UC Irvine
UC Irvine Electronic Theses and Dissertations

Title
A Theoretical and Empirical Approach to Network Effects

Permalink
https://escholarship.org/uc/item/7m44s0zq

Author
Gutierrez, Victor Guillermo

Publication Date
2018

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA,
IRVINE

A Theoretical and Empirical Approach to Network Effects

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Economics

by

Victor Guillermo Gutierrez

Dissertation Committee:
Associate Professor Jiawei Chen, Chair
Professor Linda R. Cohen
Professor Amihai Glazer

2018

©2018 Victor Guillermo Gutierrez

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank the University of California Irvine for the great human and academic resources offered to me.

To the School of Social Sciences and the support of the Merit Fellowship in Economics that allowed me to focus in school and the pursue of knowledge.

To my professors, specially to my committee, Jiawei Chen, Linda Cohen, and Amihai Glazer, for their disposition, patience, and their invaluable advice. Their support was transcendental.

To my parents, wife, and relatives. Their support and encouragement allowed me to pursue my academic goals.

CURRICULUM VITAE

Victor Guillermo Gutierrez

Education

PhD in Economics **2018**
University of California Irvine *Irvine, California*

Masters in Economics **2015**
University of California Irvine *Irvine, California*

Bachelors in Economics **2012**
Autonomous University of Baja California *Tijuana, Mexico*

International Baccalaureate Diploma **2008**
Federal High School Lazaro Cardenas *Tijuana, Mexico*

Experience

Part-time Faculty **2016-2018**
Orange Coast College *Costa Mesa, California*

Teaching Assistant **2013-2018**
University of California Irvine *Irvine, California*

Analyst of Market Research **2012-2018**
Testa Marketing *Tijuana, Mexico*

Economics Analyst **2012-2013**
Center for Economic Studies of the Private Sector *Tijuana, Mexico*

Economics Intern **2012**
Metropolitan Planning Institute of Tijuana *Tijuana, Mexico*

Analyst's Coordinator **2011-2012**
Public Opinion Institute of Baja California *Tijuana, Mexico*

Analyst Intern **2011**
Economists Association of Baja California *Tijuana, Mexico*

ABSTRACT OF THE DISSERTATION

A Theoretical and Empirical Approach to Network Effects

by

Victor Guillermo Gutierrez

Doctor of Philosophy in Economics

University of California, Irvine, 2018

Associate Professor Jiawei Chen, Chair

This dissertation approaches network effects from different perspectives and methodologies. Social network effects of immigrants have an impact on the decision of their destination, the presence of this characteristic is exploited in the first chapter to statistically isolate the effect they have on the voting behavior of the district's representative. Generally, more immigration is correlated with a more liberal voting behavior. The second chapter takes a theoretical approach, a network adoption model is constructed that deviates from the usual linear assumption and incorporates time of adoption and a type-dependent utility function that allows a different utility derivation dependent on the type of members in the networks and not only in size as the linear model assumes. Using simulations, it is found that more heterogeneous populations and the presence of early adopters allow the emergence of clusters and the survival of small platforms even in the presence of strong network effects.

INTRODUCTION

This dissertation approaches network effects from different perspectives. It adds a more social input to the analysis by linking the demographic characteristics of the population to the theoretical and empirical analysis.

There are four main themes present in this work: immigration, determination of policy, cluster formation, and concentration in the Operating Systems market. The big themes are all linked by the presence of networks in one way or another. Network effects have been studied extensively in the literature, but this work includes the use of demographic characteristics of the population to derive additional insights.

The first chapter recognizes the effect that networks of migrants have on the decision of new immigrants about where to go. Migrants want to be part of a network, and historically they go to places where established groups of immigrants are present. Network effects are present and are usually the main determinant for the choice of destination. This characteristic is the exploited to isolate the effect this growing networks have on the voting behavior of representatives. In this case, network effects are used as an isolation tool that allow the calculation of important behavior that is relevant to the determination of policy in the United States and that goes beyond simple correlation.

The second chapter takes a theoretical approach. A network model is constructed where an important assumption is relaxed. The linear assumption of network effects assumes the size of the network is the most important characteristic of value derivation and in consequence of network selection. The model presented in the chapter deviates from this assumption by allowing the users to derive different utility depending on the member's characteristics in the network and not only the size. This approach is innovative and allows the study of network selection and the emergence of clusters. By using simulations, the model can be completely controlled and it allows the analysis of markets under a high variety of parameter values.

1 The effect of non-U.S. citizens on the ideological position of their representatives

1.1 Abstract

This paper determines the effects of non-U.S. citizens on the ideological position of the representatives of their respective congressional districts. I use the DW-NOMINATE scale and the National Journal House ratings to measure the ideological position of representatives, and the percentage of non-U.S. citizens as a measure of immigrant presence in the congressional districts. The data shows a significant but small connection between the immigrant populations and the general ideological position of Congress; larger immigrant populations translates into more liberal representatives. When ideological positions by subject are considered similar effects are found in economic and foreign issues, the exception are social issues where more immigration appears to cause more conservative behavior.

1.2 Introduction

This paper relates two topics, immigration and policy determination. The goal is to better understand the relation between these two subjects, specifically the effect of the inflow of non-U.S citizen population on the voting behavior of the representatives in Congress. Non-U.S. citizens have limited participation in politics, they can't participate in the direct form of voting, but they can influence Congress in various ways. By joining groups or supporting interest groups immigrant's preferences can play a key role in the national agenda setting, and their presence can affect the preferences of voters in their districts. The main goal of this paper is to find if these alternative channels of influence are being used and more importantly to determine the extent to which they are effective.

1.3 Literature Review

The consequences of immigration in the United States have been studied intensively; there is an ongoing and repeated narrative that links immigrant populations to fiscal, social, and cultural consequences (Chavez 2008, Hopkins 2010, Brader et al 2008, Santa Ana 2004). However, the impact that immigrants have on politics is less clear. Many studies demonstrate the growing strength of minorities on voting, particularly of the latino electorate (de la Garza et al 1992, DeSipio 1996, Alvarez and Garcia Bedolla 2003, Abrajano and Alvarez 2010), other studies have found strong support of immigrants to the Democratic Party (Wong et al 2011, Alvarez and GarciaBedolla 2003, Hajnal and Lee 2011). However little work has been done in the effects that immigration has on the ideological positions of the representatives and on policy.

Immigration can affect policy and the ideological position of the representatives by changing preferences of the native population, Bauer et al. (2000) started a young literature on natives' attitudes toward immigrants looking at cross-country survey data, this paper and others (see Dustman and Preston (2006)) such as that of Mayda (2006) find evidence of a robust relationship between attitudes towards immigration and both security concerns and cultural and national identity issues.

Other extensive studies analyze the more general topic of factors that influence voting behavior in the House of Representatives and in the Senate. The analysis of roll call data that started with the important contribution of Poole ad Rosenthal (1985, 1991, 2000), as followed by many studies searching for links between roll call voting and: polarization (Garand 2010; Heberlig, Hetherington, and Larson 2006; Hetherington 2001, McCarty, Poole and Rosenthal 2006), responsiveness of representatives (Gailmard and Jenkins

2009, Griffin and Newman 2005), the president (Berry, Burden and Howell 2010; Lebo and OGeen 2011) among others, however the relation between immigration and roll call data has not been explored sufficiently.

This paper analyzes the relation between representative's voting behavior and immigration using an instrumental variable and controlling for demographic and economic variables related to the national districts.

1.4 Description of data

The data of non U.S. citizen population comes from the American Community Surveys (ACS). The total population and the number of residents that do not have the U.S. citizenship was used to calculate the percentage that the immigrant population represents to each congressional district.

$$\%Foreign_{it} = \frac{TotalNon - U.S.citizens_{it}}{TotalPopulation_{it}} \tag{1}$$

where i represents a specific congressional district and t the year of the estimates.

To measure ideological position I used the DW-NOMINATE scaling method developed by political scientists Keith T. Poole and Howard Rosenthal. The procedure calculates two measures to form a "spacial coordinate" for each member of congress. One of these coordinates can be interpreted as the familiar left-right or liberal-conservative spectrum. The method used to calculate the coordinates is based on the assumption that each member of Congress has a utility function that is bell-shaped, symmetric and single-peaked, the ideal point represents the individual's most preferred outcome, and individuals most desire outcomes are the closest to this point. Given the roll-call voting behavior of each member, the ideal point can be estimated and used as an approximation of his or her ideology. This data is available in the webpage of the authors.

Political Position is the DW score and represents the position of the representative of congressional district i in year t , this measure ranges from -1 (most liberal) to 1 (most conservative). The average position across the 1,770 observations was 0.10, which is a slight conservative position.

One of the most important variable to explain representative's ideological position is party affiliation. *Party* is a dummy variable that takes the value of 1 when the representative is a Democrat and a value of 0 if he or she is a Republican, the source of party data is also from Pool and Rosenthal.

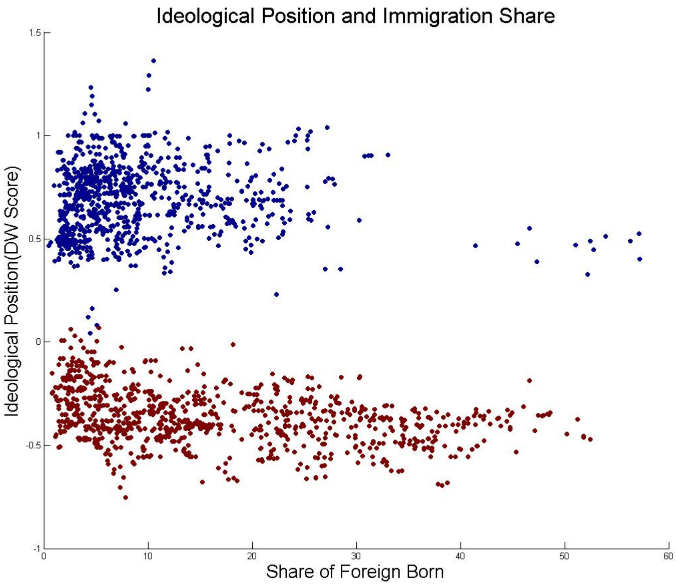

Figure 1: Ideology and immigrants by Congressional District

As shown in figure 1, when the DW scores are graphed with immigration data, there are two clusters of representatives which can be interpreted as a polarization of Congress as stated extensively in the literature. After a visual inspection it seems to be more liberal representatives in districts with higher share of immigrants, suggesting a negative relation between the DW score and the immigration share in the district.

Age represents the median age of the population in the respective district, *Age Native* the median age of only native population and *Age Foreign* of only the foreign population in the congressional district. Age could have an important effect on the voting behavior of the representative, a relative young population will demand different services and they will probably have a different set of opinions and preferences when compared with older populations.

Information related with ethnicity per congressional district is also used, I use the percentage of white, African American, Asian, Native Hawaiian, Hispanic and other races, over the total population of the respective district. There is research suggesting an important relation between ethnic heterogeneity and variables like public opinion towards immigrants, public good provision and cooperation.

As used in Mauro (1995), Canning and Fay (1993) and Easterly and Levine (1997) and many others, I used the ethnolinguistic fractionalization variable. This measure is computed as one minus the Herfindahl index of ethnolinguistic group shares, and reflects the probability that two randomly selected individuals from a population belonged to different race groups. The fractionalization measure adds important information about diversity, concentration, and possible integration among races which can be important to determine coordination and public opinion. The formula is the following and was used to calculate the index for each district:

$$Fractionalization = 1 - \sum_{i=1}^{N} S_{ij}^2 \tag{2}$$

% Bachelors is a variable measuring the percentage of the population with a bachelors degree, *% Bachelors Native* represents the same information but only for native population and *% Bachelors Foreign* for foreign born. It is reasonable to expect education to be an important variable to explain the representative's positions. More educated citizens tend to be more engaged and informed, which we will expect to have an effect on the behavior of their politicians.

The interaction and integration between immigrants and native groups seems like a reasonable characteristic to explain general opinion towards this minority group and subsequently the behavior of their representative. Language is important for integration, cooperation, and communication. The variable *%Low English Foreign* shows the percentage of the foreign population that can't speak English "very well".

Unemployed is the unemployment rate in the district, *Unemployed Native* is the rate for the native population and *Unemployed Foreign* for the foreign born population.

Income represents the income in todays dollars of the median household, meaning that 50% of the population in the district earns less that the specified amount and the other half has income above that. *%Below Poverty* is the percentage of the total population that is below the poverty line.

Income F/N is the proportion that the median foreign born income represents from the median income of the native population.

%Agriculture is the percentage of workers in the district that is employed in the agriculture industry and *%Construction* measures the percentage in the construction industry. Historically, these sector tend to be immigrant

intensive, and by adding this controls is possible to identify districts where immigration could have a bigger impact in the economy, in public opinion, and representative's behavior. *%Foreign Workers* is the proportion of foreign workers of the total workers in the district, this variable tries to incorporate some information related with foreign participation in the economy and labor force.

The native population's opinion towards the immigration community, as discussed before, is important to determine the effect the minority group will have on policy. Public opinion is affected by the perception of fairness in the utilization of public goods, *%Public Assistance* measures the percentage of the population that receives governmental cash assistance.

Table 1: Descriptive Statistics

	Average	Standard Deviation	Maximum	Minimum
DW	0.156	0.547	1.361	-0.751
Political Party	0.51	0.5	1	0
%Foreign Born	12.77	11.15	57.20	0.45
Age	37.129	3.359	51.4	25.7
Age Native	35.523	5.118	51	16.3
Age Foreign	40.093	4.016	55.7	29
%White	73.971	17.563	97.3	12.3
%African American	12.665	14.691	68.6	0.3
%America/Alaska Native	0.782	1.808	23.4	0
% Asian	4.707	6.168	53	0.2
%Native Hawaiian	0.154	0.68	12.2	0
%Other Race	5.225	6.753	58.6	0.1
%Hispanic	15.671	17.401	87.2	0.5
%Bachelors Degree	17.558	5.509	37.7	4.6
%Bachelors Native	18.275	6.054	42.3	5.8
%Bachelors Foreign	15.791	5.491	37.1	2.9
%Low English Foreign	46.952	11.574	80.9	16.1
%Unemployed	5.59	1.691	14.7	2.2
%Unemployed Native	5.689	1.79	15	2.2
%Unemployed Foreign	5.131	1.925	15.5	0
Median Income	53,535	14,148	113,376	23,291
%Below Poverty Line	14.279	5.571	38.8	3.1
% Employed Agriculture	1.903	2.437	28.1	0
% Employed Construction	6.89	1.844	20.9	2
% In Public Assistance	2.34	2.027	11.6	0
Fractionalization	0.316	0.17	0.809	-0.49
% Foreign Workers	14.171	14.902	95.749	0

To account for possible different effects of immigrants on different policy topics, I also use an alternative measure for ideology, a vote rating published in the National Journal that assigns to each member of Congress a number that reflects how liberal his or her voting behavior was compared to other members of Congress. There is such rating for each member in three general classifications: economic, social (abortion rights, gun control, etc.) and foreign policy (war funding, foreign aid). For example, if a member receives a rating of 30 in the economic classification that means that the representative voted more liberally than 30% of the House in economic issues. The rating system used was devised by political analyst Bill Schneider. [1]

Table 1 summarizes the variables used in this paper. All variables except vote ratings by subject, are available from 2007 to 2013, covering four periods of Congress and reaching a number of observations of 1,770. In the case of foreign born data, besides having information from 2007 to 2013, data from 1980 was also found, the data from 1980 will be used to construct an instrumental variable. The vote rating by subject are only available after 2010.

1.5 Endogeneity problem

Regressions using Ordinary Least Squares (OLS) assumes there is no effect of the dependent variable on the independent variables. If the regression analysis is made using the data described earlier, the possibility exists of having biased coefficient results because of the endogeneity problem. The typical regression techniques rely on the assumption that the dependent variable has no effect in the explanatory variable, when this effect exists the results are biased. In this case this problem may be present because it is plausible to think that the ideological positions of the representative affects the decision of immigrants to go to that specific congressional district. It is possible that immigrants go to places where the positions of the representatives are more preferable to them. This situation will cause spurious results.

As in Saiz (2007), Ottaviano and Peri (2007) and Gonzanlez and Ortega (2013), I instrument immigration using historical information on immigrant networks. The idea is that the decision of immigrants about where to go is correlated with previous establishments of immigrants; and because the establishment of immigrants in the past is unlikely to be correlated with actual political and economic conditions we can use this information to construct a variable that avoids the endogeneity problem.

I used data from 1980 Census to extract information about non-U.S. born population in that year. The objective of using 1980 data is to use the information to find immigrants networks that attracted more immigrants afterwards. The period of analyses in this paper is from 2007 to 2013, and

[1] National Journal Vote Ratings, https://ballotpedia.org June, 2016.

I will use the 1980's networks to estimate today's immigrants flow. If there is a two way relation between the representative's ideological position and immigration, using the 1980's networks eliminates the endogeneity problem because their establishment, 37 years ago, is unlikely correlated with actual political and economic conditions.

The constructed instrument by eliminating the impact that the representative's position has on immigration flows it allows to isolate the desired effect of immigration on ideological position in Congress.

The main geographical unit of observation is the congressional district, however the limits and the territory they represent is not fixed in time and are adjusted after each decennial census. The number of districts each state has also changes accordingly to movements in population, and the numbering of districts can also be different after each census. These changes make districts less comparable, specially when you try to match districts separated by several decades.

In order to have more comparable geographical units I use the Census Tracts. These tracts are relatively small areas defined for the purpose of taking the census. According the the census definition: "The primary purpose of census tracts is to provide a stable set of geographic units for the presentation of statistical data" [2]. Tracts change less frequently and less drastically than the districts.

The Census web-page also provides the relationship files which allow to match congressional district across different geographical units like counties or census tracts. I use these files to match each 2007 congressional districts with their respective tracts, and the same exercise was made for the post-2010 districts. Then I used another relationship files from the census to match the district's tracts with the respective 1980's tracts, few changes occurred in this period related with tracts boundaries. The information was in text files and the construction of the database was a labor intensive task. The resultant data allows the identification of 1980's immigration networks in today's congressional districts bounaries .

$$\%Foreign_{i,t} = [\frac{Immigrants_{i,1980}}{Immigrants_{TotalUSi,1980}}] * Immigrants_{TotalU.S.,t} \quad (3)$$

After the matching process was done, I calculated the share of immigrants in each congressional district using the total immigrants in the country in

[2]United States Census Bureau, https://www.census.gov July 2016

1980. For example, using the 2011 boundaries districts, the first district of Alabama has 165 tracts, after matching those tracts with the 1980's tracts I find that the territory that is today the congressional district one had 0.05% of the immigrant population in the country in 1980. Then I used the total population of foreign born in the country in the respective year to estimate the number of immigrants in today's congressional districts, see equation (3). In the Alabama example, I multiplied the 37.9 million foreign-born population in the country in 2011 times the 0.05 to get the estimate immigrant population: 18,950. After estimating the number of immigrants in each congressional district then I computed the share that those immigrants represent with respect to the total population in the respective year. The 18,950 estimated foreign born in the first district of Alabama represent 2.6% of the district's population, these shares were used in the regression analyses.

Table 2: Changes in % Foreign

	Original %Foreign	IV %Foreign
%Foreign	12.77	13.53
Standard Deviation	11.15	17.92

1.6 Results

Table 3 shows the results of the linear regression model using Ordinary Least Squares (OLS) and two stage regression. The regressions were calculated using the program Matlab and the number of observations is 1,770.

The explanatory variable *% Foreign* has a small negative but significant effect among all specifications except when the regions controls are included. For model 1 through 5 the effect is significant to the 99% confidence level, for model 6 is not significant but is still negative. The results suggests that as the proportion of immigrants in a specific congressional district increases the representative will have a slightly more liberal position. The effect seems to be low, a one percent increase of the proportion of foreign born in a district according to model 5 has an average effect on ideological position of -0.015 which translates to a 0.75% more liberal representative, using the DW scale.

The variable *Party*, as expected, has a big and significant effect on ideological position, all six specifications have *Party* having a significant effect at the 99% confidence level . The results suggest that the party affiliation of the representative explains almost all the variance in the first dimension of the DW-NOMINATE scale. The representative's behavior after being elected may be influenced in big part by the national party agenda and the party's objectives.

According to the regression, being a democrat representative reduces the DW nominate score by almost one point in average, which translates to being 50% more liberal. *Party* is, by far, the most important variable to explain voting behavior. This result is in accordance to the findings of polarization in congress among the literature.

%Age has a positive effect and significant in two out of four specifications, the results suggests that as the median age in the district increase the representative tends to be more conservative, the effect appears to be greater than immigration but not as large as Party.

The different ethnic groups have in general positive effects, an increase in the share of Whites, African American, Asian increases the liberal position of the representatives. The fractionalization index that measures ethnic diversity has a significant positive effect among the six specifications. The effect varies from 0.13 to 0.26 which translates into a 6.5% to 13% more conservative representative. Ethnic fractionalization appears to change voting behavior in congress into a more conservative position.

The variables related with education do not show significant effect. *%Bachelors*,*%Bachelors Native*,*%Bachelors Foreign*, *Unemployment*, and *% Low English Foreign* do not show any significant effect.

Table 3: Regression Results. Dependent variable: Ideological Position

	OLS			IV. Second Stage		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	0.716***	-0.273	-0.554	0.714***	-0.696	-0.675
	(0.006)	(0.45)	(0.446)	(0.006)	(0.517)	(0.512)
Political Party	-1.031***	-0.988***	-0.976***	-1.032***	-0.980***	-0.973***
	(0.008)	(0.008)	(0.008)	(0.007)	(0.009)	(0.009)
%Foreign Born	-0.003***	-0.007***	-0.005***	-0.002***	-0.015***	-0.006
	(0)	(0.001)	(0.001)	-0.0003	(0.005)	(0.005)
Age		0.013**	0.008		0.038**	0.013
		(0.006)	(0.006)		(0.016)	(0.016)
Age Native		-0.012***	-0.009*		-0.032**	-0.012
		(0.005)	(0.005)		(0.013)	(0.013)
Age Foreign		-0.003*	0		-0.004**	-0.000
		(0.002)	(0.002)		(0.001)	(0.001)
%White		0.011**	0.013***		0.013***	0.013***
		(0.004)	(0.004)		(0.004)	(0.004)
%African American		0.008*	0.01**		0.010**	0.010**
		(0.004)	(0.004)		(0.004)	(0.004)
%America Native		0.005	0.008		0.006	0.008
		(0.005)	(0.005)		(0.005)	(0.005)
% Asian		0.009**	0.01**		0.013***	0.011**
		(0.005)	(0.005)		(0.005)	(0.005)
%Native Hawaiian		0.02	0.026*		0.020	0.026*
		(0.015)	(0.015)		(0.015)	(0.014)
%Other Race		0.01**	0.014***		0.012**	0.014***
		(0.004)	(0.004)		(0.004)	(0.004)
%Hispanic		0.002**	0		0.002***	0.000
		(0.001)	(0.001)		(0.001)	(0.001)
%Bachelors Degree		0.004	0.005		-0.003	0.003
		(0.005)	(0.005)		(0.007)	(0.007)
%Bachelors Native		-0.002	-0.005		0.007	-0.003
		(0.004)	(0.004)		(0.007)	(0.007)
%Bachelors Foreign		0.001	0		0.001	0.000
		(0.002)	(0.002)		(0.001)	(0.001)
%Low English Foreign		0	0		0.000	-0.000
		(0.001)	(0.001)		(0.000)	(0.000)

Continues in next page

Table 4: Continuation Regression Results.Dependent variable: Ideological Position

	OLS			IV. Second Stage		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
%Unemployed		-0.002	-0.004		-0.002	-0.005
		(0.022)	(0.022)		(0.022)	(0.021)
%Unemployed Native		0	-0.002		0.001	-0.002
		(0.02)	(0.02)		(0.020)	(0.019)
%Unemployed Foreign		0.004	0.004		0.006	0.004
		(0.004)	(0.003)		(0.003)	(0.003)
Median Income		0	0		1.144	9.154
		(0)	(0)		(6.843)	(6.864)
%Below Poverty Line		-0.001	0.001		8.628	0.000
		(0.002)	(0.002)		(0.001)	(0.001)
% Employed Agriculture		0.003	0.001		0.002	0.000
		(0.002)	(0.002)		(0.001)	(0.001)
% Employed Construction		0.011***	0.006**		0.012***	0.005*
		(0.003)	(0.003)		(0.002)	(0.003)
% In Public Assistance		-0.01***	-0.007***		-0.009***	-0.006***
		(0.002)	(0.002)		(0.002)	(0.002)
Fractionalization		0.173***	0.131**		0.261***	0.145*
		(0.057)	(0.058)		(0.078)	(0.079)
% Foreign Workers		-0.001	-0.001**		-0.001**	-0.001*
		(0)	(0)		(0.000)	(0.000)
Income Foreign/Native		-0.104***	-0.104***		-0.103***	-0.104***
		(0.032)	(0.032)		(0.032)	(0.031)
Congress 111th		0.034***	0.039***		0.022	0.037**
		(0.013)	(0.013)		(0.014)	(0.015)
Congress 112th		0.036***	0.037***		0.027**	0.036***
		(0.012)	(0.012)		(0.013)	(0.013)
Congress 113th		0.055***	0.053***		0.039**	0.050***
		(0.014)	(0.014)		(0.016)	(0.016)
Region Midwest			0.069***			0.073***
			(0.013)			(0.012)
Region South			0.096***			0.100***
			(0.014)			(0.013)
Region West			0.07***			0.080***
			(0.015)			(0.014)
N	1,770	1,770	1,770	1,770	1,770	1,770
R2	0.924	0.9321	0.9341	0.9235	0.9314	0.9337
F	10745	795.84	745.35	10670	787.4	740.43

Dependent Variable: DW score(-1 = most liberal; 1= most conservative).
Standard errors in parenthesis. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

The percentage of the total workers employed in the construction industry, *%Construction*, has a positive and significant effect. Districts that are more Construction intensive have on average more conservative representatives. The construction industry tends to attract more immigrant populations, is possible that native workers that live in these district have less positive opinions towards immigration. These opinions are reflected on their representative behavior by making him or her vote in a more conservative way.

% In Public Assistance, which measures population under government assistance programs has a small negative but significant effect. If a higher share of the population is under assistance government the representative will be on average more liberal.

Income F/N, which measures differences in income among native and foreign populations is also a big negative and significant effect. As foreign workers increase their income compared with native workers the representatives in the districts will be on average more liberal.

Finally, when dummy variables for controlling the number of Congress and regions are added, these variables are significant and important for explaining the independent variable.

There exists the possibility that immigrants have an effect on ideological position through the determination of the party in power. To test this hypothesis I ran a regression where the dependent variable was *Party* and the share of immigrants in the district as the explanatory variable. Table 1.6 shows the results. The R square is very low showing the little explanatory power of the share of immigrants in each district. However the effect is significant and the coefficient suggests that as the share of immigrants increases it is less likely that the district has a Democratic representative. This can be caused because an inflow of non-U.S. may change the political preferences of natives, making them to vote for more conservative representatives that are more likely to vote against immigration.

Table 5: Regression results. Dependent variable: Party(Democrat)

Variable	**Coefficient**	**t-stat**
Intercept	0.65	20.29
Share Immigrants*	-0.02	-5.07
R square	0.06	
Observations	435	

Table 6: 1st Stage Regression Results. Dependent variable: % Foreign

	Model 1	
Constant	-35.341***	(7.151)
Political Party	0.072***	(0.007)
%Foreign New	0.559***	(0.137)
Age	2.847***	(0.062)
Age Native	-2.34***	(0.05)
Age Foreign	-0.281***	(0.025)
%White	0.211***	(0.071)
%African American	0.172**	(0.069)
%America and Alaska Native	0.059	(0.087)
% Asian	0.396***	(0.073)
%Native Hawaiian	-0.123	(0.238)
%Other Race	0.063	(0.073)
%Hispanic	0.212***	(0.013)
%Bachelors Degree	-0.571***	(0.088)
%Bachelors Native	0.839***	(0.07)
%Bachelors Foreign	0.018	(0.024)
%Low English Foreign	-0.023**	(0.011)
%Unemployed	0.484	(0.354)
%Unemployed Native	-0.061	(0.321)
%Unemployed Foreign	0.153***	(0.056)
Median Income	0	(0)
%Below Poverty Line	0.024	(0.031)
% Employed Agriculture	0.119***	(0.032)
% Employed Construction	0.432***	(0.047)
% In Public Assistance	0.024	(0.036)
Fractionalization	11.013***	(0.906)
% Foreign Workers	-0.066***	(0.007)
Income Foreign/Native	0.163	(0.514)
Congress 111th	-1.79***	(0.211)
Congress 112th	-1.281***	(0.191)
Congress 113th	-1.878***	(0.225)
Region Midwest	-0.974***	(0.209)
Region South	-1.272***	(0.224)
Region West	-2.208***	(0.233)
N	1,770	
R2	0.924	
F	10745	

Dependent Variable: % Foreign Born in each Congressional District
Standard errors in parenthesis. Significance levels: * p¡0.1, ** p¡0.05, *** p¡0.01.

Table 6 shows the results of the first stage regression from specifications four, five, and six. The dependent variable is the original %Foreign used in the OLS regressions, and the explanatory variables are the control variables including the %Foreign New which are the new percentages calculated using the 1980 immigration networks. The 1st stage regression shows a high R2, a high F which validates the overall significance of the model and the new estimate of foreign born using networks is significant to explain the original data at the 99% confidence interval. The results suggest that the instrument is valid.

1.7 Results by topic

Non-U.S. citizens may be more engaged in specific political topics and it is possible that the effects they have on ideological position depends on which type of policy is being voted. To analyze this possibility I use the National Journal House Ratings as the dependent variable and the same set of explanatory variables. The ratings rank the representatives in a liberal measure, a higher number means the representative is more liberal when compared with other representatives in Congress. For this analysis there were only two years of ratings, 2011 and 2013, which covers the 112th and 113th Congress. The number of observations is 867.

The results of the second stage regression are shown in table 7. The coefficients of the share of immigrants denominated %Foreign Born has a negative effect and is significant when all controls are included. On average, an increase in the immigrant population will reduce the representative's liberal ranking in social issues, which include: abortion rights, gun control, etc. When considering economic issues the effect from immigration is positive and significant, more foreign born appears to have more liberal voting behavior related with economic issues. Foreign born seems to make representatives to behave more liberal when voting for foreign related policy: war funding, foreign aid, etc.

The results suggests that as immigrants represent a higher share of the total population the representative of that Congressional district is more likely to be in a higher percentile of the liberal measure, this is true in foreign and economic issues. The social related voting behavior shows a different result, more immigrants translate into more conservative representatives, this is interesting and it was not expected.The party is still by far the most important variable to explain political position with democrats being more liberal independently of the policy topic.

The results by subject are in accordance with the results from the previous section in the case of economic and foreign issues. Immigrants seem to have an effect on the voting behavior of their representatives, specifically by voting in a more liberal way.

Table 7: Regression Results. Dependent variable: Liberal Score by Subject:

	Social		Economic		Foreign	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	-32.821	-8.93	60.178	86.259	33.727	34.707
	(69.852)	(69.087)	(66.881)	(66.932)	(69.754)	(70.204)
Political Party	45.276***	43.404***	44.005***	42.917	44.493	44.212***
	(1.391)	(1.404)	(1.332)	(1.36)	(1.389)	(1.426)
%Foreign Born	-0.046	-0.134**	0.148**	0.088*	0.093*	0.068
	(0.066)	(0.067)	(0.064)	(0.065)	(0.066)	(0.069)
Age	0.542	0.608	0.532	0.457	0.195	0.276
	(0.635)	(0.628)	(0.608)	(0.609)	(0.634)	(0.639)
Age Native	-0.258	-0.107	0.163	0.267	0.177	0.191
	(0.525)	(0.516)	(0.503)	(0.5)	(0.524)	(0.524)
Age Foreign	-0.152	-0.267	-0.283	-0.358	-0.08	-0.038
	(0.269)	(0.266)	(0.258)	(0.258)	(0.269)	(0.27)
%White	0.5	0.259	-0.548	-0.753	-0.146	-0.229
	(0.69)	(0.681)	(0.66)	(0.659)	(0.689)	(0.692)
%African American	0.764	0.649	-0.221	-0.378	0.028	-0.051
	(0.664)	(0.657)	(0.636)	(0.637)	(0.663)	(0.668)
%American Native	1.364	0.897	0.035	-0.241	-0.157	-0.37
	(0.836)	(0.826)	(0.8)	(0.801)	(0.835)	(0.84)
% Asian	0.935	0.688	-0.421	-0.601	0.248	0.091
	(0.709)	(0.699)	(0.679)	(0.678)	(0.708)	(0.711)
%Native Hawaiian	1.97	1.066	-0.607	-1.143	-0.66	-0.977
	(2.229)	(2.194)	(2.135)	(2.126)	(2.226)	(2.23)
%Other Race	0.765	0.295	-0.493	-0.833	-0.052	-0.236
	(0.713)	(0.708)	(0.683)	(0.686)	(0.712)	(0.72)
%Hispanic	-0.075	0.06	0.06	0.12	-0.007	0.006
	(0.128)	(0.129)	(0.123)	(0.125)	(0.128)	(0.131)
%Bachelors Degree	-2.068**	-2.809***	-1.155	-1.545	-0.07	-0.25
	(0.965)	(0.958)	(0.924)	(0.929)	(0.964)	(0.974)
%Bachelors Native	1.786**	2.477***	0.758	1.176	0.224	0.382
	(0.764)	(0.759)	(0.732)	(0.735)	(0.763)	(0.771)
%Bachelors Foreign	0.403	0.676**	0.266	0.395	-0.086	-0.072
	(0.311)	(0.31)	(0.298)	(0.3)	(0.311)	(0.315)
%Low English Foreign	-0.129	-0.031	-0.16	-0.089	-0.238	-0.226*
	(0.122)	(0.122)	(0.117)	(0.118)	(0.122)	(0.124)

Continues in next page

Table 8: Continuation Regression Results. Dependent variable: Liberal Score by Subject

	Social		Economic		Foreign	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
%Unemployed	0.309	-1.276	-5.628	-6.671	-3.081	-5.15
	(3.932)	(3.938)	(3.765)	(3.815)	(3.926)	(4.002)
%Unemployed Native	-0.061	0.605	3.948	5.002	3.29	5.007
	(3.529)	(3.543)	(3.379)	(3.433)	(3.524)	(3.6)
%Unemployed Foreign	0.35	0.637	1.617**	1.744	0.607	0.953
	(0.745)	(0.744)	(0.713)	(0.721)	(0.744)	(0.756)
Median Income	0	0	0	0***	0***	0
	(0)	(0)	(0)	(0)	(0)	(0)
%Below Poverty Line	-0.032	-0.003	0.023	-0.03	0.109	0.292
	(0.297)	(0.31)	(0.284)	(0.3)	(0.296)	(0.315)
% Employed Agriculture	-0.086	-0.145	-0.055	-0.034	0.099	0.054
	(0.295)	(0.301)	(0.283)	(0.291)	(0.295)	(0.306)
% Employed Construction	-0.801	-0.414	-0.752	-0.773	-0.225	-0.456
	(0.488)	(0.52)	(0.467)	(0.504)	(0.487)	(0.528)
% In Public Assistance	1.31***	0.786*	1.31***	0.971	1.021	0.894**
	(0.402)	(0.404)	(0.385)	(0.392)	(0.401)	(0.411)
Fractionalization	-22.46**	-18.046**	-13.389	-13.049	-12.505	-11.447
	(8.844)	(8.874)	(8.468)	(8.598)	(8.832)	(9.018)
% Foreign Workers	-0.103	-0.1	-0.039	-0.006*	-0.094*	-0.08
	(0.083)	(0.083)	(0.079)	(0.081)	(0.082)	(0.085)
Income Foreign/Native	6.567	6.652	8.152	7.594	-2.247	-2.504
	(6.229)	(6.127)	(5.964)	(5.936)	(6.22)	(6.226)
Congress 111th		0.155		-0.818		-2.545**
		(1.183)		(1.146)		(1.202)
Region Midwest		-5.557***		-6.136		-0.418
		(1.988)		(1.926)		(2.02)
Region South		-12.034***		-7.293		-0.704
		(2.127)		(2.061)		(2.162)
Region West		-2.113		-3.101		1.903
		(2.224)		(2.154)		(2.26)
N	867	867	867	867	867	867
R2	0.7387	0.7498	0.7538	0.7586	0.737	0.7392
F	87.86	80.7	95.14	84.63	87.07	76.32

Dependent Variable: Liberal Score, higher number means more liberal ranking in Congress
Standard errors in parenthesis. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

1.8 Conclusion

The data shows a small significant relation between the population of immigrants and the general voting behavior of Congress, more immigrant intensive populations have the effect of moving the representative to a more liberal position, except when we consider social policy the effect is reversed to a more conservative behavior. However, party affiliation is by far the most important variable to explain voting behavior.

Non-U.S. citizens can't directly express their preferences by voting. The channels that they can use to make an impact on policy are more limited, they can join an interest group, support campaigns and assist to manifestations, they can also contribute to national debates, the data shows that the effect they have is stronger in specific issues. There is evidence that suggests that they also have an impact in which party wins in their district. These results also show how non-voters population can have an influence in policy decision of the country, but importantly it helps to raise the issue related to the representative's lack of connection with their population.

Table 9: Regression Results. Dependent variable: Liberal Score by Subject. No controls

	Social	Economic	Foreign
	Model 1	Model 2	Model 3
Constant	22.965***	25.108***	24.138***
	(0.748)	(0.717)	(0.735)
Political Party	47.408***	46.123***	46.657***
	(1.225)	(1.174)	(1.204)
%Foreign Born	0.15***	0.205***	0.187***
	(0.035)	(0.033)	(0.034)
N	867	867	867
R2	0.717	0.7329	0.7243
F	1094.5	1185.1	1135

Dependent Variable: Liberal Score, higher number means more liberal ranking in Congress
Standard errors in parenthesis. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

2 Who Drives the Market? Market Structure and Network Effects with Type-dependent Utility

3 Abstract

This research paper contributes to the literature on market dominance on network goods. I find that when agents care about the characteristics of the members, asymmetric equilibriums can be found where several platforms survive. I also investigate which platforms survive. While early adopters generally determine the characteristics of the surviving network, under some conditions it is the late adopters' preferred platform that becomes dominant.

3.1 Introduction

Market dominance is closely related with total surplus. The number of firms surviving in one market relates with the options available to the consumer and with the price strategy of the participants. A better understanding of the mechanisms that determine the structure of the market allows us to find strategies that will improve social benefit. Given the tendency of high concentration on markets with network goods, the study of the market dominance in these markets becomes even more relevant.

Markets with network effects are those where having multiple individuals buying the same good has a positive impact on the personal utility of the agent. Examples range from telephones to social networks. Michael Katz and Carl Shapiro define network effects as follows: The utility that a given user derives from the good depends upon the number of other users who are in the same network as he or she (Katz and Shapiro 1985). We can identify direct and indirect network effects. The first one arises when by employing the same software or using the same standard users can communicate with each other more easily. The indirect effects arise when the generalized use of a good generates a broader range of associated goods and services, which enhance the utility of the basic good.

In the literature of network effects, the linear network effect is one assumption which is common across models. The linear effect assumption says that the function used to describe the network effects is linear with respect to the number of members in the network. The linearity implies that the added benefit to the personal utility made by the second or third adopter is the same as the contribution made by the n^{th} adopter. In some cases the linear assumption is dropped and the models allow the network effects to be concave (Katz and Shapiro, 1985). Concavity means the added value is high in the case of early adopters and low when subsequent adopters are incorporated. In both cases, in the linear and the concave assumption, it is implicitly assumed that the benefit that a user derives from the network depends only on the number of members. This paper challenges this assumptions.

The functional form of the network effect is an important assumption that plays a key role in the results of the model. Linearity implies that in order to maximize utility all members should join one network. As a result, models with linear effects predict a single dominant firm in the market. This result however, fails to match reality: in most cases where network effects are present we find multiple firms surviving even in the presence of a dominant firm. An example is the software market. There are products like MacOs and Linux with a low but stable market participation participation, despite the high share of the market that Windows has. Another consequence of linearity is that early adopters determine the dominant firm.

In this paper, I propose a modified Hotelling model where agents care about the characteristics of other members in the network. There is heterogeneity in the adoption behavior and there is heterogeneity in the preferences of the members. In contrast with the literature, the functional form of the network effect in this paper does not depend only on the number of agents in the network. The characteristics of the other members are important, and the similarity with other members is the key source of utility. The motivation of this assumption is to adopt a real world attribute. Agents in a network usually pay attention to characteristics of other members. They usually prefer to be with other agents that are similar to them.

The assumption of different adoption behaviors is also an important modification that deviates from the literature. The different adoption rates incorporates the observation that different groups of agents enter the market at different rates. In the real world, some groups are more likely to early adopt a new technology while other groups may wait until the market is almost saturated.

I find that market concentration decreases when agents care more about the characteristics of other members. The proposed model suggests that formation of clusters in a non-dominant network is possible even in the presence of strong network effects. Under some conditions, agents that share some characteristics can choose a non-dominant option and form a cluster. The result is a market with several firms sharing the market. The necessary conditions to allow cluster formation in a network market are type-dependent valuation, heterogeneity of preferences and a high number of options available. The result helps to explain the fact that several small platforms can survive even in the presence of a dominant firm.

The model also challenges the classical result that early adopters usually determine which firm is the dominant one in a market with network effects. While early adopters generally determine the characteristics of the surviving network, under some conditions it is the late adopters' preferred platform that becomes dominant. If agents care only about the number of members on a network, they choose the option with the highest number of members. The network with most members will be the one chosen by the first adopters. When the attributes of the members are important, a high number of late adopters can generate a cluster and decide the dominant firm .

The results are relevant to policy because they highlight the importance of the functional form of the network effects. The results also show the relevance of the initial population shares of the different types, the adoption rates, and the heterogeneity of preferences. The different adoption behaviors are also key to understand markets with network effects because they can drive the results in an important way.

3.2 Literature Review

Katz and Shapiro's (1985) paper is one of the first works written about markets with network effects. They used a simple Cournot model where firms offer homogeneous products and choose output levels. Consumers are assumed to be homogeneous in their valuation of the network externality. The network effect function, which measures the level of utility derived from others, is modeled to be increasing but with decreasing marginal benefits. The network effect is concave, and only the number of members is important to personal utility, the network structure is irrelevant. The work of Katz and Shapiro is important because they were one of the first to model a network effect. Their model is relevant also because is a good example of a model were only the number of members is important. The model suggests two possible equilibriums, symmetric oligopoly and an asymmetric oligopoly. However the last possibility is only mentioned but not fully analyzed.

Farrell and Saloner (1985) also analyze the standardization problem where network effects are present. They use a model with sequential decisions. It is assumed that the network effect is the same across firms and only depends on the number of members, the structure of the network is irrelevant. When they allow complete information, the unique perfect equilibrium involves all firms joining one standard. The linearity of the network effect causes the emergence of a unique dominant firm.

Arthur (1989) proposes a simple model with two new technologies that compete for adoption. An agent comes into the market and decides at time t to buy from firm A or B. There are different preferences but similar individual internalization of network effects. An important characteristic of Arthurs model is that he first considers three possible cases for the network effects. He allows for increasing, constant and decreasing returns, all three cases are linear and only depend on size. Under constant and decreasing returns, the market is shared. He finds that in the increasing case the adoption process becomes a random walk with absorbing barriers, meaning that after some threshold is achieved one standard will dominate the market. He then allows for non-linear effects and derives an interesting result: dominance by a single technology is no longer inevitable if the improvement function is bounded. In the bounded case, some historic events dynamically lead to a shared market, other events may lead to dominance.

Arthur's result is similar to the findings presented in the current paper: when the effect function is non-linear, the market will probably be shared. The model presented in this paper contributes to the literature because it goes beyond the non-linear function by modeling a type-dependent valuation. Arthurs results also highlight the relevance of the network effects assumption on the final results.

The most recent work has not payed much attention to the functional form of network effects. More recent work takes more focus on dynamic settings. In Cabral (2007) consumers die with a constant hazard rate and are replaced by new consumers, with network effects being weakly increasing. Le Guel (2011) takes Arthurs model and relaxes the assumption of exclusivity. Markovich and Moenis (2008) analyze competition dynamics in the presence of indirect network effects when firms also invest in quality improvements, finding that the network externalities can cause a positive effect in competition. Jullien (2001) allows for differentiated products and price discrimination.

The evolution of the network effects and the way they are incorporated in the framework varies across authors. In some cases like Farrell and Saloner (1985) there is heterogeneity in the value consumers have about the network. In Katz and Shapiro (1985) and Church and Gandal (1992) the effects are positive but the marginal benefit is decreasing. Arthur (1989) considers three linear cases and one with bounded effects. As the results show, the way the effects are incorporated in the models has a transcendental impact in the nature of the results.

Given the relevance of the assumption, the study in more depth of the functional form of network effects becomes very important. Swan (2002) highlights that the linear effect assumption is a very strong one, he finds that two conditions must hold to the linear effects to be present.

In the first condition the subscriber of a network must be equally likely to benefit from all members. A telephone network is a good example; the first condition poses that an individual is equally likely to call all the people that own a telephone. This conditions is implausible. In the telephone case, it is more reasonable to believe that a person calls more often to certain subset of the members to whom he shares some characteristics with, like tastes or subjects to talk about.

The second condition that must hold to have linearity is that all types of consumers must have the same diffusion function. This means they must be equally likely to join the market in any given time. Again, this condition seems unrealistic, because different characteristics of the agents can influence the time at which the buy a product or join a network. Engineers, for example, are probably more likely to adopt a technology before designers.

Swan (2002) finds that under the more relaxed assumptions, the individual utility takes the form of an s-shaped curve for the median adopter when plotted against network size. One of these conditions requires that the benefit of a new member to a specific type of member declines the greater the differences between them. Another condition requires that the adoption rates take the form of a logistic curve. With these assumptions, the median adopter network utility function has rapid increases in benefits around the 50% of the extent of the network, just in the time when agents close to the median will most likely start to join the network. He derives little benefit from first and late adopter because few characteristics are shared. This causes the functional form of the network effect to take the s-shaped form. These assumptions seem a more realistic approximation of reality in many markets where not only the size of the network matters but also the composition of it.

The model I developed incorporates the more realistic assumptions that Swan uses. The objective is to analyze the consequences on market structure.

3.3 Model

The operating system (OS) market will be used as an illustrative example of an application of the proposed model. This market has important characteristics which are represented in the model. The most relevant are high switching costs and important network effects through file sharing, software compatibility, and complementary services. The concept of network is used as the group of users using the same OS.

There are five strategies, and each strategy can be interpreted as a choice of network, platform or technlogy. In the OS example it is interpreted as a choice of one operative system, which it will be represented by the index i. The finite strategy set is defined then as:

$$i = \{1, 2, 3, 4, 5\} \tag{4}$$

There are N players, each of which belongs to one of five populations or types, which is represented by the index j:

$$j = \{1, 2, 3, 4, 5\} \tag{5}$$

Each type has a natural preference for a specific network and incurs a cost when choosing a non-preferred one. This can be interpreted as a transportation cost (TC) or as the disutiliy from not joining the ideal network or buying the preferred OS. The TC increases in a quadratic way as the distance traveled increases. For simplicity, it is assumed that a type one agent prefers to join network number one, a type two agent prefers network number two and so on. The transportation cost is minimized at the ideal choice, when i=j. The TC, which depends on the type and the network chosen, is defined by:

$$TC(i, j) = -(i - j)^2 \tag{6}$$

The different types of agents represent the heterogeneity of agents in the real word. One interpretation is that j represents different groups of users that share some general characteristics, such as their ideal choice of OS and software preferences.

In the OS hypothetical example, the different types of agents are represented by engineers, designers, and students. The idea is that engineers in general share some common preferences because of their similar profile. It is the same idea with designers and students. The model has five types of agents and five platforms, and these number are important to the development of the model. However, the illustrative example only considers 3; the other two types and platforms are still there and play an important role in the dynamics, but they are ignored for convenience of the illustrative explanation.

Continuing with the example, students prefer Windows because of their friendly user interface, but Mac OS is ideal for designers because of the software options. Linux is best suited for engineers. The TC is then the cost that an engineer will incur if he or she buys Windows instead of Linux, the ideal choice for an agent of his or her characteristics. The TC increases quadratically as the elected OS differs from the ideal choice.

A consumer that joins a network also derives utility from the agents that are already members of that specific network. The size of the network has a positive effect on the individual utility. However, each type prefers to be with agents of the same type and the utility derived from not-similar types decreases exponentially as the distance between them increases. These differences in the utility derived from different types of agents is expressed by the function:

$$g(\alpha, j) = e^{-(j-\alpha)^2 \frac{1}{h}} \tag{7}$$

which represents the utility that a type α derives from joining a network with one type j already in it. The form of the function $g(\alpha,j)$ is similar to the normal distribution plot, so if the index j would be continuous, then the function would take the bell shaped form characteristic of the normal distribution with the highest value when $j=\alpha$. This means that a type α derives the highest benefit from another α member and the benefit decreases as the new member differs from him.

The total utility of a type α of joining network i at time t is given by:

$$U(\alpha, i, t) = -(i-\alpha)^2 + \beta \sum_{j=1}^{5} [e^{[-(j-\alpha)^2 \frac{1}{h}]} * S_{i,j,t}] \tag{8}$$

Where $S_{i,j,t}$ is an indicator of the number of agents of type j that are members of network i at time t. β is the parameter that measures how important the network is to the total utility of the player.

The parameter h determines how fast the utility decreases as the distance among types increases. For low values of h the gain of a type α joining a network with one member type $i=\alpha+1$ is very low compared to the utility when $i=\alpha$. The agent highly values being with similar types. This type of valuation will be called: Type-dependent utility". For very high values of h the utility is the same across i, agents derive the same utility from new members regardless of their type. This is similar to the linear network effect function that is assumed in most models where only the number of members matters and not their characteristics. This form of valuation will be called: Linear utility".

In the OS example, under the Type-dependent utility" designers derive a higher utility when they join a network with a high number of designers. And they derive a lower utility from a network with a high number of students. The concept is that when a designer buys an OS and joins a network with a lot of colleagues, he will easily find software developed and adapted to the profession. He will also find designer-specific complementary services, and he can learn, from more experienced designers, how to use the OS in an optimal and designer-oriented way. If the designer joins a network with many students in it, he will still derive some utility. He can share knowledge with students, but they have different characteristics and needs, so it is very unlikely that he can find designer-oriented services. The tutorials and software will be developed for students and not suited for designers. In the

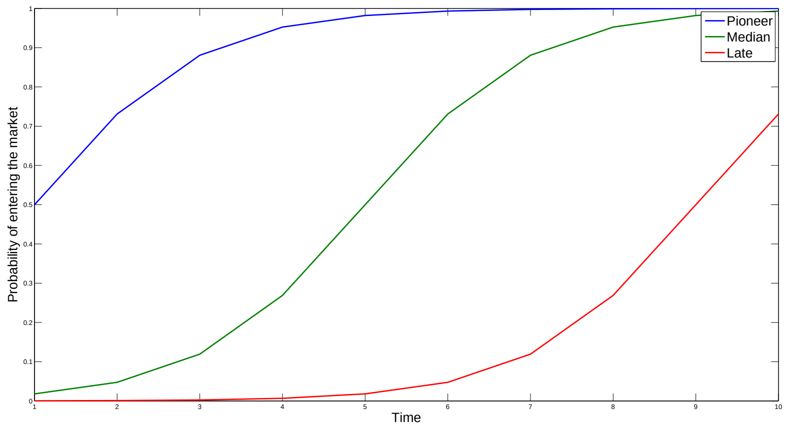

Figure 2: Probability of entering the market of different types

Linear utility" case the designer does not care about the composition of the network. He only cares about the number of agents in it. In this case joining a network with ten designers or another with ten students would derive him the same utility.

At a given time t, agents choose the network that maximizes their utility using all information of the compositions from the five established networks in the previous period *t-1*.In the model, agents are myopic, they are not forward looking agents and do not predict the future evolution of the game. Players have only one opportunity to make a decision, the reason to have this assumption is to represent markets with high switching costs. The time at which they make their choice depends on their type, the intention of the assumption is to represent the heterogeneity of adoption rates on the population. Type 1 players are more likely to enter the market, and make a choice, at the initial stages of the game. I will refer to this type as Pioneers". If they didn't enter the market at a specific time t the probability of entering increases as time passes. Type 5 players are the Late adopters" and the probability to enter the market for them is really low at the initial stages but high at the final stages. The adoption behavior of agents of type 2, 3, and 4 are in between the pioneers and late adopters. The probability of a type j to enter the market a time t is resented by the following function:

$$P(j,t) = \frac{1}{1 + e^{-(t-j)}} \tag{9}$$

The way *P(i,j)* evolves is shown in figure 2. Notice that the probability is always increasing. By the last period it is almost sure that all players have entered the market. Pioneers are more likely to enter in the initial stages, but when time reaches the mid stage the median adopters start to enter the market rapidly. The same happens to the late adopters in the final stages of the game. The different adoption behaviors in the model is another way to represent the heterogeneity of agents in reality. Usually different groups adopt at different rates; for example, is more likely that engineers adopt a new technology earlier than students.

In the OS example, I will assume engineers have a higher probability of adopting an OS in the early stages of the game, making them the pioneers. They have abilities that allow them to adopt an OS that hasn't been used before, and their willingness to adopt a new technology is probably higher than those in the other groups. Designers are the median adopters and students the late adopters.

Finally, let $d_{j,t}$ be a variable that takes the value of one if an agent of type j enters the market at time t and 0 otherwise. Then, $d_{j,t}$ is a random variable that follows a Bernoulli distribution with probability of success equal to *P(j,t)*. After they decide" to enter the market players choose the network that maximizes their utility given previous information. The maximization problem of a type α that enters at time *t+1* is:

$$\underset{i}{\text{Maximize}} \quad U(\alpha, i, t) = -(i-\alpha)^2 + \beta \sum_{j=1}^{5} [e^{[-(j-\alpha)^2 \frac{1}{h}]} * S_{i,j,t}] \tag{10}$$

The timing of the game is resumed in the following table:

Timing	
t=1	Realization of $d_{j,1}$ for all agents, those who enter choose network
t=2	Realization of $d_{j,2}$ for those that $d_{j,1}$=0. Those who enter choose network based on network composition in *t=1*
t=n	Realization of $d_{j,n}$ for those that $d_{j,n-1}$=0. Those who enter choose network based on network composition in *t=n-1*

Given the characteristics of the model and the functional form of the network effects, the initial distribution of types will play an important role in the determination of market structure. Three initial distributions are considered, the first one with a high share of pioneers agents, the second one with high share of type three, and in the third distribution there is a high share of late adopters.

3.4 Results

The model was simulated 1,000 times for each set of parameters values. The software used to simulate was Matlab. The number of players was set to *N=1000* and the number of periods to *T=20*. Three different values of β were considered. β*=0* represents the limiting case where network effects are not present, only the intrinsic value of the product is important. When β*=0.5* network effects are medium and when β*=1* the effects are strong and the network is very valuable in the derivation of agent's utility.

When network effects are positive I consider two values for the parameter *h*. When *h=1* the agent cares a lot about the composition of the network. The utility derived from joining a platform with different types is very low, but joining a network with a lot of same types is high. When *h=1*, the type-dependent utility case is being considered. Using the OS example, engineers highly value to be in a network with their colleagues and derive low utility from designers or students.

When *h=1000*, the player derives virtually the same utility from all types. The player does not care about the composition, and only the number of adopters matters because everyone is equally important. When *h=1000*, the classic case of Linear utility" is bing considered. Engineers do not care about the profession or profile of other members, they only care about network size.

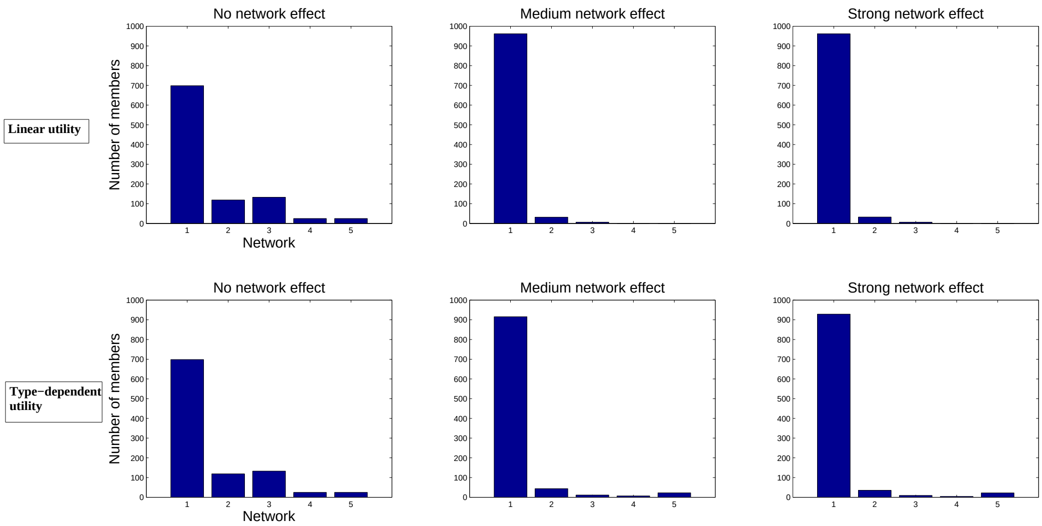

Figure 3: Average final structures. Initial distribution: High share of pioneers.

Market structure, in this case, consists in the distribution of agents among all networks. A concentrated market has a high share of agents in one or a few groups, this is the market structure that is common in the typical models. A low concentration level means that agents are more or less equally distributed. Each simulation ended after the last agent chose a network. For each set of parameters and after each simulation one market structure was computed. Figures 3, 4 and 5 illustrate the average final market structure of the simulations, that is, the average number of agents in each firm after the 1,000 repetitions.

Figure 3 illustrates the case when the pioneers represent a high share of the population. When network effects are positive, it doesn't matter if agents only consider the size of a platform or if they care about the composition, the final structure of the market always has Network 1 as the dominant.

The previous result occurs because the pioneers are the first to enter the market and they choose they preferred network. A network effect is created that makes subsequent adopters find more profitable to join the pioneers ideal choice.

The next table shows recapitulated information of the OS example used to illustrate the model:

Example. The OS market		
Type of agent	*Adoption rate*	*Ideal choice of OS*
Engineer	Pioneer	Linux
Designer	Median adopter	Mac OS
Student	Late adopter	Windows

When engineers (pioneers) represent a high share of the population, the model suggests they will enter the market first and they will adopt OS 1 (Linux), which is their ideal choice. When the low number of designers and students enter the market they also choose Linux because of the high network effects induced by engineers. It doesn't matter if students and designers highly value to be with their peers, the utility derived from the high number of pioneers is large enough to make Linux the best option.

Proposition 1: *When the share of pioneers is high, the final structure will have a single dominant firm. This applies for the linear and for the type-dependent utility case.*

Figure 4 represents the case when the share of median adopters is high. In the case of linear utility, the final average distribution has one platform as the dominant. The reasoning is similar than before, the pioneers create a network effect that makes median and late adopters choose the dominant option. However, when individuals care about the composition of the platforms, the final structure of the market is a duopoly. Network 1 has a high share of the market because, as before, the initial adopters create a network effect for subsequent adopters. However, as the high number of type three agents enter the market they form a cluster in network three. The previous result occurs because they prefer that platform and because they highly value being with similar agents. The high number of median adopters also allows them to create a strong enough network effect to make this option payoff dominant. The cluster formation process allows the concentration in the type-dependent case to be considerable low when compared to the linear case.

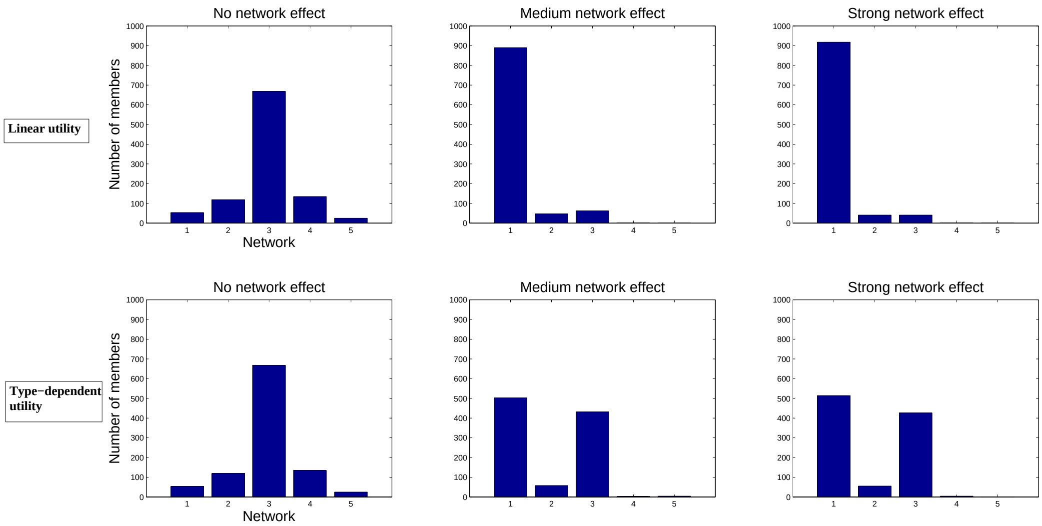

Figure 4: Average final structures. Initial distribution: High share of median adopters

Proposition 2: *When the initial distribution has a high share of median adopters, the final market in the type-dependent utility will have a lower concentration when compared to the linear utility case.*

Figure 5 represents the case when there is a high share of late adopters. In the case of linear utility, Network 1 one dominates the market because the effect created by the pioneers, the high share of late adopters does not have an effect. When the type-dependent utility is considered, the final structure is completely different. Platform 5 gets the highest share because the high number of type five agents allows them to form a cluster which maximize their utility, despite the initial effect crated by the pioneers in platform 1. An important difference is that the pioneers role is diminished and they don't determine the dominant firm in the final structure, which did occur in the linear case despite the low number of them.

In the OS example, when the population has a high number of students (late adopters) and only the sizes of the network matters, students will choose option number one (Linux) because of the effect created by the engineers. When composition is important, the process is different. By the time students adopt an OS, engineers and designers will already have their respective networks, but because the non-students members are low in number and students value poorly to be with them, they can start a new cluster in their ideal choice (Windows). It doesn't matter if engineers choose first, students will form a cluster in their ideal option making Windows the dominant OS. This result shows that under some conditions, even in the presence of network effects, the first adopters may not determine the dominant standard. I will return to this result later.

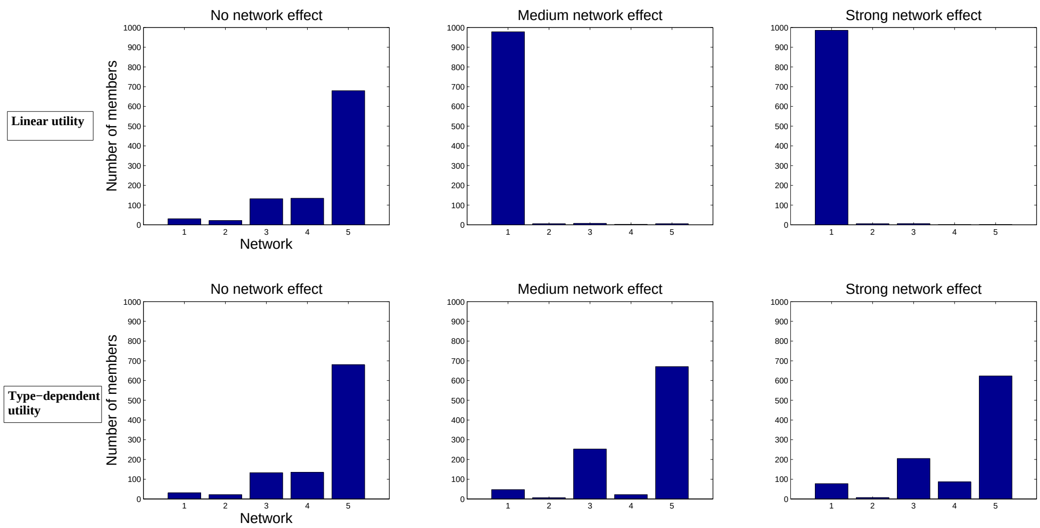

Figure 5: Average final structures. Initial distribution: High share of late adopters

Proposition 3: *When the share of late adopters is high, the final market structure in the type-dependent utility will have a lower concentration when compared to the linear utility case.*

The final structure can vary significantly across simulations. The range of final structures of the simulations is broad because the different realizations of the probability of entering the market. The different realizations cause different decisions to be made at different times and consequently different final market structures to result. Previous figures show the average final structure and do not show all possible final compositions.

The variation in results is specially true when there is a high share of late adopters. Figure 6 illustrates the highest and lowest market concentration in the case of high share of late adopters. In the maximum concentration case, firm number four has more than 80% of the market and with only two other firms surviving but with marginal participation. It is interesting that OS 5 has zero members. In the minimum concentration case, the highest share is near 40%, and all five firms have adopters. In the type-dependent case, the variance of results increases as the share of late adopters increases.

The reason there is a higher variation in market concentration when there is a high number of late adopters is because of the different realizations of entering the market. In the first stage of the game, the probability a pioneer enters the market is high (50%), and increases fast as time goes by. Because of this, all pioneers enter the market in a short period of time. Late adopters, in the other hand, start the first stage with a really low probability (0.0001%), which increases gradually at first and rapidly in later stages. This causes the general adoption rate of late adopters to vary in each simulation. In the case the realizations of these probabilities make late adopters enter in the very last stages, there will not be enough momentum created to form a new cluster in OS 5. This is the reason why in some cases, as seen in figure 6, OS 5 doesn't have any members.

In order to be enough momentum to let students form a cluster in Windows, it is necessary that some students enter the market in the early or mid stages of the game, when the network effects of engineers or designers are not high enough. This will allow subsequent students to join the first students adopters in the Windows network and form a cluster with them.

Proposition 4: *In the type-dependent utility case, as the share of late adopters increases the uncertainty of the final structure increases.*

Pioneers have an important role in the determination of the dominant firm in the final market structure. Their early entrance to the market creates a network effect that makes posterior adopters to value higher the pioneers' ideal choice.

In the linear case, pioneers will always determine the dominant firm. This is true even in the case when they represent a low share of the population. In the type dependent case, pioneers are important in two cases. The first

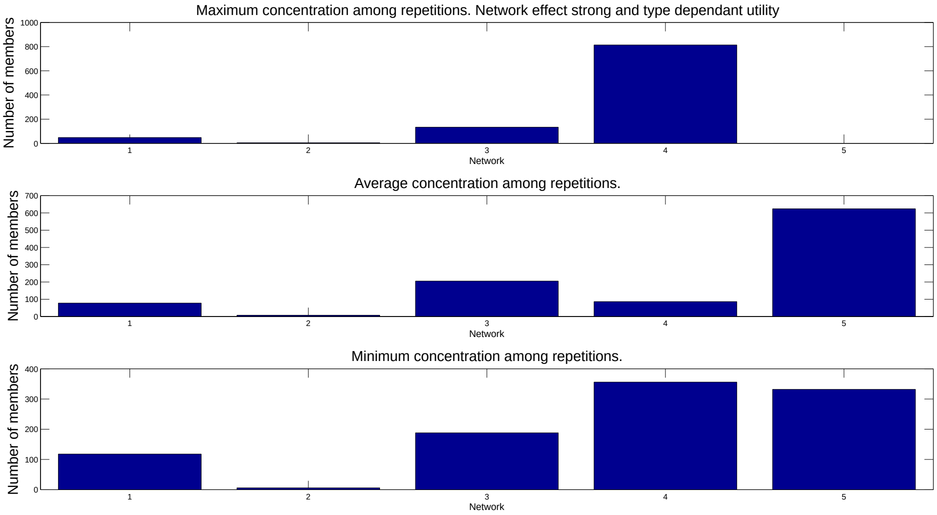

Figure 6: Range of market concentrations. Type-dependent utility. Initial distribution: High share of late adopters

one is when they represent the highest share, in this case they determine the unique dominant firm. The second case is when median adopters represent the highest share, in this case pioneers choose one of the two dominant firms. There is one situation, in the type-dependent case, when pioneers do not determine a dominant firm: when late adopter represent the highest share of the population.

Proposition 5: *In the linear utility case, pioneers always determine the dominant firm of the final structure.*

Proposition 6: *In the type-dependent utility case, and when the share of pioneers is low, they do not determine the dominant firm in the market. While in the high median and in the high pioneers case they do.*

In the OS example, proposition five and six mean that when only the size of the network matters engineers will always determine the dominant or one of the dominant firms in the market. If the composition of network is important and agents value higher to be with similar agents, then engineers will be important to the determination of the final structure only in two cases: when engineers represent the highest share of the population and when designers represent the highest number . When students represent the highest number of agents, engineers do not have a key role.

Results summary			
Distribution. (High share of:)		*Linear utility*	*Type-dependent utility*
Pioneers	Mean Gini Index	9,256	8,663
	S.D. of Gini	102.8	236.9
	Market driver	Pioneers	Pioneers
Median adopters	Mean Gini Index	8,570	4,510
	S.D. of Gini	849.5	328.4
	Market driver	Pioneers	Pioneers and Medians
Late adopters	Mean Gini Index	8,619	4,504
	S.D. of Gini	75	632.1
	Market driver	Pioneers	Late and Medians

With the objective of analyzing the robustness of the model, the simulation was made with 3 firms and 3 type of agents. The results showed that the type-dependent utility still decreased the market concentration when compared to the linear utility case. However, the change in concentration was considerable lower in the 3 firms case when compared to the 5 firms case. The result suggests that as the options available and the heterogeneity of the agents increase the type-dependent utility effect increases.

When a low number of networks is available there is less difference among them. The TC of choosing a not-ideal network can not increase much because of the "closeness" of the options. If there are just a few types of agent and they are similar to each other, the incentives to create a cluster in a different network are less. In the same way as before, early adopters will start a network effect in their ideal choice. Late adopters will be less likely to form a new cluster because they are not that different from the pioneer and the TC of choosing the pioneers' choice is low. When options are limited and the differences among types is low, the results for the type-dependet utility are really close to the linear-utility case.

Proposition 7: *The relevance of the type-dependent utility increases as the number of networks and the heterogeneity of agents increase.*

3.5 Conclusions

When the structure of the network plays an important role to the derivation of utility of the agents and when there is heterogeneity in the adoption rates, the final market structure may significantly change when compared to the typical case where only the size of the network matters and all members adopt at the same rate.

I find that allowing different valuations from different types of agents may help to explain why in the real world it is possible to find cases where several firms survive, despite existing in markets with strong network effects. The model also helps to understand the development of consumer clusters in certain networks. Social groups, clubs, social media, and fraternities are only examples of markets where groups of similar persons join the same options and form a cluster.

Who drives the market? as we know, pioneers have a very important role to determine dominant platforms, but in the presence heterogeneity, type-dependent utility, and under some initial conditions, median and even late adopters can have an important effect on market structure.

The results of this model also highlight the importance of two important features of the market: the initial distribution of types and the speed of adoption behavior.

The initial composition of the population plays an important role in the final concentration of the market. A high share of early adopters can lock the market to a monopoly, a high share of late adopters may result in low concentrated market. The initial composition of agents is a key factor in determining which platforms will survive. The speed of adoption plays a similar role. If late adopters enter the market too late, they will not be able to form a new cluster in their ideal platform, which will allow pioneers to determine the unique dominant network. If at least a few late adopters enter in relatively early stages, they will be more likely to develop network effects that allow them to construct clusters in non-dominant firms.

This new information is relevant to the decision making involved in the development and application of public policies. There are government programs aimed to accelerate the adoption of certain technologies with the objective to raise social benefit. This model suggest that if some types react differently to the policy, the market will more likely change significantly. Markets must be regulated accordingly to their specific characteristics. Heterogeneity can't be ignored.

Further work can include firms that behave strategically, charge prices and innovate, agents that are forward looking and may choose several networks to join, the creation and elimination of firms and the change in population size. This paper also tries to incorporate to the general literature the importance that type-dependent valuations and the heterogeneity in adoption have on markets. And how the omission of this features can lead to a very different predictions of the same market.

References

[1] Abrajano, Marisa and Hajnal Zoltan. *White Backlash: Immigration, Race, and American Politics*. University of California San Diego, 2008.

[2] Ansolabehere, Stepen, James M. Snyder Jr. and Charles Stewart III *Candidate Positioning in US House Elections.* American Journal of Political Science 45(1): 136-159, 2001.

[3] Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg *Fractionalization.* Journal of Economic Growth 8: 155-94

[4] Brader, T., Valentino, N. A. and Suhay, E. *What Triggers Public Opposition to Immigration? Anxiety, Group Cues, and Immigration Threat.* American Journal of Political Science, 52: 959978., 2008.

[5] Bauer, Thomas; Lofstrom, Magnus and Zimmermann Klaus. *Immigration Policy, Assimilation of Immigrants and Natives' Sentiments towards Immigrants: Evidence from 12 OECD-Countries*. 2000.

[6] Borjas, George J. *Heaven's Door: Immigration Policy and the American Economy.* . Princeton University Press, 2001.

[7] Canning D., and M. Fay *The effects of transportation networks on Economic Growth* . Columbia University Working Paper.

[8] Carroll, Royce, Jeffrey B. Lewis, James Lo, Keith T. Poole, and Howard Rosenthal. *Comparing NOMINATE and IDEAL: Points of Difference and Monte Carlo Tests*. Legislative Studies Quarterly 34(4): 555591., 2009.

[9] Dustmann, Christian, and Ian P. Preston. *Racial and economic factors in attitudes to immigration.*. The B.E. Journal of Economic Analysis and Policy , 2007.

[10] Easterly, W., and R. Levine *Africa's Growth Tragedy: Policies and Ethnic Divisions* . Quarterly Journal of Economics 111(4), 1203-1250, 1997.

[11] Garand, James C. *Income Inequality, Party Polarization, and Roll-Call Voting in the US senate.*. Journal of Politics 72(04): 11091128, 2010.

[12] Gonzalez, Libertad and Ortega, Frances *Immigration and Housing Booms: evidence from Spain.* Institute for the Study of Labor,2009.

[13] Hopkins, Daniel J *Politicized places: Explaining where and when immigrants provoke local opposition.*. American Political Science Review 104 (1) 40-60, 2010.

[14] Mauro, P *Corruption and Growth* . Quarterly Journal of Economics 110(3), 681-712, 1995.

[15] McCarty, Nolan M., Keith T. Poole, and Howard Rosenthal. *Polarized America: The Dance of Ideology and Unequal Riches* MIT Press Cambridge, MA., 2006.

[16] National Journal Vote Ratings *https* : $//ballotpedia.org/National_journal_vote_ratings$ June 2016.

[17] Ottaviano, Gianmarco I.P. and Peri, Giovanni *The effects if Immigration on US Wages and Rents: a General Equilibrium Approach CReAM Discussion Paper No 13/07, 2007.*

[18] Poole, Keith T., and Howard Rosenthal *A Spatial Model for Legislative Roll Call Analysis.*. American Journal of Political Science 29(2): 357384, 1985.

[19] Saiz, Albert *Immigration and Housing Rents in American Cities* Journal of Urban Economics 61:345-371, 2007.

[20] Poole, Keith T., and Howard Rosenthal *A Spatial Model for Legislative Roll Call Analysis.*. American Journal of Political Science 29(2): 357384, 1985.

[21] Valentino, Nicholas, Ted Brader, and A. E. Jardina. 2013. *Immigration Opposition Among U.S. Whites: General Ethnocentrism or Media Priming of Attitudes About Latinos*. Political Psychology 34 (2) 149-166, 2013.

[22] Arthur, Brian *Competing technologies, increasing returns, and lock-in by historical events* The Economic Journal 99, 1991.

[23] Cabral, Luis *Dynamic price competition with network effects* New York University and CEPR, 2007.

[24] Chen, Doraszelski and Harrington. *Avoiding market dominance: product compatibility in markets with network effects* The RAND journal of economics, 2009.

[25] Church, Jeffrey and Gandal, Neil *Network effects, software provision, and standardization* The Journal of Industrial Economics, 1992.

[26] Church, Jeffrey and Gandal, Neil *Indirect network effects and adoption externalities* Foerder Institute for Economic Research , 2002.

[27] Easly, David and Kleinberg, Jon *Network effects, Networks, Crowds and Markets: Reasoning about high connected world* 2010.

[28] Farrel, Joseph and Saloner, Garth *Standardization, compatibility and innovation* The Rans Journal of Economics 16 70-83, 1985.

[29] Katz, Michael and Shapiro, Carl. *Network externalities, competition and compatibility,* American Economic Association, 1989.

[30] Markovich, Sarit and Moenius, Johannes. *Winning while losing: competition dynamics in the presence of indirect network effects,* International Journal of Industrial Organization, 2008.

[31] Swan, Peter *The functional form of network effects* Information Economics and Policy, 2002.