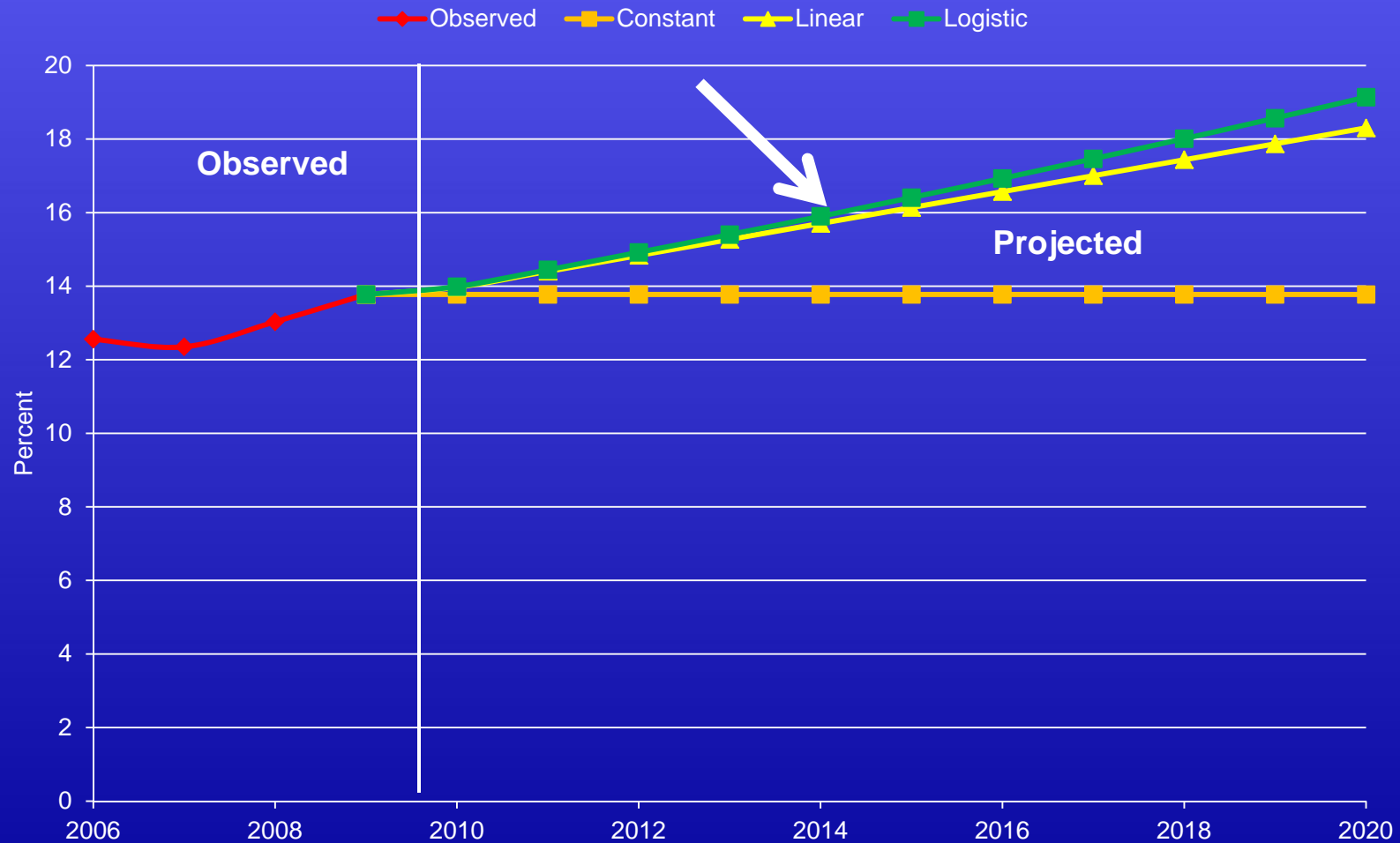
Comparison of Models: Example 1

Increase in LOTE use from 2006 to 2009



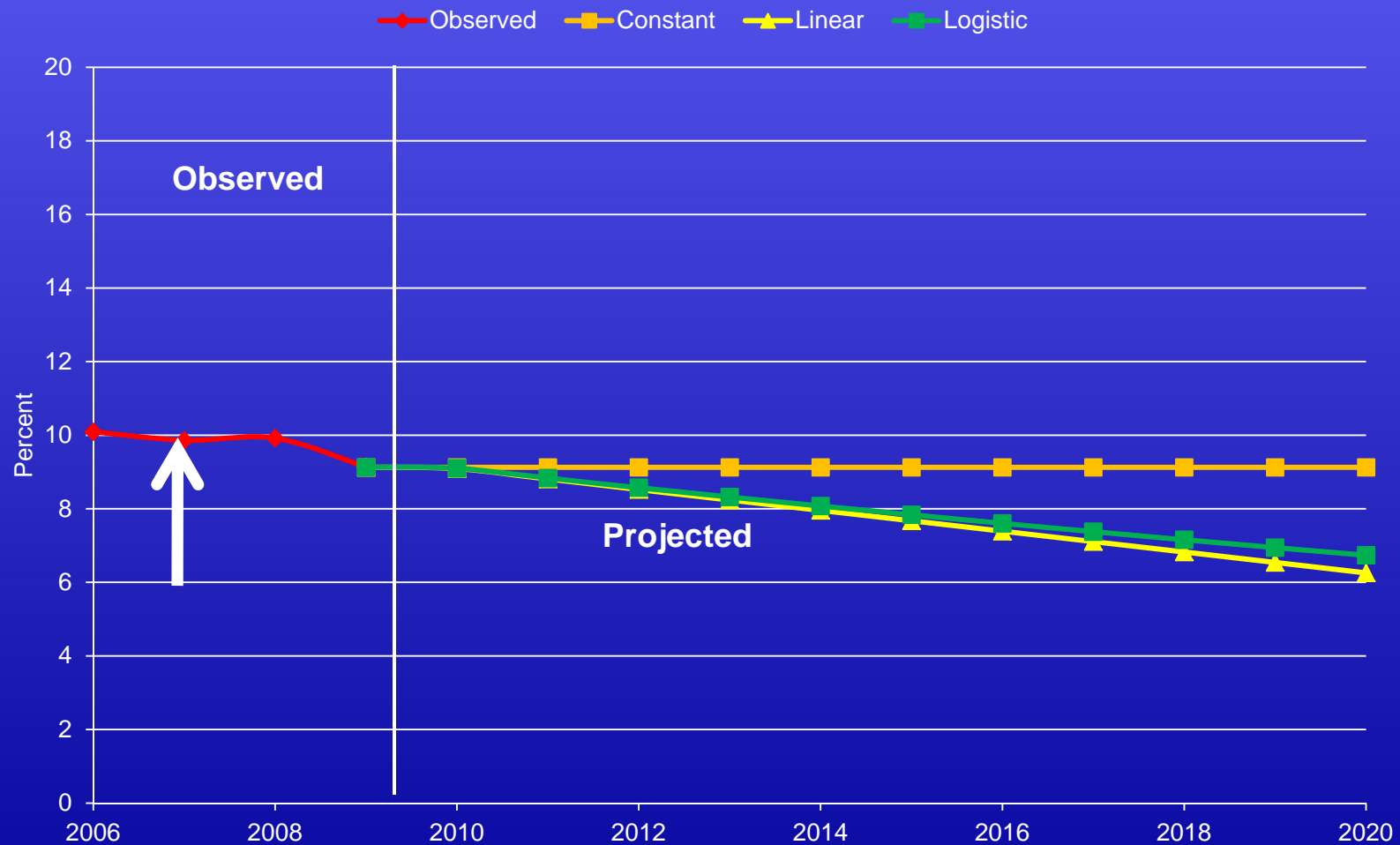
Comparison of Models: Example 1

Increase in LOTE use from 2006 to 2009



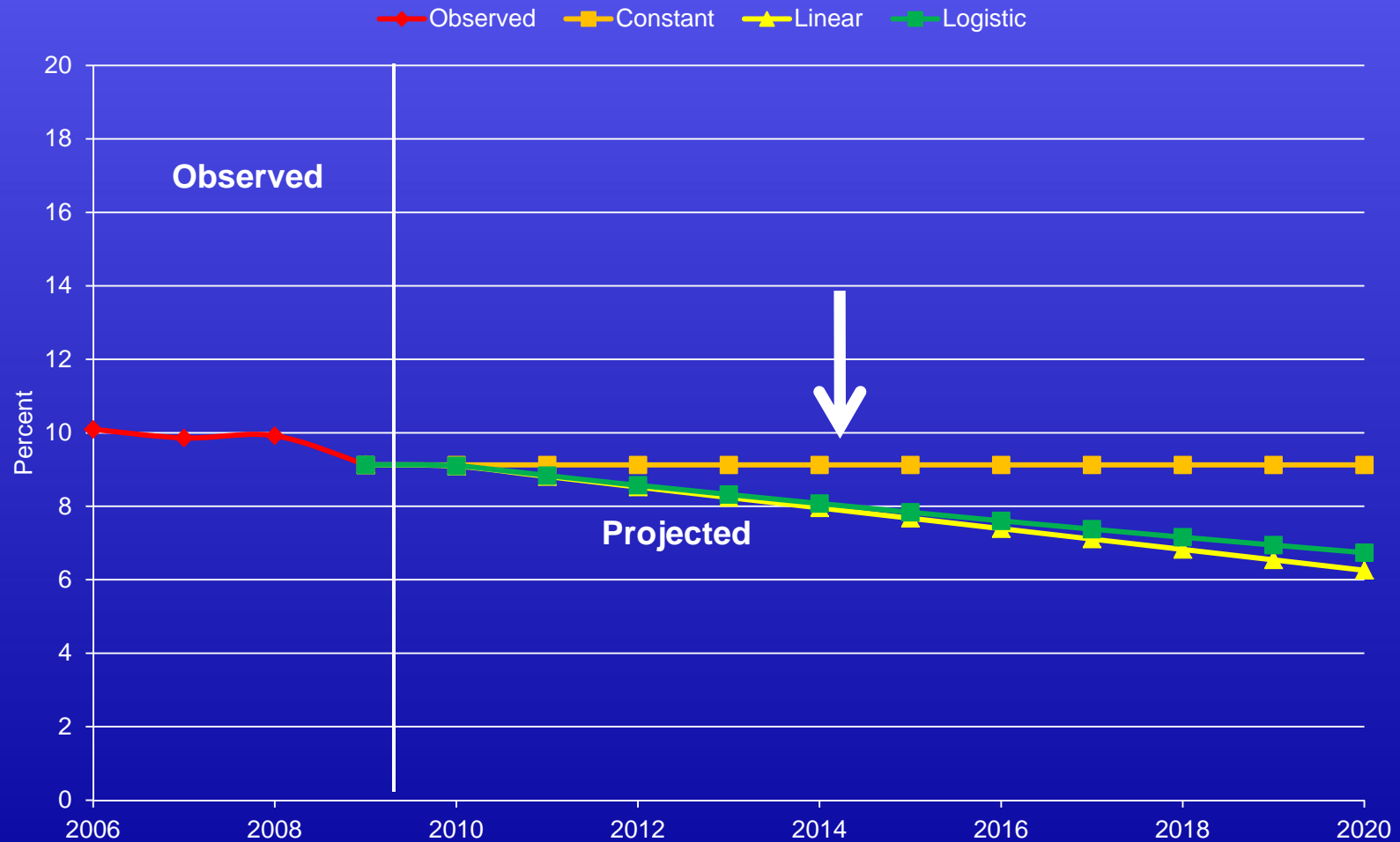
Comparison of Models: Example 2

Decrease in LOTE use from 2006 to 2009



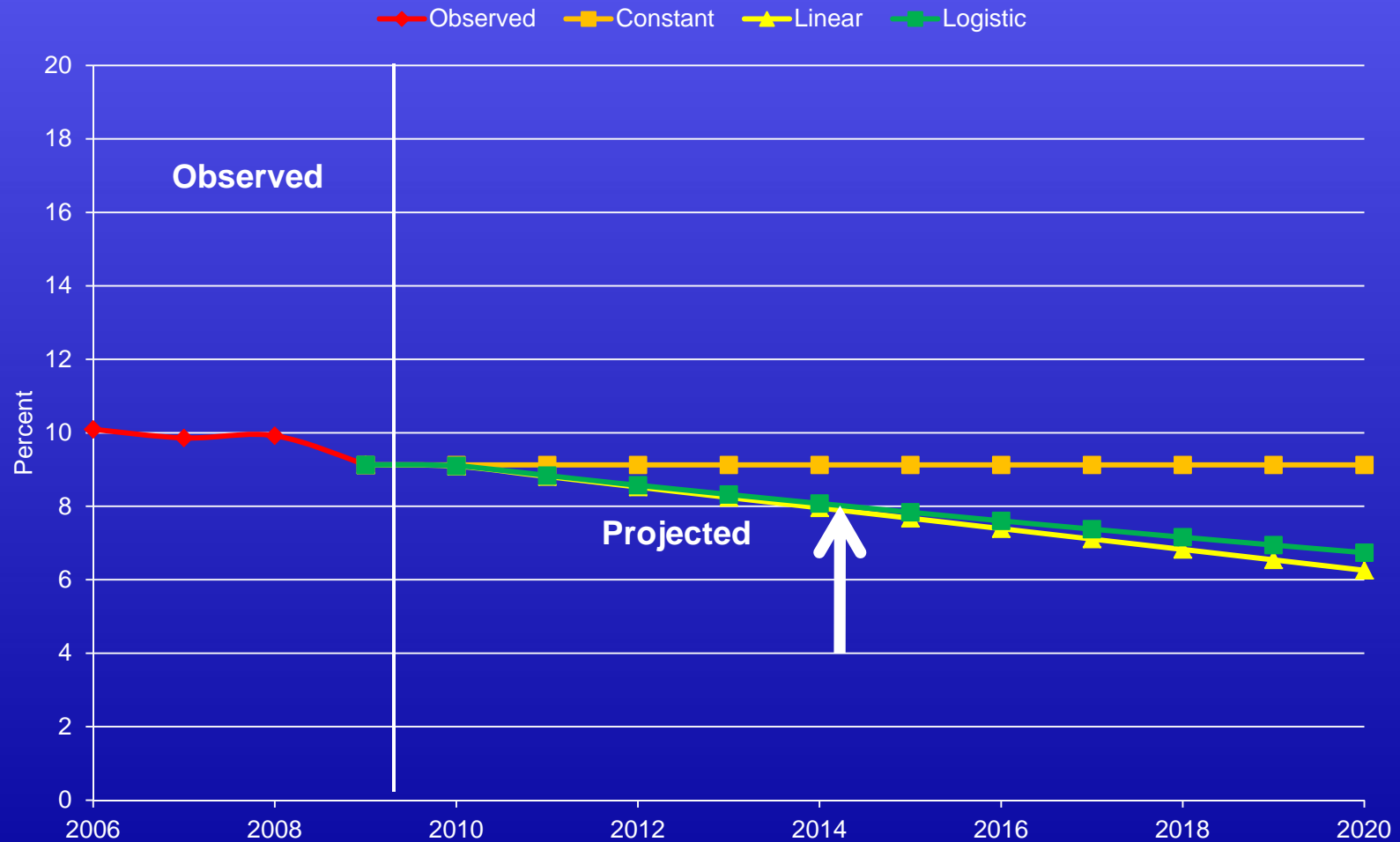
Comparison of Models: Example 2

Decrease in LOTE use from 2006 to 2009



Comparison of Models: Example 2

Decrease in LOTE use from 2006 to 2009



Results

Projected Number of LOTE Speakers

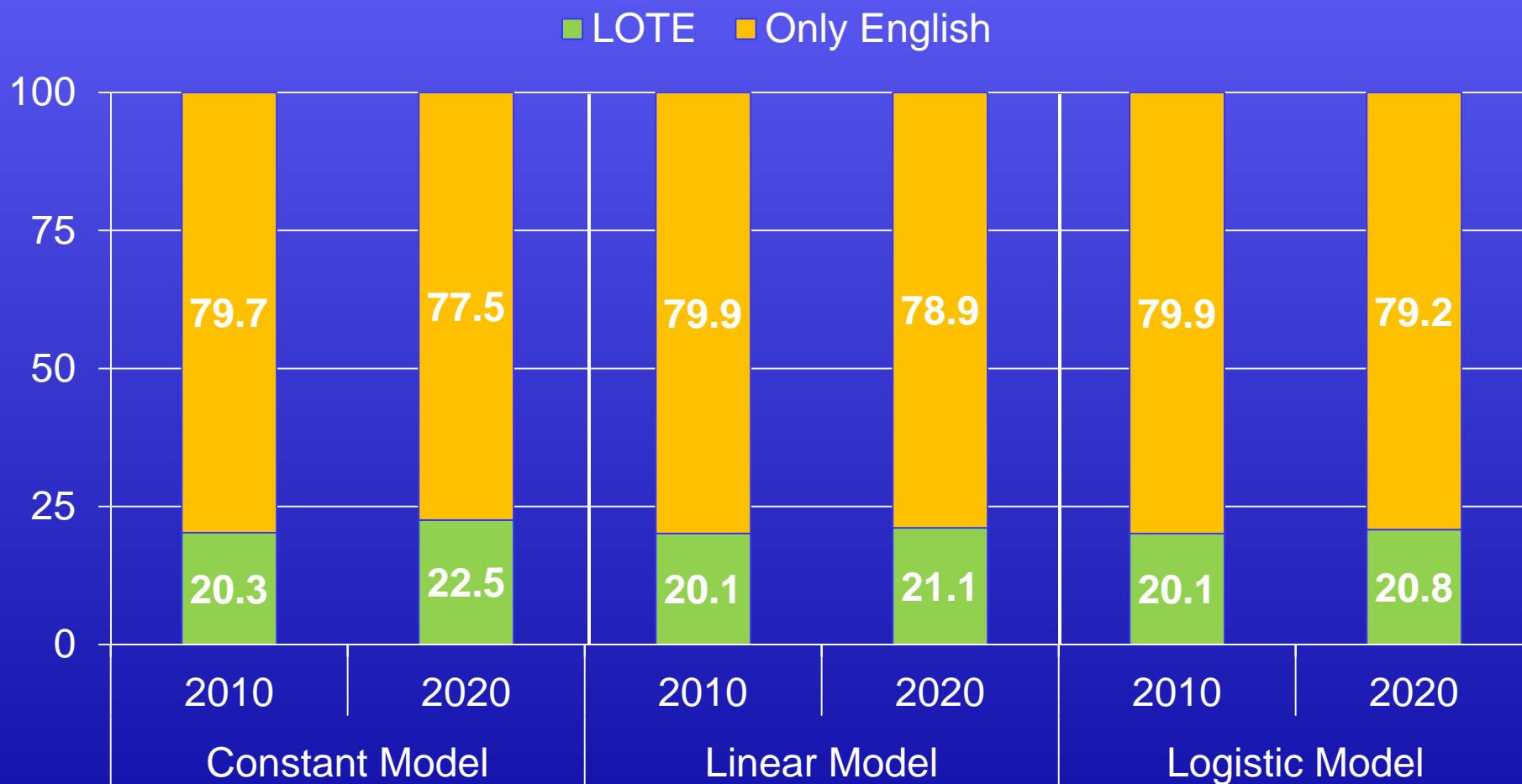


Numbers in thousands
Based on the 2008 National Population Projections

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

Projected Percent Speaking LOTE and English Only

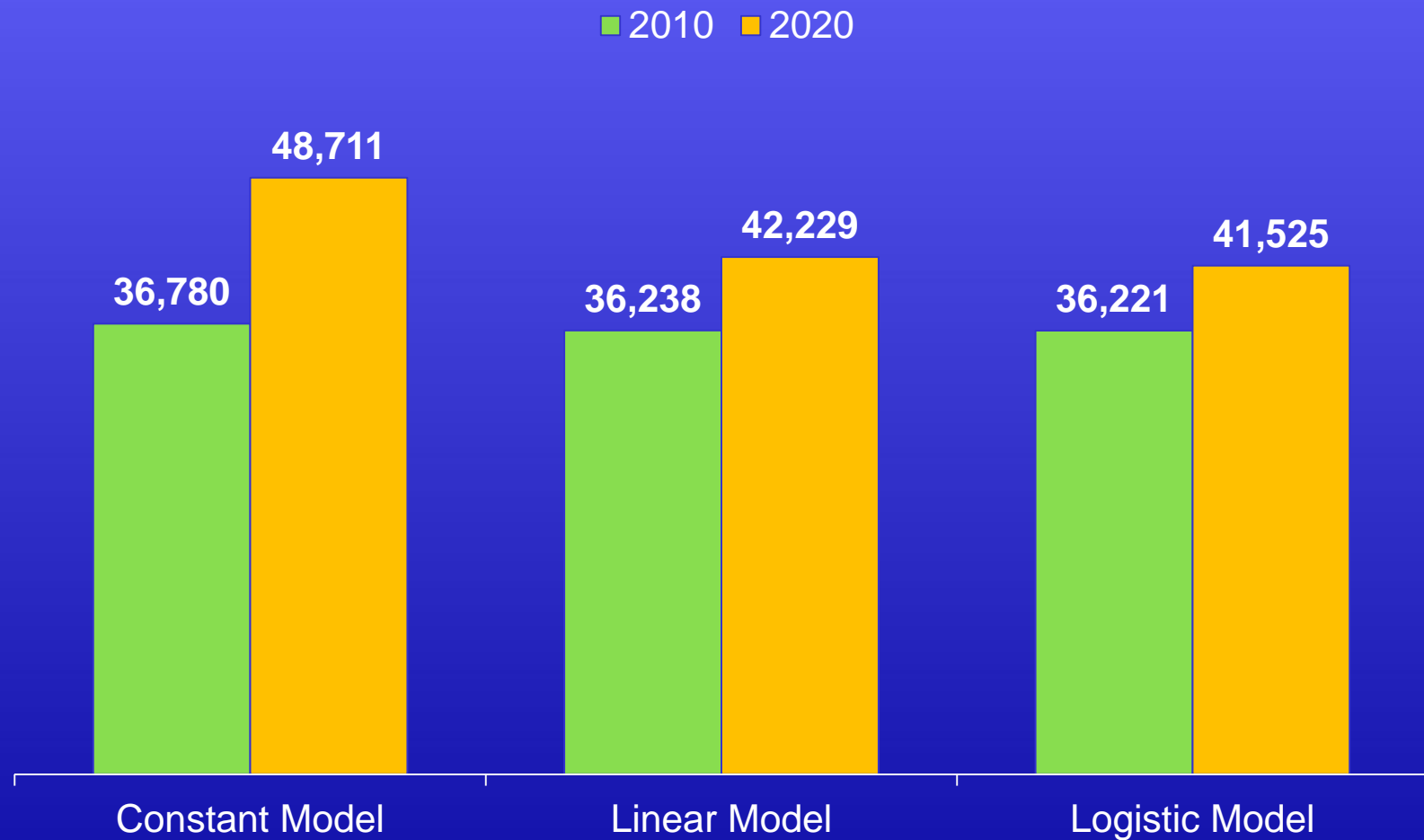


Based on the 2008 National Population Projections

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

Projected Number of Spanish Speakers



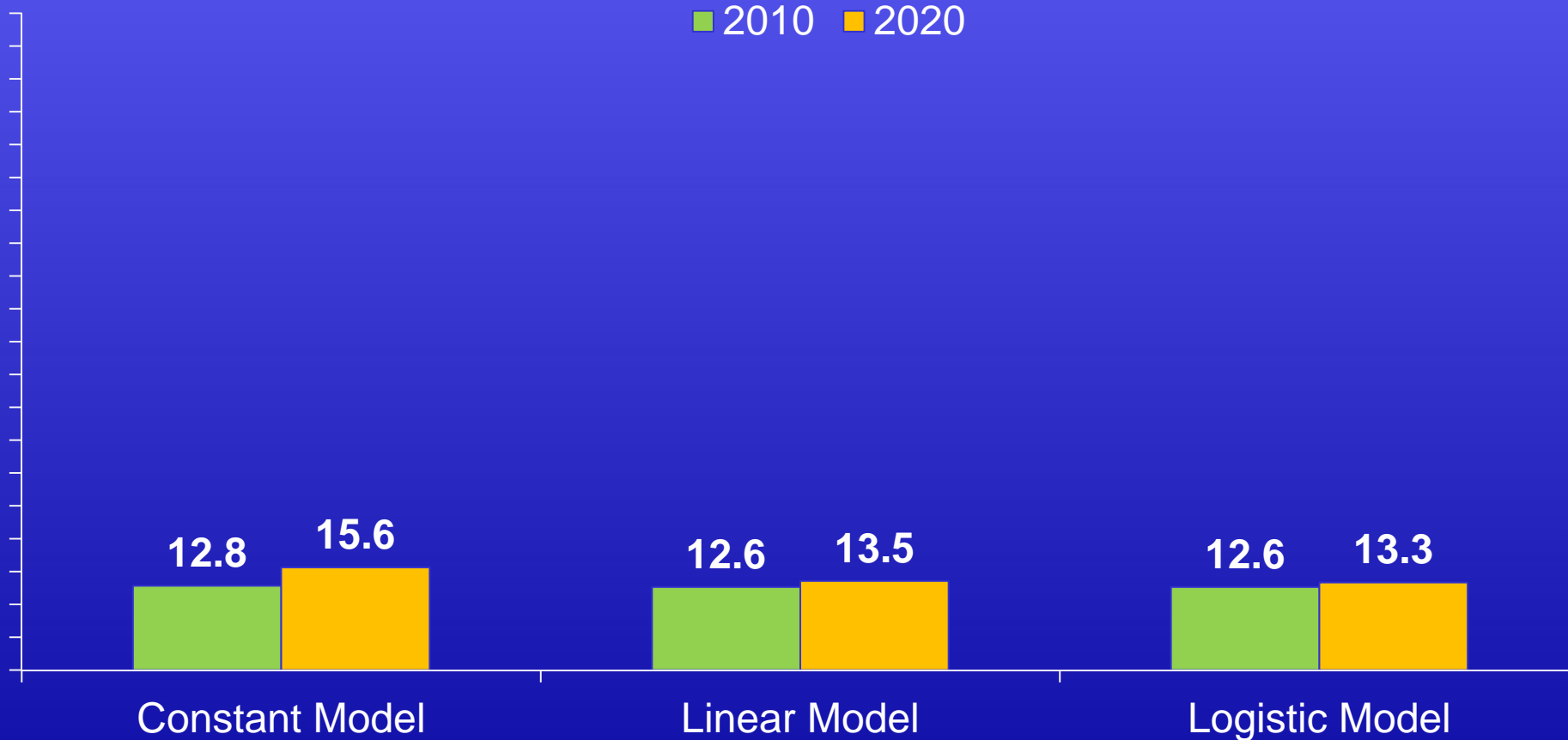
Numbers in thousands
Based on the 2008 National Population Projections

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

Projected Percent of the Population 5 Years and Older Speaking Spanish

■ 2010 ■ 2020



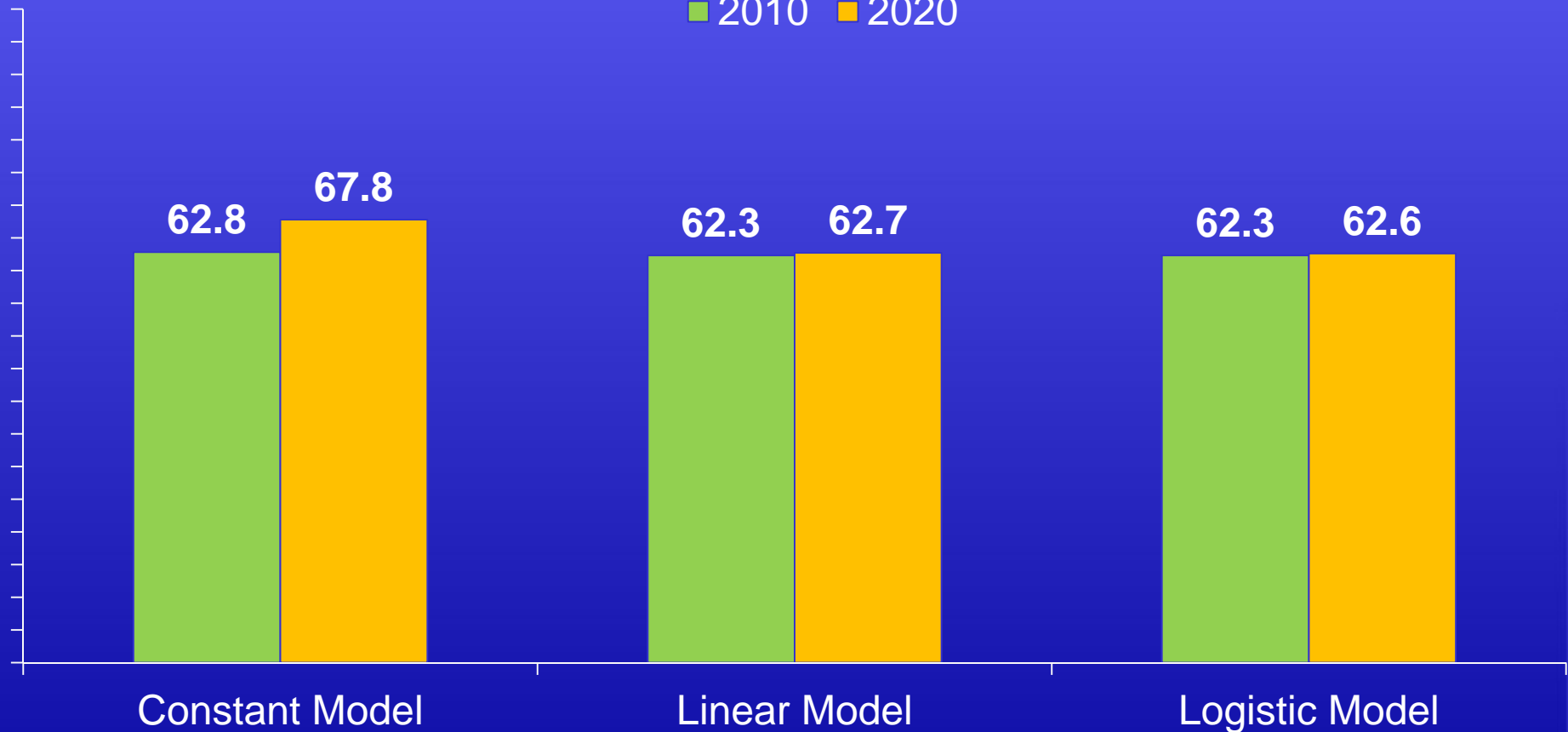
Based on the 2008 National Population Projections

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

Projected Percent of the LOTE Speakers that Speak Spanish

■ 2010 ■ 2020

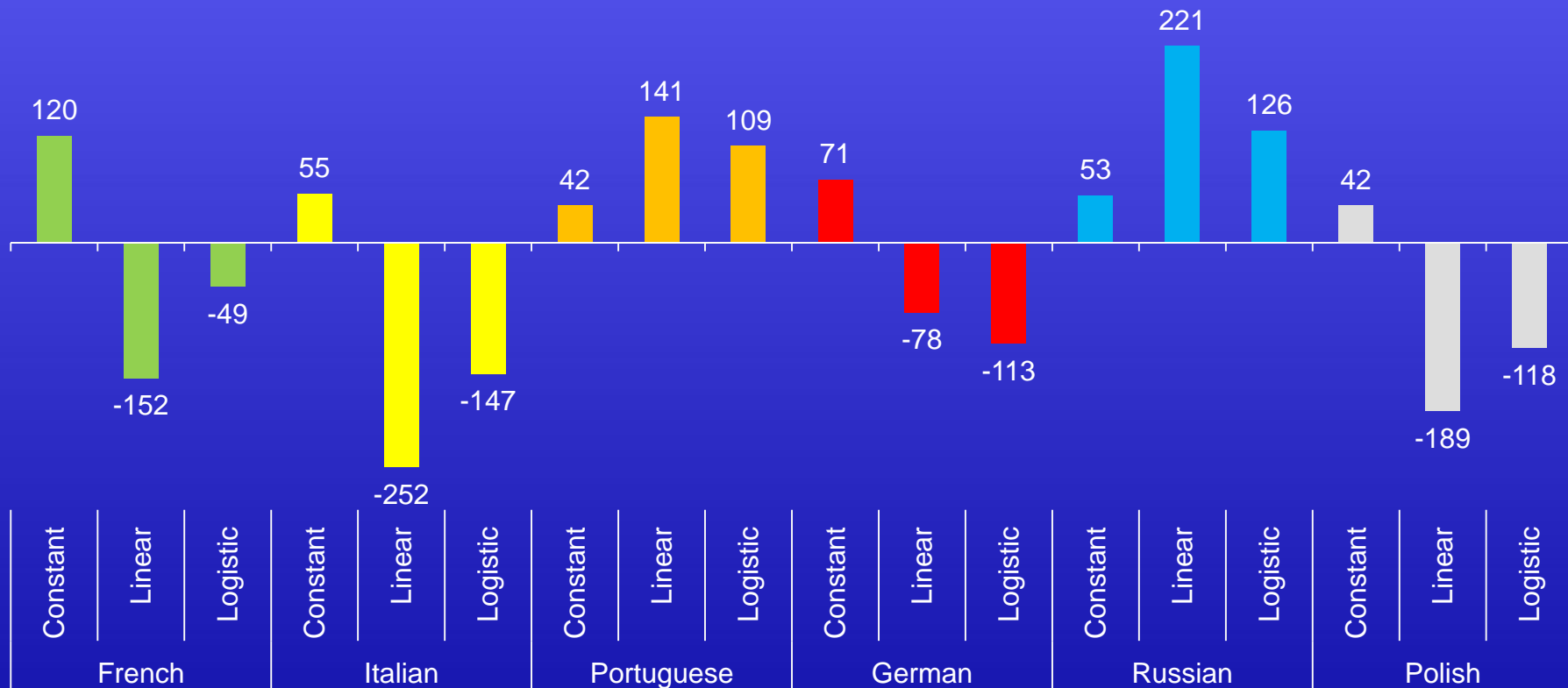


Based on the 2008 National Population Projections

U S C E N S U S B U R E A U

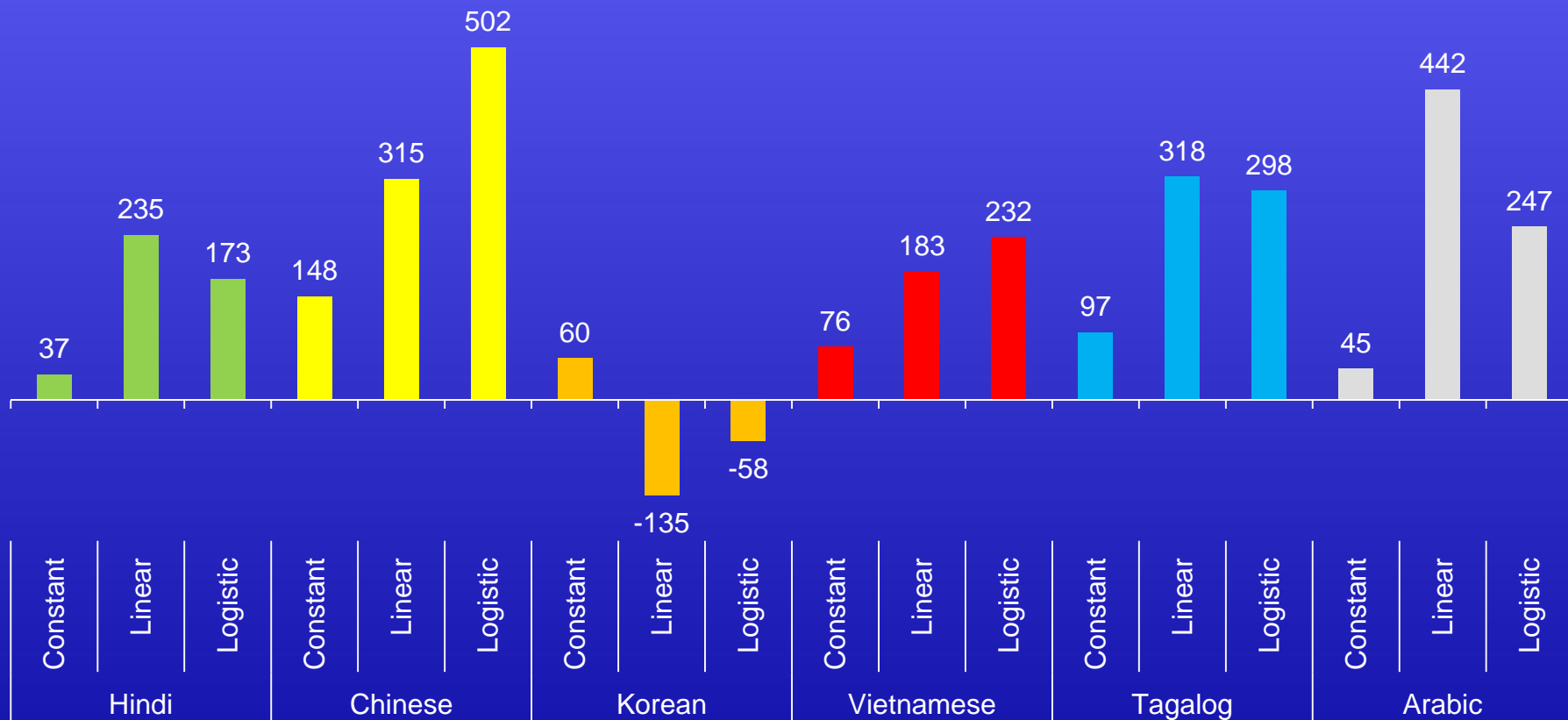
Helping You Make Informed Decisions

Change between 2010 and 2020 in the Projected Number Speaking European Languages



Numbers in thousands
Based on the 2008 National Population Projections

Change between 2010 and 2020 in the Projected Number Speaking Non-European Languages



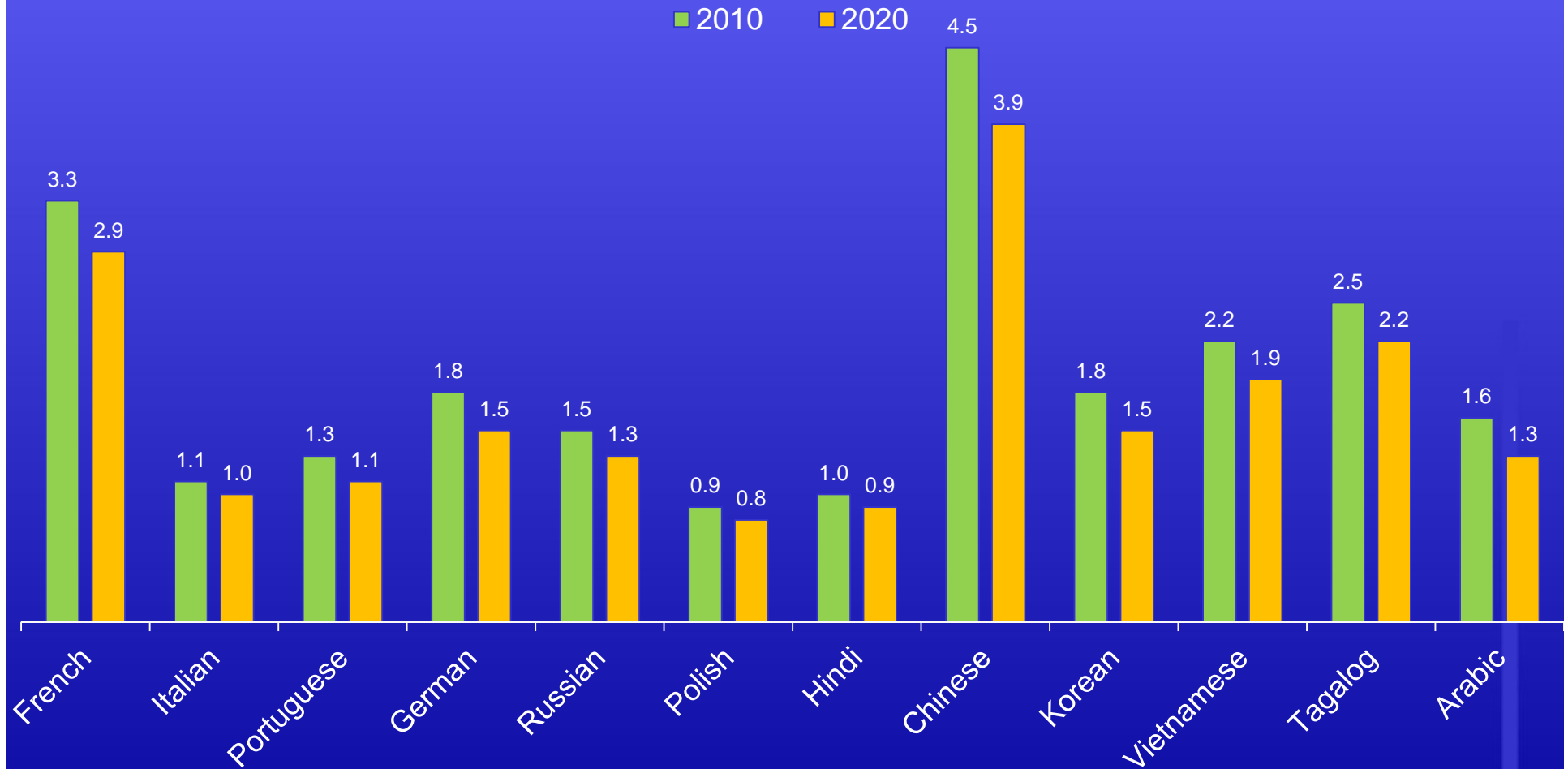
Numbers in thousands

Based on the 2008 National Population Projections

USCENSUS BUREAU

Helping You Make Informed Decisions

Distribution of LOTE Speakers by Language: Constant Model

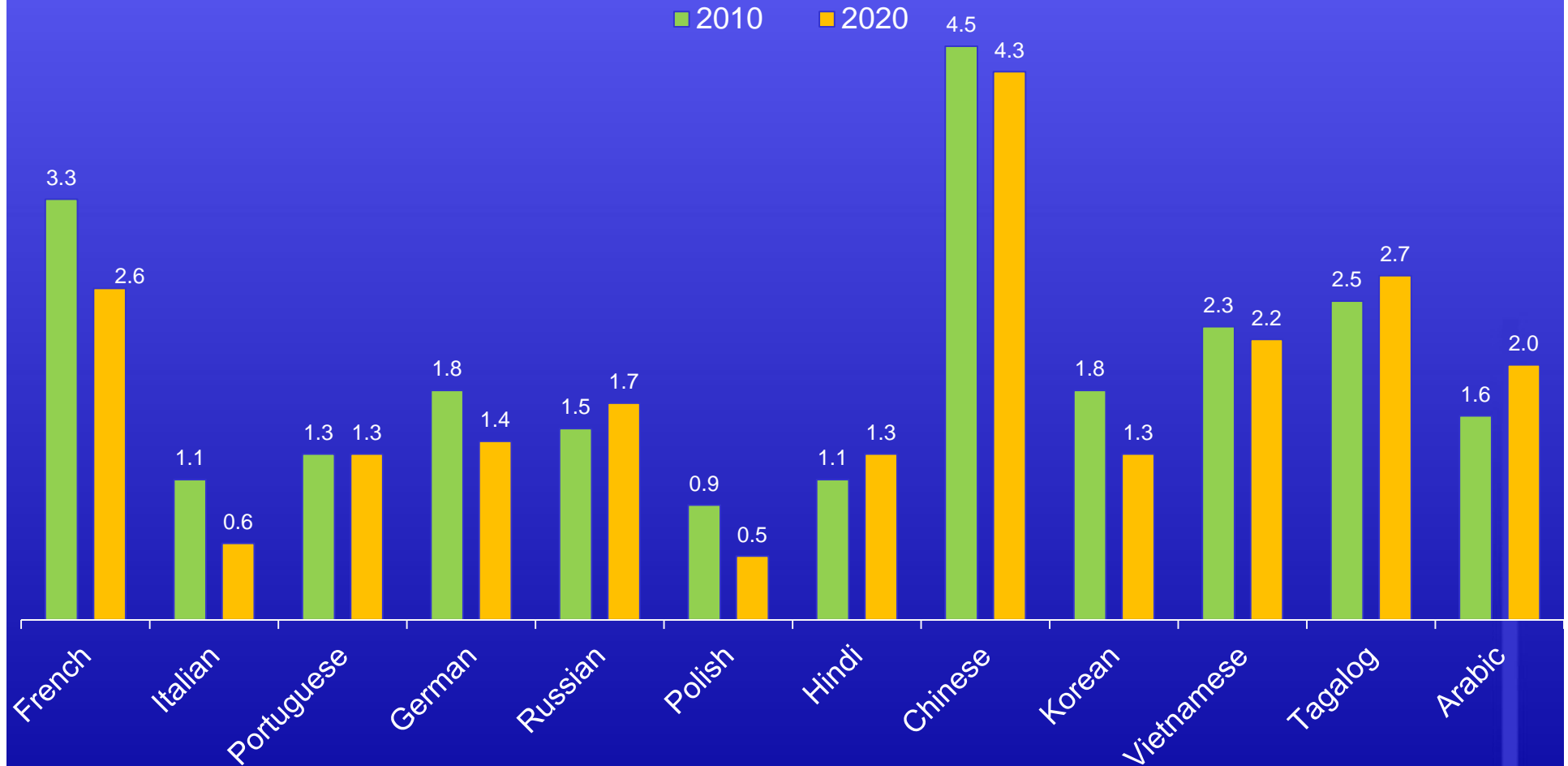


Percent of all LOTE Speakers

USCENSUS BUREAU

Helping You Make Informed Decisions

Distribution of LOTE Speakers by Language: Linear Model

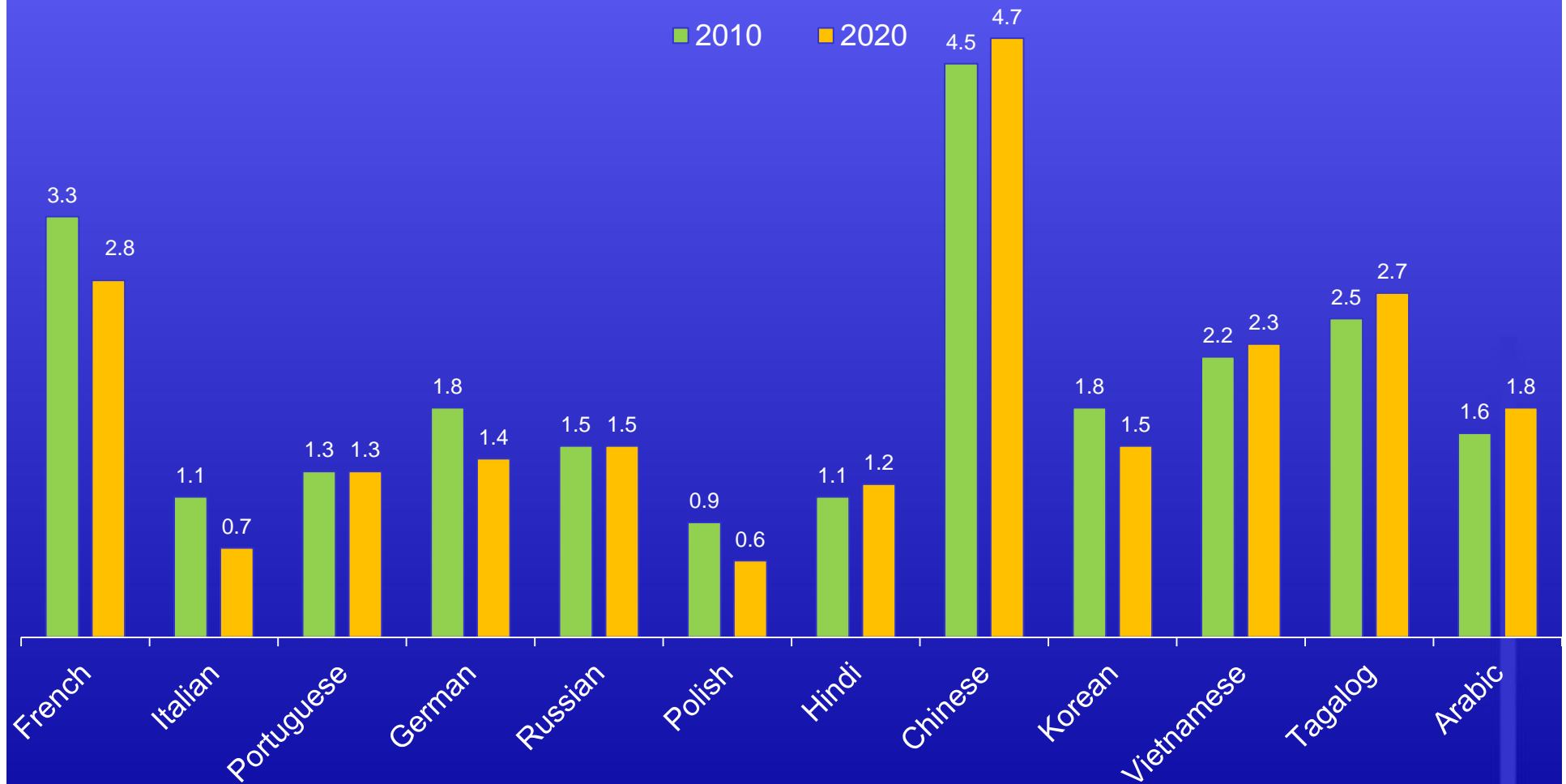


Percent of all LOTE Speakers

USCENSUSBUREAU

Helping You Make Informed Decisions

Distribution of LOTE Speakers by Language: Logistic Model



Percent of all LOTE Speakers

USCENSUS BUREAU

Helping You Make Informed Decisions

Summary

- The constant growth assumption is simplistic, resulting in an increase in LOTE use for all languages.
- The linear and logistic assumptions are more realistic and provide results that are very similar.
- The use of LOTE is projected to increase over the next ten years.
 - Spanish is projected to remain the most frequently spoken.
 - Portuguese, Russian, Hindi, Chinese, Vietnamese, Tagalog, and Arabic are all projected to increase.
 - French, Italian, German, Polish, and Korean are projected to decline.

Future Directions

- Include 2010 ACS data and use 2010-Census based population projections when the data are available
- Project by age groups instead of single years
- Use 3-year ACS files as basis of time series
 - 2006-2008
 - 2007-2009
 - 2008-2010
- Project by nativity (foreign and native born)
- Project by birth cohorts instead of age

More Information

- <http://www.census.gov/acs/www/>
- <http://www.census.gov/population/www/projections>
- Questions:
 - Population Projections Branch:
301-763-2428
jennifer.m.ortman@census.gov
 - Education and Social Stratification Branch:
301-763-2464
hyon.b.shin@census.gov