Immigrant Locations and Native Residential Preferences in Spain: New Ghettos?

Jesús Fernández-Huertas Moraga[1]
Ada Ferrer[2]
Albert Saiz[3]

[1]FEDEA, IAE (CSIC) and IZA
[2]IAE (CSIC)
[3]MIT

UAB Seminar
Bellaterra, 6^{th} November 2013

International Migration and Residential Segregation

Fact: International migration from poor to rich countries has been increasing during the past 20 years.

Questions: How do natives' residential location decisions respond to immigrant arrivals? Do natives contribute to immigrant residential segregation?

International Migration and Residential Segregation

Fact: International migration from poor to rich countries has been increasing during the past 20 years.

Questions: How do natives' residential location decisions respond to immigrant arrivals? Do natives contribute to immigrant residential segregation?

The Spanish Experience

Fact 1: Spain received an enormous inflow of international immigrants in 1998-2008: from 3 per cent to 13 per cent of the Spanish population.

Fact 2: This increase was part of the general tendency but faster and larger than in any other rich country in the last 20 years.

Questions: How did Spanish natives' residential location decisions respond to immigrant arrivals? Did Spanish natives contribute to immigrant residential segregation?

The Spanish Experience

Fact 1: Spain received an enormous inflow of international immigrants in 1998-2008: from 3 per cent to 13 per cent of the Spanish population.

Fact 2: This increase was part of the general tendency but faster and larger than in any other rich country in the last 20 years.

Questions: How did Spanish natives' residential location decisions respond to immigrant arrivals? Did Spanish natives contribute to immigrant residential segregation?

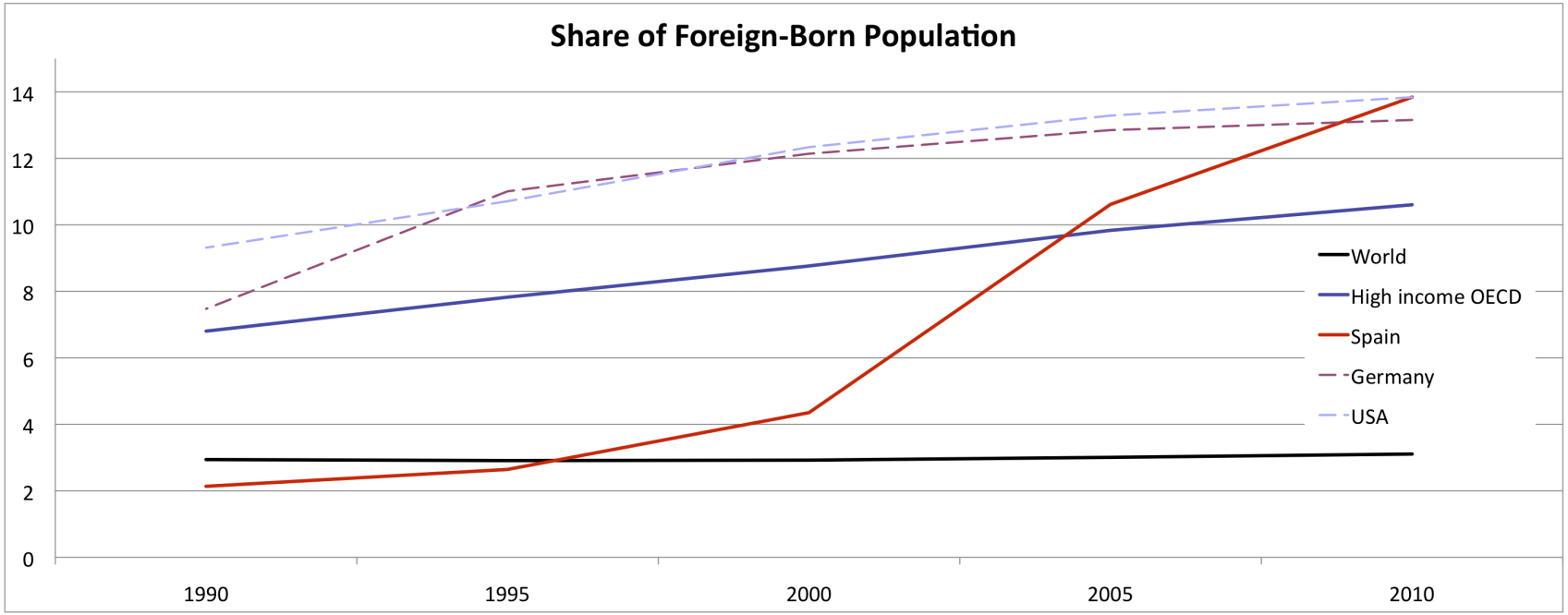
The Spanish Experience

Fact 1: Spain received an enormous inflow of international immigrants in 1998-2008: from 3 per cent to 13 per cent of the Spanish population.

Fact 2: This increase was part of the general tendency but faster and larger than in any other rich country in the last 20 years.

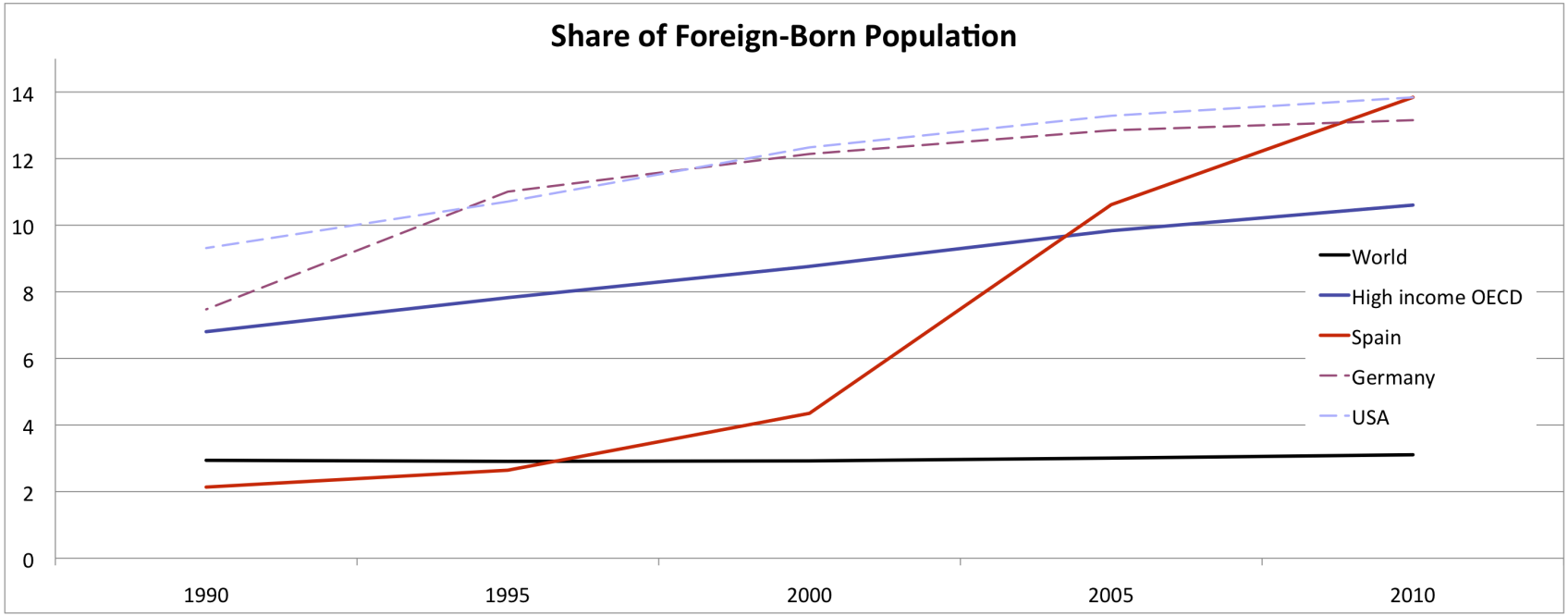
Questions: How did Spanish natives' residential location decisions respond to immigrant arrivals? Did Spanish natives contribute to immigrant residential segregation?

Comparing Spain with other destinations



Spain received 5.5 million immigrants in 1990-2010, second only to the US: 19.6 million. Germany is third: 4.8 million.

Comparing Spain with other destinations



Spain received 5.5 million immigrants in 1990-2010, second only to the US: 19.6 million. Germany is third: 4.8 million.

What the paper does

Comprehensive study of the response of Spanish natives' residence location decisions to immigrant arrivals, combining microdata on exact addresses with distance to amenities and socioeconomic characteristics of neighborhoods at baseline.

In an environment of massive inflows, with the Spanish population growing by 10 percent because of immigration between 1998 and 2008, we find that:

- Immigrants displaced natives from city centers and centers of satellite towns in metro areas.
- New neighborhoods in suburbs saw both immigrant and native arrivals.
- Overall effect on average immigrant segregation neutral.

What the paper does

Comprehensive study of the response of Spanish natives' residence location decisions to immigrant arrivals, combining microdata on exact addresses with distance to amenities and socioeconomic characteristics of neighborhoods at baseline.

In an environment of massive inflows, with the Spanish population growing by 10 percent because of immigration between 1998 and 2008, we find that:

- Immigrants displaced natives from city centers and centers of satellite towns in metro areas.
- New neighborhoods in suburbs saw both immigrant and native arrivals.
- Overall effect on average immigrant segregation neutral.

What the paper does

Comprehensive study of the response of Spanish natives' residence location decisions to immigrant arrivals, combining microdata on exact addresses with distance to amenities and socioeconomic characteristics of neighborhoods at baseline.

In an environment of massive inflows, with the Spanish population growing by 10 percent because of immigration between 1998 and 2008, we find that:

- Immigrants displaced natives from city centers and centers of satellite towns in metro areas.
- New neighborhoods in suburbs saw both immigrant and native arrivals.
- Overall effect on average immigrant segregation neutral.

What the paper does

Comprehensive study of the response of Spanish natives' residence location decisions to immigrant arrivals, combining microdata on exact addresses with distance to amenities and socioeconomic characteristics of neighborhoods at baseline.

In an environment of massive inflows, with the Spanish population growing by 10 percent because of immigration between 1998 and 2008, we find that:

- Immigrants displaced natives from city centers and centers of satellite towns in metro areas.
- New neighborhoods in suburbs saw both immigrant and native arrivals.
- Overall effect on average immigrant segregation neutral.

What the paper does

Comprehensive study of the response of Spanish natives' residence location decisions to immigrant arrivals, combining microdata on exact addresses with distance to amenities and socioeconomic characteristics of neighborhoods at baseline.

In an environment of massive inflows, with the Spanish population growing by 10 percent because of immigration between 1998 and 2008, we find that:

- Immigrants displaced natives from city centers and centers of satellite towns in metro areas.
- New neighborhoods in suburbs saw both immigrant and native arrivals.
- Overall effect on average immigrant segregation neutral.

Previous literature

- US-centered literature. It uses 10-year aggregated data. Examples: Cutler, Glaeser and Vigdor (2008); Saiz and Wachter (2011).
- Scandinavian datasets. Inflows are much smaller than in the Spanish case. Examples: Edlin, Fredriksson and Aslund (2003); Piil Damm (2009); Dahlberg, Edmark and Lundqvist (2012).
- Spain. Studies with limited scope. Examples: Pareja-Eastaway (2009) focuses on just one city; Bosch, Carnero and Farré (2010) show the existence of ethnic discrimination in the rental market; Ballester and Vorsatz (2013) focus on a cross-section when introducing a new measure of segregation.

Previous literature

- US-centered literature. It uses 10-year aggregated data. Examples: Cutler, Glaeser and Vigdor (2008); Saiz and Wachter (2011).
- Scandinavian datasets. Inflows are much smaller than in the Spanish case. Examples: Edlin, Fredriksson and Aslund (2003); Piil Damm (2009); Dahlberg, Edmark and Lundqvist (2012).
- Spain. Studies with limited scope. Examples: Pareja-Eastaway (2009) focuses on just one city; Bosch, Carnero and Farré (2010) show the existence of ethnic discrimination in the rental market; Ballester and Vorsatz (2013) focus on a cross-section when introducing a new measure of segregation.

Previous literature

- US-centered literature. It uses 10-year aggregated data. Examples: Cutler, Glaeser and Vigdor (2008); Saiz and Wachter (2011).
- Scandinavian datasets. Inflows are much smaller than in the Spanish case. Examples: Edlin, Fredriksson and Aslund (2003); Piil Damm (2009); Dahlberg, Edmark and Lundqvist (2012).
- Spain. Studies with limited scope. Examples: Pareja-Eastaway (2009) focuses on just one city; Bosch, Carnero and Farré (2010) show the existence of ethnic discrimination in the rental market; Ballester and Vorsatz (2013) focus on a cross-section when introducing a new measure of segregation.

The Data

- Microdata from the Spanish Municipal Registry (Padrón): population registered in Spanish municipalities as of January 1st yearly from 1998 till 2008.
- Registration gives access to municipal and regional services. For example, schooling and health.
- Undocumented migrants were allowed to register (since January 2000) and registration was used to legalize during amnesties.
- Person characteristics: street address, place of birth, date of birth, nationality, gender, education (unreliable).

The Data

- Microdata from the Spanish Municipal Registry (Padrón): population registered in Spanish municipalities as of January 1st yearly from 1998 till 2008.
- Registration gives access to municipal and regional services. For example, schooling and health.
- Undocumented migrants were allowed to register (since January 2000) and registration was used to legalize during amnesties.
- Person characteristics: street address, place of birth, date of birth, nationality, gender, education (unreliable).

The Data

- Microdata from the Spanish Municipal Registry (Padrón): population registered in Spanish municipalities as of January 1st yearly from 1998 till 2008.
- Registration gives access to municipal and regional services. For example, schooling and health.
- Undocumented migrants were allowed to register (since January 2000) and registration was used to legalize during amnesties.
- Person characteristics: street address, place of birth, date of birth, nationality, gender, education (unreliable).

The Data

- Microdata from the Spanish Municipal Registry (Padrón): population registered in Spanish municipalities as of January 1st yearly from 1998 till 2008.
- Registration gives access to municipal and regional services. For example, schooling and health.
- Undocumented migrants were allowed to register (since January 2000) and registration was used to legalize during amnesties.
- Person characteristics: street address, place of birth, date of birth, nationality, gender, education (unreliable).

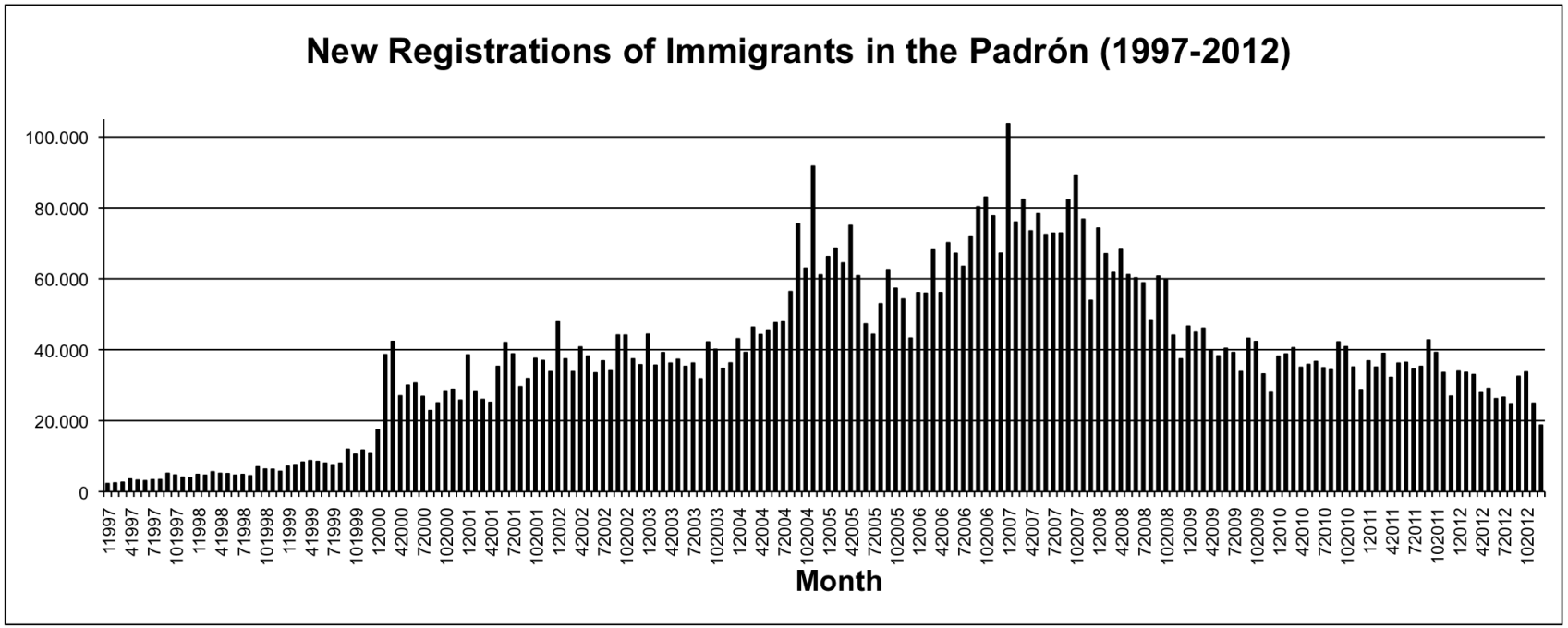
Defining Immigrants

Living in Spain (1-1-2008)	People	Share
Total:	46,157,822	100.0%
Immigrants (foreign-born)	6,044,528	13.1%
Spanish	*1,037,663*	*2.2%*
Foreigners	*5,006,865*	*10.8%*
Natives (Spanish-born)	40,113,294	86.9%
Spanish	*39,851,397*	*86.3%*
Foreigners	*261,897*	*0.6%*

Population Sizes

Year	Population	Immigrants	Share
1998	39,852,650	1,173,767	2.9%
1999	40,202,158	1,259,054	3.1%
2000	40,499,790	1,472,458	3.6%
2001	41,116,842	1,969,269	4.8%
2002	41,837,894	2,594,052	6.2%
2003	42,717,064	3,302,440	7.7%
2004	43,197,684	3,693,806	8.6%
2005	44,108,530	4,391,484	10.0%
2006	44,708,964	4,837,622	10.8%
2007	45,200,737	5,249,993	11.6%
2008	46,157,822	6,044,528	13.1%

Law Changes

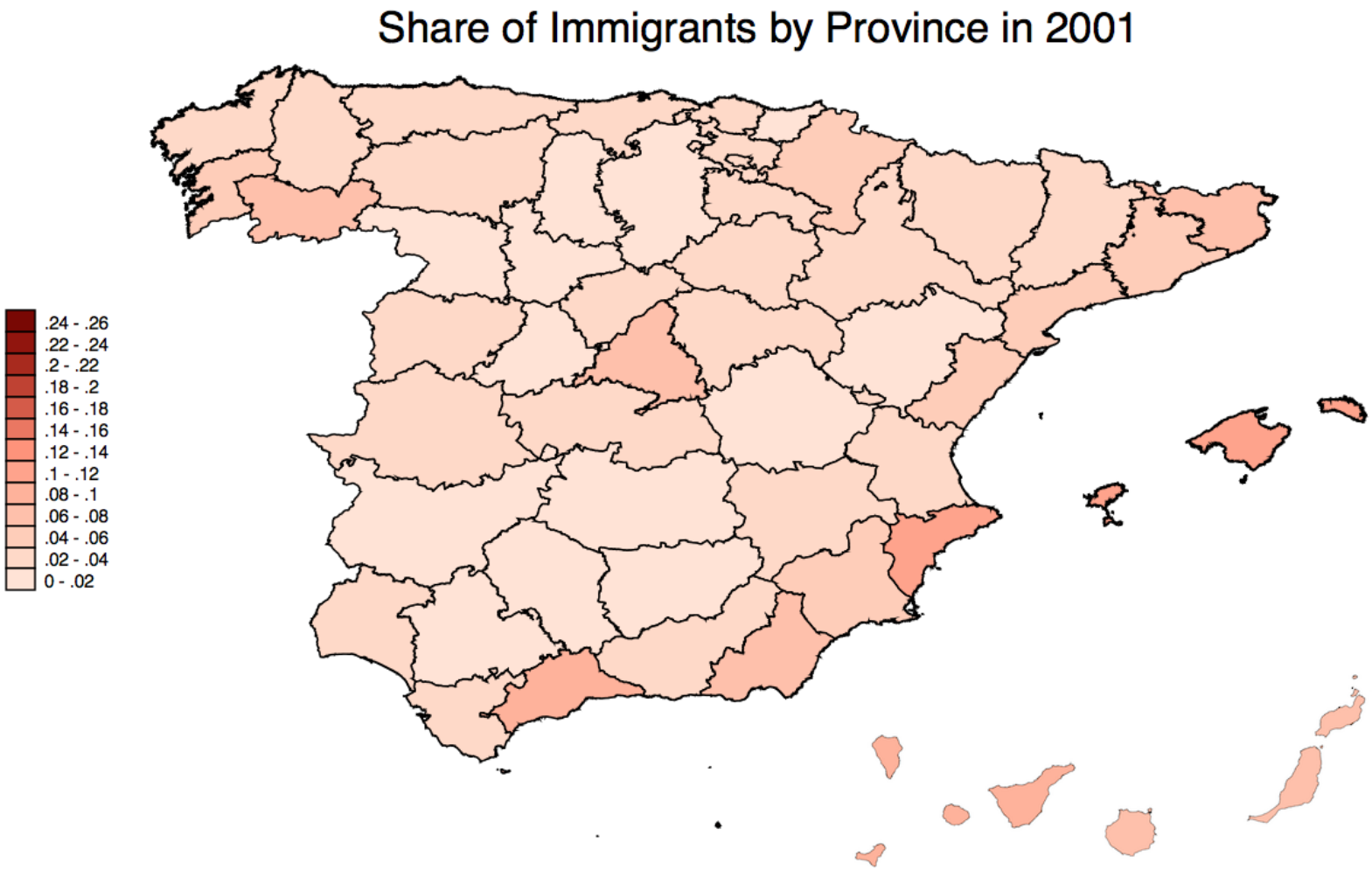


- January 2000. Law 4/2000.
- November 2004. 2005 Amnesty is announced.
- January 2007. Romania and Bulgaria enter the EU.

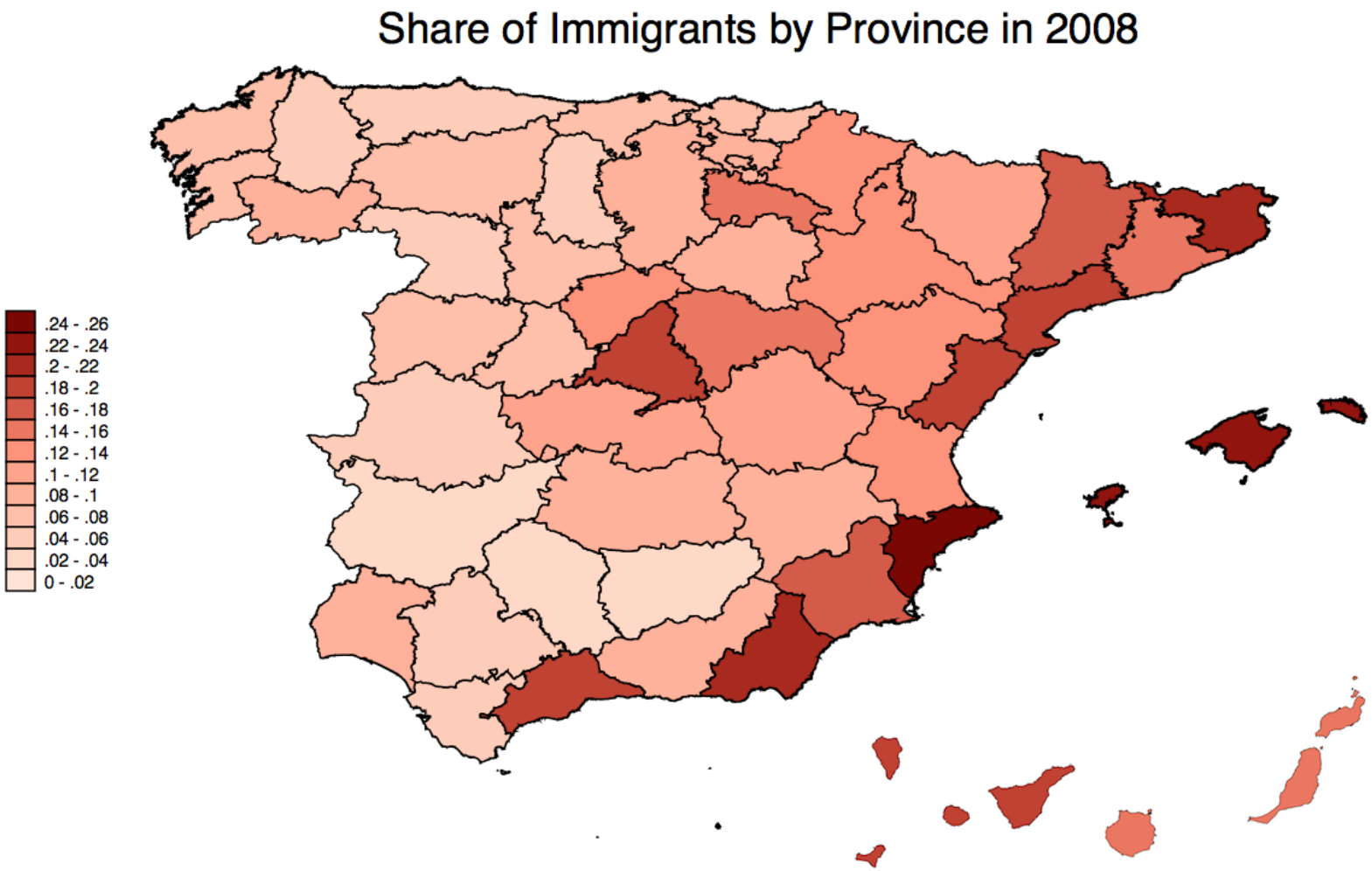
Composition

Year	Share of Total Immigration from:				
	Eastern Europe	Latin America	Subsaharian Africa	Maghreb	Other
1998	1.9%	24.7%	3.0%	17.9%	52.5%
1999	1.9%	25.3%	3.3%	16.9%	52.6%
2000	2.7%	26.5%	3.5%	17.4%	49.8%
2001	5.2%	32.8%	3.7%	16.7%	41.6%
2002	7.5%	37.7%	3.6%	15.9%	35.4%
2003	9.9%	40.3%	3.4%	14.7%	31.6%
2004	11.9%	42.2%	3.4%	14.2%	28.4%
2005	13.9%	40.9%	3.6%	14.0%	27.6%
2006	15.1%	39.3%	3.7%	13.7%	28.2%
2007	16.9%	38.7%	3.5%	12.9%	28.0%
2008	19.2%	38.0%	3.4%	12.3%	27.0%
Immigrants in 2008	1,161,242	2,298,727	208,497	745,637	1,630,425
2001-2008 increase	1,058,819	1,652,575	135,755	416,542	811,468
Shares	26.0%	40.6%	3.3%	10.2%	19.9%

2001 Map


Share of Immigrants by Province in 2001
.24 - .26
.22 - .24
.2 - .22
.18 - .2
.16 - .18
.14 - .16
.12 - .14
.1 - .12
.08 - .1
.06 - .08
.04 - .06
.02 - .04
0 - .02

2008 Map


Share of Immigrants by Province in 2008
.24 - .26
.22 - .24
.2 - .22
.18 - .2
.16 - .18
.14 - .16
.12 - .14
.1 - .12
.08 - .1
.06 - .08
.04 - .06
.02 - .04
0 - .02

Concentration in 2008

Madrid, Canary Islands and 11 Mediterranean provinces concentrate 75.2 percent of the immigrant population and 53.3 percent of the native population.

Spanish metro areas, defined by Ministerio de Vivienda (2007), concentrate 72.7 percent of the immigrant population and 66.9 percent of the native population.

2008	Natives	Immigrants	Municipalities	% Inmigrants
Metro Areas	66.9%	72.7%	744	14.1%
Rest of Spain	33.1%	27.3%	7,368	11.0%

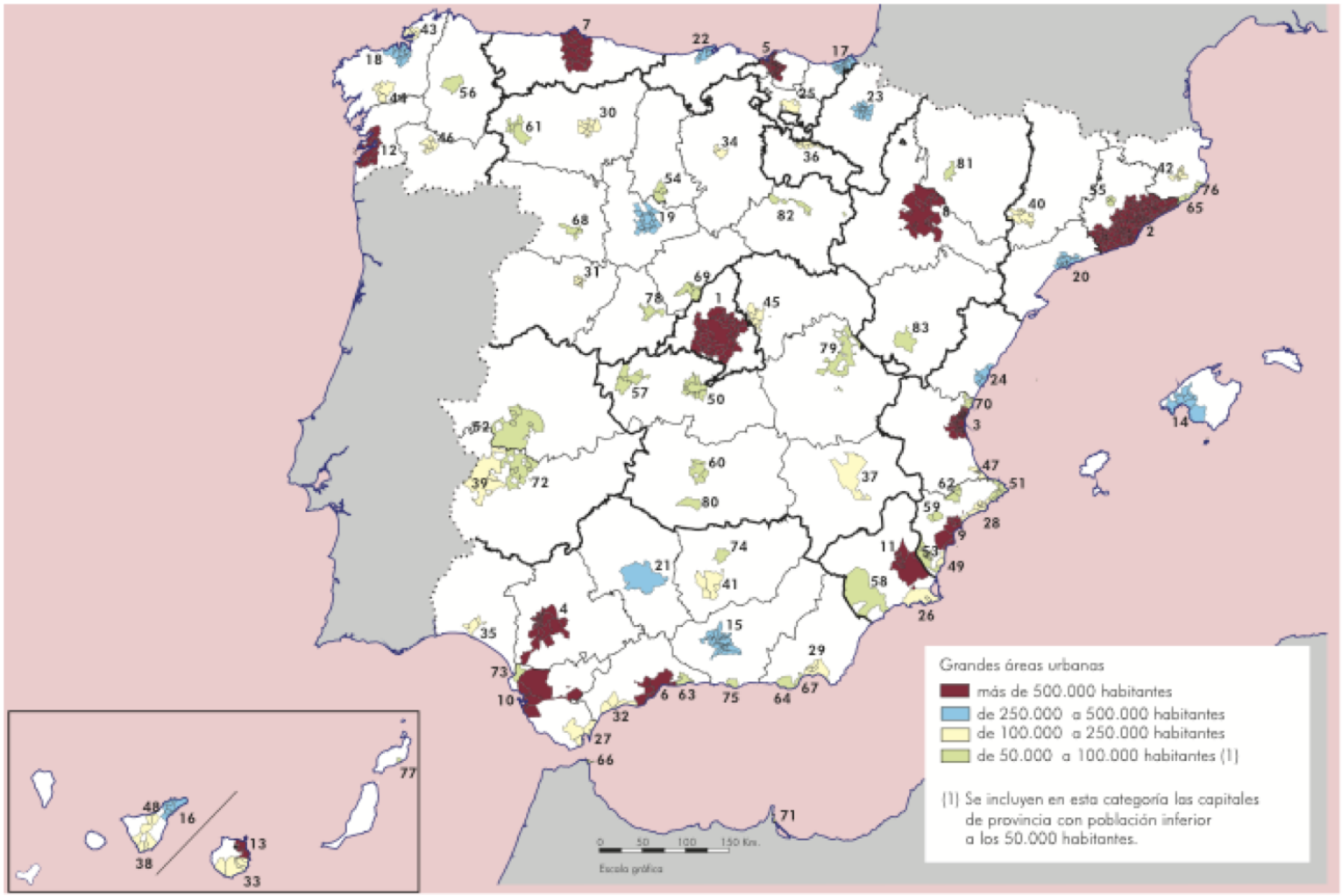
Concentration in 2008

Madrid, Canary Islands and 11 Mediterranean provinces concentrate 75.2 percent of the immigrant population and 53.3 percent of the native population.

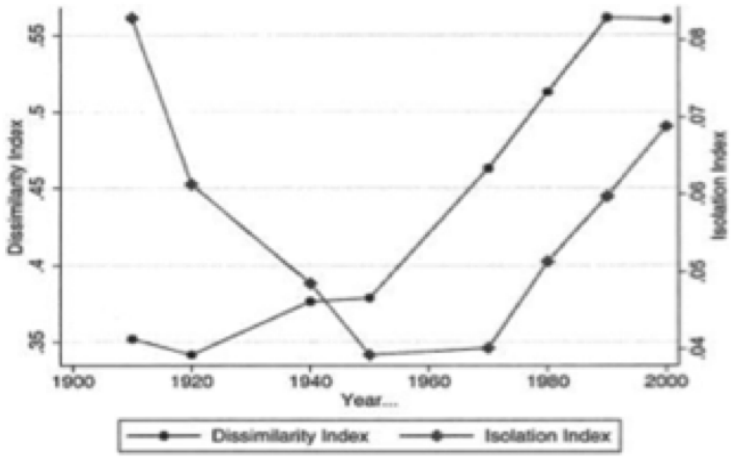
Spanish metro areas, defined by Ministerio de Vivienda (2007), concentrate 72.7 percent of the immigrant population and 66.9 percent of the native population.

2008	Natives	Immigrants	Municipalities	% Inmigrants
Metro Areas	66.9%	72.7%	744	14.1%
Rest of Spain	33.1%	27.3%	7,368	11.0%

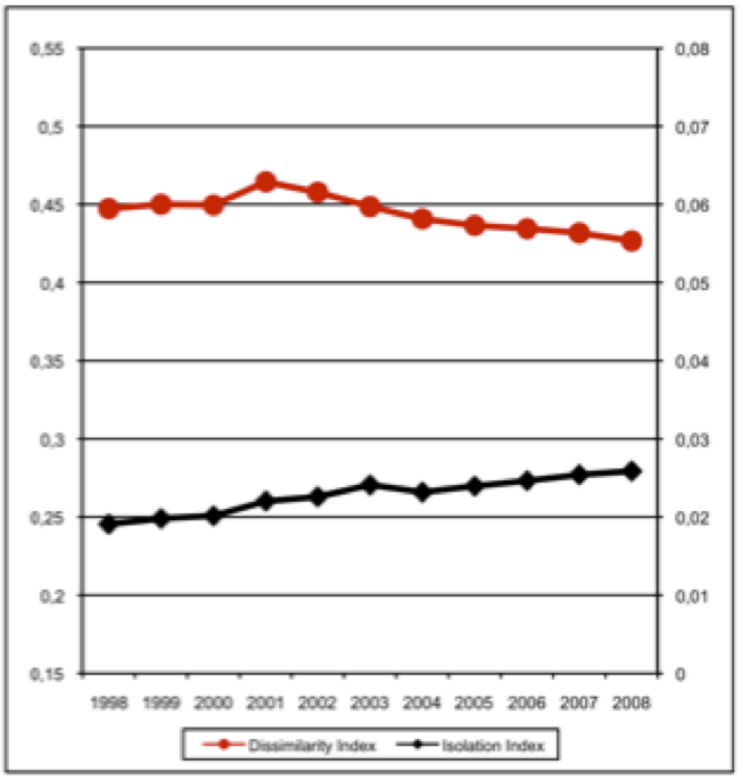
Spanish Metro Areas



Segregation in Metro Areas. Comparison with the US



US



SPAIN

Address Characteristics

We geocoded our Padrón data by matching each address with addresses from ESRI StreetMap Premium Europe NAVTEQ 2009 Release 2. We end up with 7,568,601 uniquely identified addresses.

For each address, we calculated its distance to a series of 62 features (points of interest) from the map server, such as hospitals, exit roads, schools, bus stops, metro stops, etc. In the end, for each address, we have six different measures of amenities for each of the 62 points of interest. ▸ POIs

Address Characteristics

We geocoded our Padrón data by matching each address with addresses from ESRI StreetMap Premium Europe NAVTEQ 2009 Release 2. We end up with 7,568,601 uniquely identified addresses.

For each address, we calculated its distance to a series of 62 features (points of interest) from the map server, such as hospitals, exit roads, schools, bus stops, metro stops, etc. In the end, for each address, we have six different measures of amenities for each of the 62 points of interest. ▸ POIs

Gravities

The first measure is the minimum distance between each address and each of the points of interest.

The other five measures are gravities: sums of points of interest in Spain weighted by distance. That is:

$$g_i^{p,\alpha} = \sum_{n_p=1}^{N_p} d_{i,n_p}^{-\alpha}$$

where i is an address, p is a point of interest (i.e. hospitals), N_p is the number of points of interest p in the radius where i is located, α is a coefficient that takes values $\{0.5; 1; 2; 3; 4\}$ and d_{i,n_p} is the distance between address i and point of interest n_p (i.e. one particular hospital).

Gravities

The first measure is the minimum distance between each address and each of the points of interest.

The other five measures are gravities: sums of points of interest in Spain weighted by distance. That is:

$$g_i^{p,\alpha} = \sum_{n_p=1}^{N_p} d_{i,n_p}^{-\alpha}$$

where i is an address, p is a point of interest (i.e. hospitals), N_p is the number of points of interest p in the radius where i is located, α is a coefficient that takes values $\{0.5; 1; 2; 3; 4\}$ and d_{i,n_p} is the distance between address i and point of interest n_p (i.e. one particular hospital).

2001 Census Data

The 2001 Spanish Census provides us with a set of variables referred to each of 34,251 censal sections in Spain. Censal sections are administrative divisions for electoral purposes and are supposed to have between 500 and 2,500 inhabitants. In 2001, their average population was 1,193 (s.d.=590). 94 percent of them had the correct size.

We assign each of our addresses to the 291 average characteristics of its censal section in 2001. Variables included are: age structure, education, unemployment rates, industry composition of the workforce, quality of the buildings and neighborhood, commuting habits, etc.

2001 Census Data

The 2001 Spanish Census provides us with a set of variables referred to each of 34,251 censal sections in Spain. Censal sections are administrative divisions for electoral purposes and are supposed to have between 500 and 2,500 inhabitants. In 2001, their average population was 1,193 (s.d.=590). 94 percent of them had the correct size.

We assign each of our addresses to the 291 average characteristics of its censal section in 2001. Variables included are: age structure, education, unemployment rates, industry composition of the workforce, quality of the buildings and neighborhood, commuting habits, etc.

Equation in rates

We estimate the response of natives to the arrival of immigrants with the following equation borrowed from the literature (i.e. Saiz and Wachter, 2011):

$$\frac{\Delta nat_{n,t}}{pop_{n,t-1}} = \beta \frac{\Delta mig_{n,t}}{pop_{n,t-1}} + \epsilon_{n,t}$$

where n refers to neighborhoods. $\beta < 0$ would imply displacement of natives by immigrants, endogeneity problems aside.

We estimate a long-differences model with changes between 2001 and 2008.

Problem 1: Due to the housing boom in Spain, many neighborhoods were empty in 2001 but full in 2008.

Equation in rates

We estimate the response of natives to the arrival of immigrants with the following equation borrowed from the literature (i.e. Saiz and Wachter, 2011):

$$\frac{\Delta nat_{n,t}}{pop_{n,t-1}} = \beta \frac{\Delta mig_{n,t}}{pop_{n,t-1}} + \epsilon_{n,t}$$

where n refers to neighborhoods. $\beta < 0$ would imply displacement of natives by immigrants, endogeneity problems aside.

We estimate a long-differences model with changes between 2001 and 2008.

Problem 1: Due to the housing boom in Spain, many neighborhoods were empty in 2001 but full in 2008

Equation in rates

We estimate the response of natives to the arrival of immigrants with the following equation borrowed from the literature (i.e. Saiz and Wachter, 2011):

$$\frac{\Delta nat_{n,t}}{pop_{n,t-1}} = \beta \frac{\Delta mig_{n,t}}{pop_{n,t-1}} + \epsilon_{n,t}$$

where n refers to neighborhoods. $\beta < 0$ would imply displacement of natives by immigrants, endogeneity problems aside.

We estimate a long-differences model with changes between 2001 and 2008.

Problem 1: Due to the housing boom in Spain, many neighborhoods were empty in 2001 but full in 2008.

Equation in levels

Solution 1: We estimate the classical model in levels.

$$\Delta nat_{n,t} = \beta \Delta mig_{n,t} + \gamma pop_{n,t-1} + \varepsilon_{n,t}$$

Problem 2: We need homogeneous neighborhoods. Administrative divisions change over time.

Solution 2: We use geocoded addresses to build stable neighborhoods.

Equation in levels

Solution 1: We estimate the classical model in levels.

$$\Delta nat_{n,t} = \beta \Delta mig_{n,t} + \gamma pop_{n,t-1} + \varepsilon_{n,t}$$

Problem 2: We need homogeneous neighborhoods. Administrative divisions change over time.

Solution 2: We use geocoded addresses to build stable neighborhoods.

Equation in levels

Solution 1: We estimate the classical model in levels.

$$\Delta nat_{n,t} = \beta \Delta mig_{n,t} + \gamma pop_{n,t-1} + \varepsilon_{n,t}$$

Problem 2: We need homogeneous neighborhoods. Administrative divisions change over time.
Solution 2: We use geocoded addresses to build stable neighborhoods.

Gridding Spain

We use three alternative methodologies:

- Take censal sections in 2008 and bring them back in time.
- Create squares of 0.01 degrees, approximately 1.1 km.
- Create squares of 0.005 degrees, approximately 555 meters. We prefer this measure because it gives us similar averages while it does not depend on administrative decisions. For example, censal sections with a larger share of non-voting immigrants are larger. For 2008 metro areas, the average population of the 28,870 grids is 1,070 (s.d.=2,134).

Gridding Spain

We use three alternative methodologies:

- Take censal sections in 2008 and bring them back in time.
- Create squares of 0.01 degrees, approximately 1.1 km.
- Create squares of 0.005 degrees, approximately 555 meters. We prefer this measure because it gives us similar averages while it does not depend on administrative decisions. For example, censal sections with a larger share of non-voting immigrants are larger. For 2008 metro areas, the average population of the 28,870 grids is 1,070 (s.d.=2,134).

Gridding Spain

We use three alternative methodologies:

- Take censal sections in 2008 and bring them back in time.
- Create squares of 0.01 degrees, approximately 1.1 km.
- Create squares of 0.005 degrees, approximately 555 meters. We prefer this measure because it gives us similar averages while it does not depend on administrative decisions. For example, censal sections with a larger share of non-voting immigrants are larger. For 2008 metro areas, the average population of the 28,870 grids is 1,070 (s.d.=2,134).

Gridding Spain

We use three alternative methodologies:

- Take censal sections in 2008 and bring them back in time.
- Create squares of 0.01 degrees, approximately 1.1 km.
- Create squares of 0.005 degrees, approximately 555 meters. We prefer this measure because it gives us similar averages while it does not depend on administrative decisions. For example, censal sections with a larger share of non-voting immigrants are larger. For 2008 metro areas, the average population of the 28,870 grids is 1,070 (s.d.=2,134).

Final Model

Given that we have equal-surface neighborhoods, the population at baseline variable $pop_{n,t-1}$ from the equation in levels automatically controls for population density.

The final model we estimate is:

$$\Delta nat_{n,t} = \beta \Delta mig_{n,t} + f(pop_{n,t-1}) + \Gamma' A_{n,t} + \Phi' S_{n,t-1} + \xi_{n,t}$$

where $f(pop_{n,t-1})$ is a flexible function (quartic polynomial) of population at baseline, $A_{n,t}$ is a vector of amenities (gravities) and $S_{n,t-1}$ is a vector of socioeconomic characteristics at baseline.

Final Model

Given that we have equal-surface neighborhoods, the population at baseline variable $pop_{n,t-1}$ from the equation in levels automatically controls for population density.

The final model we estimate is:

$$\Delta nat_{n,t} = \beta \Delta mig_{n,t} + f(pop_{n,t-1}) + \Gamma' A_{n,t} + \Phi' S_{n,t-1} + \xi_{n,t}$$

where $f(pop_{n,t-1})$ is a flexible function (quartic polynomial) of population at baseline, $A_{n,t}$ is a vector of amenities (gravities) and $S_{n,t-1}$ is a vector of socioeconomic characteristics at baseline.

Variables in the Model

- Population variables ($\Delta nat_{n,t}$, $\Delta mig_{n,t}$ and $pop_{n,t-1}$): We exclude children 0-15 years old to avoid population increases due to newly-born immigrant children born as natives.
- Mortality and age structure controls: We control for the baseline share of native population in age groups 15-24, 25-44, 45-64 and 65 and more, and for the baseline share of migrants.
- Additional neighborhood controls: Dummies for zero population in 2001 and for 2008; distances to population-weighted metro area and to municipality center; all gravities; 2001 Census variables: unemployment, construction, housekeeping, hotel and restaurant workers, age of buildings and p.c. indices for car use, walking to work habits, height of the buildings and neighborhood cleanliness.

Summary Statistics

Variables in the Model

- Population variables ($\Delta nat_{n,t}$, $\Delta mig_{n,t}$ and $pop_{n,t-1}$): We exclude children 0-15 years old to avoid population increases due to newly-born immigrant children born as natives.
- Mortality and age structure controls: We control for the baseline share of native population in age groups 15-24, 25-44, 45-64 and 65 and more, and for the baseline share of migrants.
- Additional neighborhood controls: Dummies for zero population in 2001 and for 2008; distances to population-weighted metro area and to municipality center; all gravities; 2001 Census variables: unemployment, construction, housekeeping, hotel and restaurant workers, age of buildings and p.c. indices for car use, walking to work habits, height of the buildings and neighborhood cleanliness.

Summary Statistics

Variables in the Model

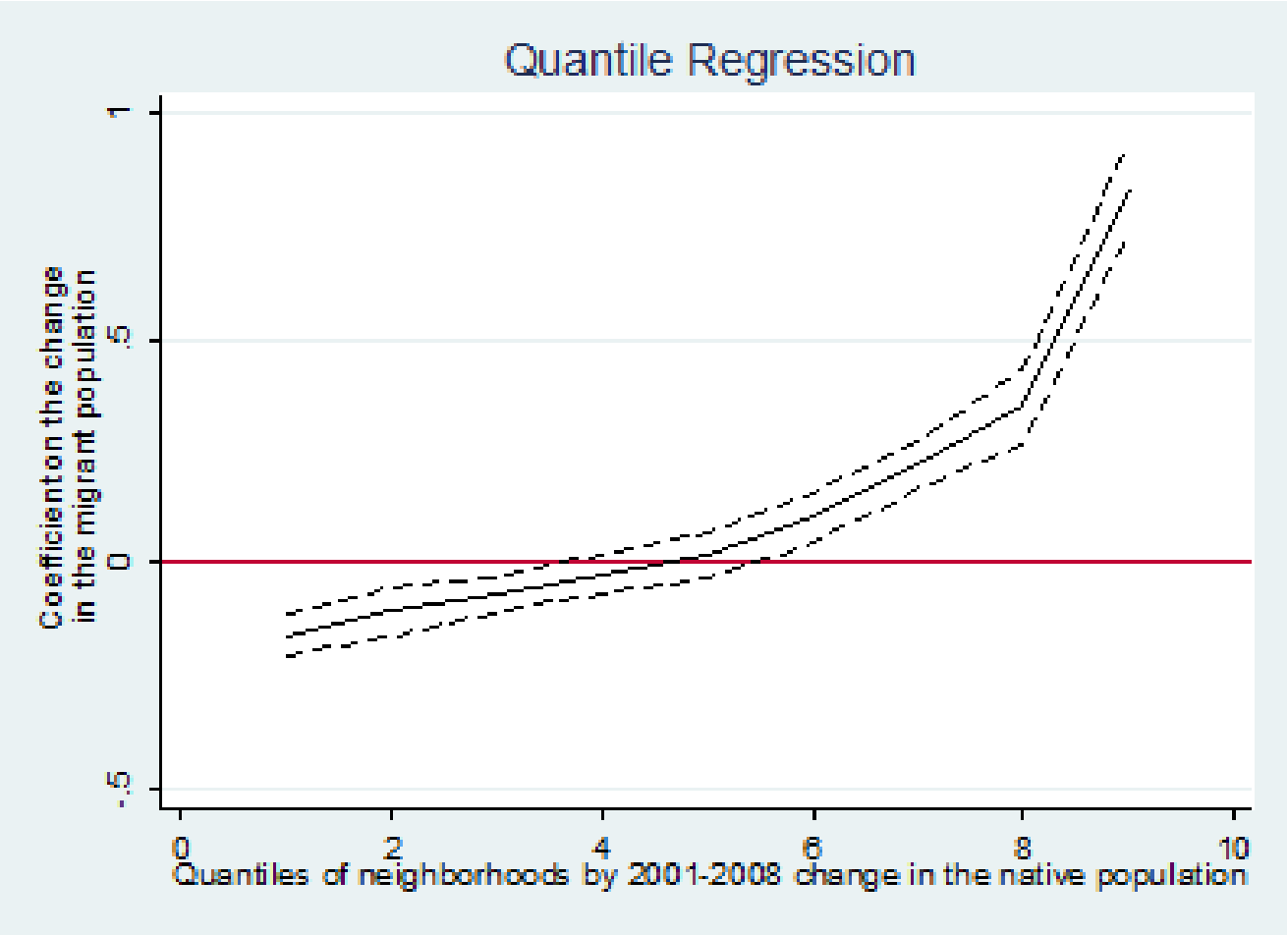
- Population variables ($\Delta nat_{n,t}$, $\Delta mig_{n,t}$ and $pop_{n,t-1}$): We exclude children 0-15 years old to avoid population increases due to newly-born immigrant children born as natives.
- Mortality and age structure controls: We control for the baseline share of native population in age groups 15-24, 25-44, 45-64 and 65 and more, and for the baseline share of migrants.
- Additional neighborhood controls: Dummies for zero population in 2001 and for 2008; distances to population-weighted metro area and to municipality center; all gravities; 2001 Census variables: unemployment, construction, housekeeping, hotel and restaurant workers, age of buildings and p.c. indices for car use, walking to work habits, height of the buildings and neighborhood cleanliness.

Summary Statistics

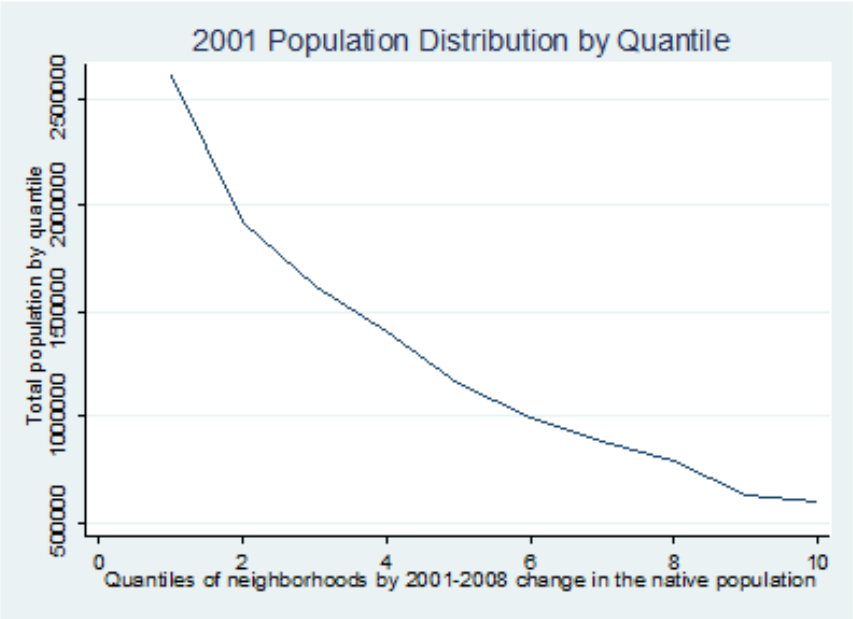
OLS: rates, levels and winsorizing

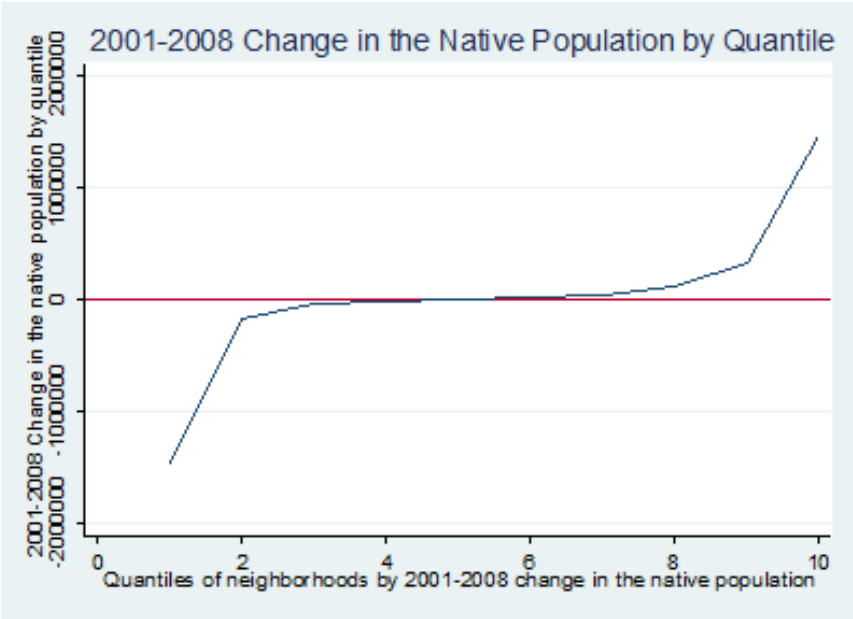
Dependent variable	$\frac{\Delta nat_{n,2001\_2008}}{pop_{n,2001}}$	$\Delta nat_{n,2001\_2008}$				
Sample	All	All	All	All	Outliers	Winsorized
Variables	(1)	(2)	(3)	(4)	(5)	(6)
$\frac{\Delta mig_{n,2001\_2008}}{pop_{n,2001}}$	0.054					
	[0.026]**					
$\Delta mig_{n,2001\_2008}$		0.038	0.055	0.078	0.643	-0.079
		[0.042]	[0.044]	[0.046]*	[0.120]***	[0.042]*
$pop_{n,2001}$		-0.076				
		[0.006]***				
Constant	-0.026	70.489	28.029			
	[0.004]***	[2.219]***	[1.938]***			
$f(pop_{n,2001})$	No	No	Yes	Yes	Yes	Yes
Controls	No	No	No	Yes	Yes	Yes
Metro Dummies	No	No	No	Yes	Yes	Yes
Adjusted R^2	0.001	0.177	0.223	0.301	0.621	0.359
Observations	26,525	26,536	26,536	26,536	4,491	22,045

Quantile regressions



Quantile regressions: descriptives





Building the Instrument

Because our neighborhoods are purely geographical, we improve upon the Saiz and Wachter (2011) instrument.

We calculate the average share of immigrants at baseline in the eight contiguous neighborhoods (population-weighted) for each of our five nationality groups.

The thought experiment is to compare identical neighborhoods with no immigrants around in 2001 with neighborhoods with many immigrants around in 2001.

Building the Instrument

Because our neighborhoods are purely geographical, we improve upon the Saiz and Wachter (2011) instrument.

We calculate the average share of immigrants at baseline in the eight contiguous neighborhoods (population-weighted) for each of our five nationality groups.

The thought experiment is to compare identical neighborhoods with no immigrants around in 2001 with neighborhoods with many immigrants around in 2001.

Building the Instrument

Because our neighborhoods are purely geographical, we improve upon the Saiz and Wachter (2011) instrument.

We calculate the average share of immigrants at baseline in the eight contiguous neighborhoods (population-weighted) for each of our five nationality groups.

The thought experiment is to compare identical neighborhoods with no immigrants around in 2001 with neighborhoods with many immigrants around in 2001.

IV Results

Dependent variable	$\Delta nat_{n,2001\_2008}$					
Model	OLS	IV	OLS	IV	OLS	IV
Sample	All	All	Outliers	Outliers	Winsorized	Winsorized
Variables	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta mig_{n,2001\_2008}$	0.078	-0.427	0.643	0.184	-0.079	-0.450
	[0.046]*	[0.140]***	[0.120]***	[0.212]	[0.042]*	[0.136]***
$f(pop_{n,2001})$	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Metro Dummies	Yes	Yes	Yes	Yes	Yes	Yes
F-test	-	26.64	-	14.08	-	21.91
Sargan test p-value	-	0.245	-	0.464	-	0.106
Partial R^2	-	0.005	-	0.016	-	0.005
Adjusted R^2	0.301	-	0.621	-	0.359	-
Observations	26,536	26,536	4,491	4,491	22,045	22,045

Robustness

Negative IV coefficients are robust to a number of alternative estimation strategies:

- Adding interactions with baseline share of migrants in the neighborhood.
- Adding maximum share of migrants among contiguous neighborhoods and its interaction with the baseline share of migrants in the neighborhood.
- IV in levels rather than in shares.
- Building the instrument pooling nationalities.

Some of the robustness checks take the IV coefficient down to -0.8 or even -1.

Robustness

Negative IV coefficients are robust to a number of alternative estimation strategies:

- Adding interactions with baseline share of migrants in the neighborhood.
- Adding maximum share of migrants among contiguous neighborhoods and its interaction with the baseline share of migrants in the neighborhood.
- IV in levels rather than in shares.
- Building the instrument pooling nationalities.

Some of the robustness checks take the IV coefficient down to -0.8 or even -1.

The Housing boom

Quantile regressions suggest two types of neighborhoods coexist (two Spains): those losing native population as migrants come in and those winning both native and immigrant population.

The two Spains are different along a number of dimensions, many related to the housing boom, like the share of open land at baseline. However, the best separating criterium is municipality size.

2001-2008 Change in:	Large Municipalities: more than 100.000	Small Municipalities: less than 100.000
Native Population	-565,999	737,487
Migrant Population	1,468,311	1,105,598

The Housing boom

Quantile regressions suggest two types of neighborhoods coexist (two Spains): those losing native population as migrants come in and those winning both native and immigrant population.

The two Spains are different along a number of dimensions, many related to the housing boom, like the share of open land at baseline. However, the best separating criterium is municipality size.

2001-2008 Change in:	Large Municipalities: more than 100.000	Small Municipalities: less than 100.000
Native Population	-565,999	737,487
Migrant Population	1,468,311	1,105,598

The Housing boom

Quantile regressions suggest two types of neighborhoods coexist (two Spains): those losing native population as migrants come in and those winning both native and immigrant population.

The two Spains are different along a number of dimensions, many related to the housing boom, like the share of open land at baseline. However, the best separating criterium is municipality size.

2001-2008 Change in:	Large Municipalities: more than 100.000	Small Municipalities: less than 100.000
Native Population	-565,999	737,487
Migrant Population	1,468,311	1,105,598

Population movements between the two Spains

Dependent variable	$\Delta nat_{c,2001\_2008}$
$\Delta mig_{c,2001\_2008}$	-1.212
	[0.586]***
$pop_{c,2001}$	0.000
	[0.065]
Constant	35454.721
	[3546.391]***
Metro Dummies	Yes
Adjusted R^2	0.768
Observations	72

We can summarize population movements between the two Spains through a simple regression where each observation, denoted by c, is just a metro area-municipality size pair. The cutoff municipality size is 100,000 in 2001.

There were only 36 metro areas with at least one municipality larger than 100,000.

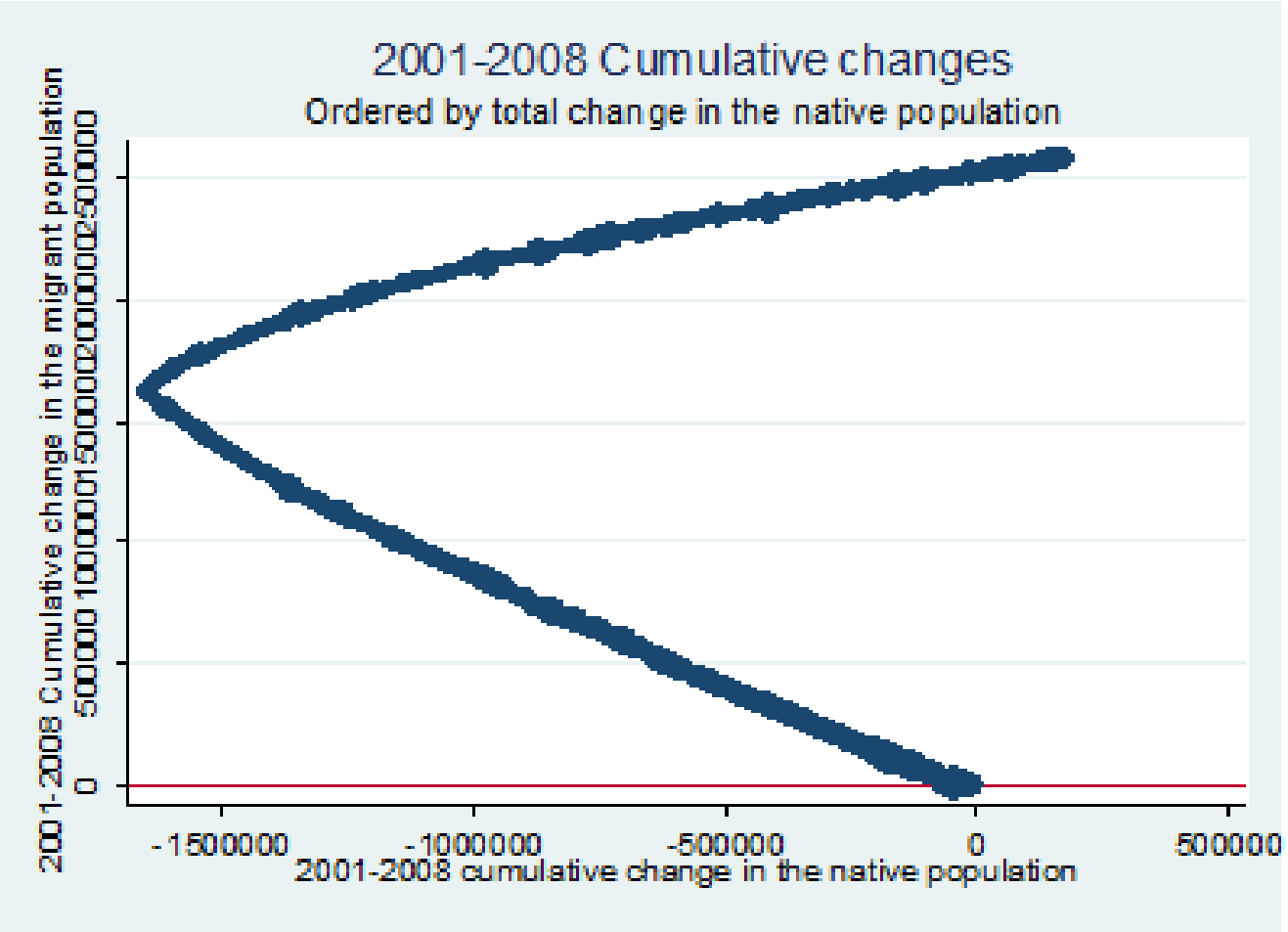
Population movements between the two Spains

Dependent variable	$\Delta nat_{c,2001\_2008}$
$\Delta mig_{c,2001\_2008}$	-1.212
	[0.586]***
$pop_{c,2001}$	0.000
	[0.065]
Constant	35454.721
	[3546.391]***
Metro Dummies	Yes
Adjusted R^2	0.768
Observations	72

We can summarize population movements between the two Spains through a simple regression where each observation, denoted by c, is just a metro area-municipality size pair. The cutoff municipality size is 100,000 in 2001.

There were only 36 metro areas with at least one municipality larger than 100,000.

Cumulative changes



Conclusions

International migration made Spain's population grow by 10 percent between 1998 and 2008.

This was the fastest and largest increase in the OECD but it is otherwise an extreme example of the migration patterns between poor and rich countries in the last 20 years.

We show how this massive inflow affected residential choices. Spanish natives escaped immigrants arriving into central cities and large towns by moving to new developments further from the center. However, some immigrants also settled there so that the resulting segregation patterns appeared unaffected.

Conclusions

International migration made Spain's population grow by 10 percent between 1998 and 2008.

This was the fastest and largest increase in the OECD but it is otherwise an extreme example of the migration patterns between poor and rich countries in the last 20 years.

We show how this massive inflow affected residential choices. Spanish natives escaped immigrants arriving into central cities and large towns by moving to new developments further from the center. However, some immigrants also settled there so that the resulting segregation patterns appeared unaffected.

Conclusions

International migration made Spain's population grow by 10 percent between 1998 and 2008.

This was the fastest and largest increase in the OECD but it is otherwise an extreme example of the migration patterns between poor and rich countries in the last 20 years.

We show how this massive inflow affected residential choices. Spanish natives escaped immigrants arriving into central cities and large towns by moving to new developments further from the center. However, some immigrants also settled there so that the resulting segregation patterns appeared unaffected.

Points of Interest

Exits (from highways or roads)	Restaurant	Performing Arts	Museum
Winery	Nightlife	Bowling Centre	City Hall
ATM	Historical Monument	Sports Complex	Police Station
Train Station	Bank	Park/Recreation Area	Post Office
Commuter Rail Station	Shopping	Casino	Department Store
Bus Station	Hotel	Convention/Exhibition Centre	Home Specialty Store
Ferry Terminal	Ski Resort	Golf Course	Pharmacy
Marina	Other Accommodation	Civic/Community Centre	Specialty Store
Public Sports Airport	Tourist Information	Amusement Park	Sporting Goods Store
Airport	Rental Car Agency	Sports Center	Medical Service
Business Facility	Parking Lot	Ice Skating Rink	Consumer Electronics Store
Grocery Store	Parking Garage/House	Tourist Attraction	Industrial Zone
Automobile Dealership	Park & Ride	Hospital	Place of Worship
Petrol/Gasoline Station	Auto Service & Maintenance	Higher Education	Embassy
Motorcycle Dealership	Cinema	School	Book Store
	Rest Area	Library	

◂ Back

Summary Statistics: main variables

Back

Variable	Average	St. dev.	Min	Max
$\Delta nat_{n,2001\_2008}$	12.56	311.34	-10773.00	5160.00
$\Delta mig_{n,2001\_2008}$	91.03	256.14	-1890.00	7014.00
$pop_{n,2001}$	819.48	1783.88	0.00	25139.00
share mig. 2001	0.08	0.16	0.00	1.00
no pop. in 2008	0.02	0.13	0.00	1.00
no pop. in 2001	0.07	0.25	0.00	1.00
age15_24	0.13	0.09	0.00	1.00
age25_44	0.31	0.15	0.00	1.00
age45_64	0.21	0.13	0.00	1.00
age65plus	0.14	0.14	0.00	1.00
log distance metro area center	1.57	0.95	-4.84	3.61
log distance municipality center	0.24	1.01	-5.75	3.49

Summary Statistics: 2001 Census variables

Back

Variable	Average	St. dev.	Min	Max
unemployment	11.20	6.89	0.00	60.39
building	10.54	6.64	0.00	48.68
hotel and rest.	5.88	5.45	0.00	53.97
housekeep.	2.80	1.45	0.00	12.50
btw41_50	3.32	5.46	0.00	99.92
btw51_60	7.04	8.69	0.00	99.62
btw61_70	12.72	12.97	0.00	100.00
btw71_80	19.30	15.26	0.00	100.00
btw81_90	15.81	14.29	0.00	100.00
btw91_00	19.38	17.57	0.00	100.00
car index	26.72	19.33	-18.71	81.49
pedestrian index	45.59	15.12	0.00	85.79
height index	60.86	79.26	-61.46	254.78
neighborhood quality index	85.43	50.60	-25.18	226.65