Diversity and Social Capital in the U.S: A tale of conflict, contact or total mistrust?*

Willie Belton†, Yameen Huq‡, Ruth Uwaifo Oyelere§

June 21, 2018

Abstract

This paper explores the relationship between ethnic fractionalization and social capital between 1990-2005. First, using data from 1990, 1997 and 2005 we test for time differences in the impact of ethnic fractionalization on social capital. Subsequently, we examine U.S. data for evidence consistent with the proposed outcomes in the conflict, contact, or hunker-down theses discussed in Putnam (2007). Putnam (2007) examines what happens to "trust" or "social capital" when individuals of different ethnicity are introduced into social, political and/or economic groups over time. Using an instrumental variable (IV) estimator, we find little evidence of heterogeneity in the impact of ethnic fractionalization on social capital over our period of analysis. In addition, using both fixed effect and IV estimators, we reject the *contact hypothesis*, but find evidence consistent with the outcomes predicted in both the *conflict hypothesis* and Putnam's *hunker-down hypothesis*, in inter-ethnic relations. Due to data limitations, we are unable to test directly which of these two thesis are more relevant for the U.S experience. However, we provide suggestive evidence in support of the *conflict hypothesis* over the *hunker-down hypothesis*. Our results suggest that between 1990-2005, as communities in the U.S became more diverse, there was a tendency for social capital to decline.

JEL classification: D71, Z10, J10, J19
Keywords: Ethnic Fractionalization; Social Capital; Trust; Diversity; Social Networks

*We are grateful for the comments and suggestions of two anaonymous referees. Comments are appreciated. Please direct correspondence to willie.belton@econ.gatech.edu or ruth.uwaifo@morehouse.edu.

†willie.belton@econ.gatech.edu, Department of Economics, 221 Bobby Dodd Way, Atlanta, GA 30332.

‡Georgia Institute of Technology, Atlanta, GA 30318.

§ruwaifo@agnesscott.edu, Department of Economics Agnes Scott College, Atlanta, Ga and IZA, Bonn Germany.

1 Introduction and Motivation

The publication of Putnam (2007), "E Pluribus Unum: Diversity and Community in the Twenty-First Century" and the significant increase in cross-national immigration of the last thirty years has been the catalyst that has merged the economics/political science/sociology/social psychology literatures in an effort to understand how ethnic diversity impacts the evolution of political, social, and economic outcomes over time. More specifically, social connectedness forms one of the basic building blocks for downstream economic, political and social outcomes. The impact of ethnic diversity on social connectedness and ultimate economic and political performance is of paramount importance as the world becomes more ethnically integrated. There are three possible outcomes of the impact of increased diversity on social connections. The first discussed in Allport (1954) is what he labeled as the *contact hypothesis*. Allport (1954) suggests that increased contact with people of different ethnicities leads to increased "inter-ethnic tolerance". In effect, the more contact we have across ethnic lines the more likely it is that we overcome ignorance and trust each other more. The second outcome can be classified as one of two dimensions of the constriction thesis. In particular, the *conflict hypothesis*, suggests that contention over scarce resources, the fear of re-distributive policies, and other zero-sum situations leads to increasing out-group distrust.[1] The third outcome is described in the *hunker-down hypothesis* which suggests that greater inter-ethnic interaction leads not only to more out-group distrust, but also to more in-group distrust as well.

As with Putnam (2007) much of the research into this relationship has attempted to examine this issue at the behavioral level using survey data. Survey data, though gathered at the individual level, is fraught with bias in terms of how questions are posed, social, economic, and political circumstances surrounding respondents, and processes used to aggregate and analyze the data. Moreover, measuring social capital has also been a challenge since definition differences exist and the concept is multidimensional in nature.

This research acknowledges the definition challenge in measuring social capital but for empirical purposes adopts the definition and measure of social capital proposed in Rupasingha, Goetz, and Freshwater (2006), hereafter (RGF). RGF (2006) uses a social capital index that is based on membership in social clubs, religious organizations, political organizations etc.[2]

Our research examines two related questions:

- Are there time differences in the relationship between ethnic diversity and social capital?
- Is there evidence supporting the *contact*, *conflict*, and/or *hunker-down* hypotheses of social interaction in the United States?

Our research is related to RGF (2006) as we make use of the same social capital index in our analysis. In addition, the variables they include in their estimation of the correlates of social capital, provide the building blocks for the empirical model we use to examine the questions highlighted above. Despite these similarities, our research differs from RGF (2006) in three ways. First,

[1] See Putnam 2007 page 142.

[2] A more detailed explanation of the index is given in section 4 of this paper.

we are focused on investigating heterogeneity in the impact of ethnic fractionalization on social capital across time rather than only identifying inputs into the production of social capital at the level of US counties. Second, rather than using the 1990-1997 time period considered by RGF (2006) we focus on a longer time period: 1990-2005. Finally, this research examines the applicability of the *contact* hypothesis or the constrict claim in the U.S. The constrict claim posits that ethnic diversity negatively affects social cohesion. Both the *conflict* and *hunker-down* propositions provide two different channels through which this happens. The conflict proposition of inter-ethnic relation proposes that as as diversity increases. distrust for other groups rises which ultimately affects social capital adversely. In contrast the hunker down proposition suggests as diversity increases social capital declines because individuals reduce social engagement with people within their ethnic group as well as those outside of their ethnic group. In particular, Putnam (2007) makes the argument that agents faced with increased ethnic diversity would "hunker-down" like a turtle in an attempt to protect themselves from the uncertainty created by engaging people of a different ethnic group.[3]

To answer the first question, similar to RGF (2006), we estimate a model of social capital production at the county level across three time periods. We control for potential factors that could impact social capital formation at the county level. We also address potential omitted variable bias using instrumental variables (IVs). We test for heterogeneity in the impact of ethnic fractionalization on social capital over time by comparing the estimated impacts across three periods using t-tests.

To address the second question we pool the data and estimate models of social capital production. The models of social interaction highlighted in Putnam (2007) lead to testable predictions. In particular, if social interactions follow the *contact hypothesis* then we should expect a positive relationship between ethnic fractionalization and social capital. If either of the other two aforementioned theses are more relevant, we should find a negative relationship.

As unbiased estimated effects are critical for identifying which of these claims are relevant, we exploit the panel nature of our data and estimate parameters using a fixed effect (FE) model specification. Using a FE specification eliminates most sources of selection bias by identifying effects using only variation over time within a county. To ensure our results are robust, we also estimate an IV model using fixed effects (FE-IV). The FE-IV model control for the possible, though unlikely, presence of time-varying unobservables at the county level that are correlated with ethnic fractionalization and social capital.[4]

Our results suggest that after controlling for potential selectivity issues, there are no significant time differences in the relationship between ethnic fractionalization and social capital between 1990-2005. Specifically, though the magnitude of the coefficient increased over time, t-tests of differences in means reveal that there are no statistical differences in coefficient size over time. We also find no evidence in support of the *contact hypothesis*. Our estimates provide evidence in support of *conflict and/or hunker-down theses*.

[3]While our data does not allow us to test directly whether conflict or hunker-down hypotheses is more relevant for the U.S experience, we provide descriptive analysis that can suggest the relevance of one theory over the other.

[4]While we cannot think of any unobservable that fits this criteria we cannot rule it out, hence, this robustness check using FE-IV.

This paper contributes to the literature by providing answers to questions that are not only timely, but could inform current discussions on the potential effects of increased diversity through immigration. The U.S. has experienced a significant influx of immigrants over the last few decades and this increase in immigrants has increased ethnic diversity within and across communities in the U.S. While Ottaviano and Peri (2012) provide evidence of the positive impact of immigrants on native wages,less is known about the impact of increased diversity on social capital formation within communities.[5] In addition, given that the past literature suggests that social capital is an important determinant of macroeconomic performance,[6] diversity may play an indirect role in affecting macroeconomic performance, therefore, its effect on social capital should be further examined. Further, since Putnam (2000) suggests that social capital may be declining within communities in the U.S., it is imperative that we examine whether diversity is a driver or it attenuates social capital. Finally, while research like Putnam (2007) have highlighted social interaction models hypothesizing the potential outcomes of increased ethnically-diverse social interaction, empirical evidence supporting or refuting these theses has not been established. Our paper fills this gap for the U.S. by providing evidence that increases in ethnic fractionalization during 1990-2005 has not fostered increased social capital formation.

The rest of our paper proceeds as follows. In section two, we review the past literature on social capital, group formation, and trust. Section three provides our conceptual framework and testable hypotheses. In section four, we provide a summary and descriptive analysis of the data sets used in this paper. Section five provides the empirical framework and justification of the modeling strategy. Section six summarizes our results, provides robustness checks and explains the limitations of the data and modeling strategy. This section also explores descriptive evidence that allows us to deduce whether the conflict or the hunker-down thesis is more consistent with the U.S data. Our summary and conclusion is found in section seven.

2 Literature Review and Background

Historically, economists have been concerned about how efficiently resources have been employed by economic agents creating long-term growth and development and ultimately solving the collective action problem. Though technical aspects have dominated the discussion in research by Solow (1956) and Swan(1956), North and Thomas (1973) examining the evolution of societies across time finds that choices made by nation-states around economic and political institutions impacted the downstream path of the nation-state. These choices influenced the actions of economic agents in ways that incentivized collective action leading to greater economic growth and development or inhibited such activity leading to economic under-performance and stagnation. Growth

[5]There is a growing literature considering the impact of immigrants on various outcomes in the U.S. Past research, like Card (2005) shows no significant effect on relative wages of native dropouts from relative supply of less-educated workers. Diette and Uwaifo Oyelere (2014) provide some evidence of small negative effects of immigrants with limited English on native students performance and Borjas (2013) suggests that the net benefit of immigrants to the native born population is trivial.

[6]See Durlauf (2002)

and development is usually a long-term process that depend on agents efficiently employing talent and resources currently for the promise of a significantly larger future pay-off. This process of investment has, in real terms, an uncertain outcome and in most cases leads involved parties to seek methods that reduces the risk of negative or unwanted outcomes. Economic agents use all available information in making their investment choices. Social connections provide a multiplicity of dimensions that offer information about an investment partner's future behavior and can be an imperfect substitute for repeated interaction. In effect, social similarities, though imperfect, can foster the production of social and economic networks that ultimately reduces uncertainty over the investment horizon. As suggested by Dasgupta (1988), "trust" among economic agents is the major ingredient necessary for growth and development to take place efficiently.

Sociologists and social psychologists, in a more precise way, examine the definition and evolution of trust, or in more broad terms "social capital". They further examine how social characteristics such as race, religion, and ethnic origin through their impact on social capital influence social, political and even economic outcomes. Van der Meer and Tolsma (2014) define social cohesion as a more fitting "neutral" term, given the variation in the definition of social capital provided by multiple authors including Bourdieu (1987), Coleman (1990), and Putnam and Nanetti (1993). Van der Meer and Tolsma (2014) define social capital "as the degree of interconnectedness between individuals that is both a result and a cause of public and civic life. It encompasses feelings of commitment, trust and norms of reciprocity, and is demonstrated by participation in networks and civic organization."(Van der Meer and Tolsma pp. 460-461). In an effort to understand the evolution of social capital social scientists have relied on early work by Blummer (1958) which attempts to address how issues of race and ethnicity play a role in constructing the social fabric of communities.

Economists have argued that social capital positively influences economic growth and development because trust reduces transaction costs and facilitates the cooperation necessary to solve the collective action problem.[7] Though most researchers support the idea that social capital leads to more efficient production, there are a few researchers who hold a contrary view.[8] They argue that the definition of social capital is problematic and they do not believe that linkages that enhance interaction are "capital" in the true sense. Our research however, leans on the definition provided in the Sociology, Social Psychology and Political Science research and is reviewed in Van der Meer and Tolsma (2014). These authors define social cohesion as a broader definition of social capital as "the degree of interconnectedness between individuals that is both a result and cause of public and civic life". We operationalize this definition using the measurement approach of RFG (2006).

Our research conceptualization has its roots in the work of Alesina and La Ferrara (2004), Putnam (2007) and RGF (2006). These papers examine the impact of ethnic diversity on social capital. Alesina and La Ferrara (2004) examines social capital indirectly, arguing that increased diversity leads communities to shift their consumption preferences for private goods over public goods. Implicit in their argument is the *conflict hypothesis*, which argues that

[7]See Colman (1988),(1990) and Putnam (1993)

[8]See Portes (1998), Arrow (2000), Solow (2000), Defilipis (2001) and Durlauf (2002)

an increase in ethnic diversity increases competition among ethnic groups for scarce resources, thus decreasing social capital.[9] Putnam (2007) advocates examining how increased diversity, within and across countries, impacts social, economic, and political outcomes. Putnam (2007) suggests that networks are as integral to production as are physical capital and labor. Since networks consist of people who trust one another due to shared common values and characteristics, it is important to understand how networks respond to change and greater heterogeneity in those key characteristics. RGF (2007) formulate a measure of social capital using county level data from the Regional Economic Information System. Their estimation of the determinants of social capital empirically confirm results found by both Alesina and La Ferrara (2004) and Putnam (2007) that increased ethnic diversity does indeed negatively impact the formation of social capital.

For empirical clarity we provide context for our analysis by examining social capital across the four dimensions suggested by Van de Meer and Tolsma (2014), formality, mode, target and geographical scope. Our measure of social capital is *formal* in the Pichler and Wallace (2007) sense and is based on data of voluntary membership in organization within a particular county within a state in the U.S. The *mode* of the measure used is not attitudinal, given that it is not developed fro survey data but is behavioral as individuals voluntarily join social groups that are consistent with their own social norms. The *target* of our measure is the general population as we try to access the impact of changing ethnic make-up of a county on the production of a formal or statistically measure of social capital. The *geographical scope* of the social capital measure is relatively large as we examine the production of social cohesion at the county level.

Using U.S. county level data, RFG (2006) developed measures of social capital which they argue are proxies for trust across communities. Following Putnam (1993), which argues that associational activities help communities solve collective action problems, RGF (2006) measure social capital by counting membership in sports clubs, religious organizations, political clubs, and the like across U.S. counties. Using these data, they created an index through the use of principal components analysis. They then examined how demographic variables, including ethnic fractionalization, explain variation in social capital across counties and time. To construct the ethnic fractionalization variable we follow the same method used in RGF (2006), which was originally constructed in Alesina et al (1999).[10] As noted by Van Der Meer and Tolsma (2014) this index is similar to the Herfindahl-Hirschman index, which can be interpreted as the probability that two randomly selected individuals living in the same geographical area have different ethnic backgrounds. The in-group relative to out-group distinction that is so prevalent in the social cohesion literature is difficult given this data does not explicitly measure attitudes across groups. However, our measure of ethnic diversity is based on Blummer (1958) and later Quillian (1995) where racial preference is determined by group position rather than individual preferences as is the case in most studies using survey data.

While RGF's measure captures multiple dimensions of social capital, it is important to mention that this measure of social capital is not without criti-

[9]Our research takes a more direct approach by following the work of Putnam (2007) which, through the use of survey data, examines directly the impact of diversity on social capital.

[10]The precise construction of the ethnic fractionalization variable can be found in the Appendix.

cism[11] For example, Knack and Keefer (1997) argue that Putnam-type social capital measures that use associational activities are unrelated to trust. Using survey data, they find that their measure of trust was at best weakly correlated with associational activities and concluded that these activities do not explain trust between individuals. However, trust is difficult to measure and the measures of trust in the World Value Survey used by these authors is imperfect. It is also clear that social capital is a multi-dimensional concept in which trust is but one of its many facets. Given these facts, it remains interesting to examine the impact of ethnic fractionalization on the production of social capital across time, even if the RGF measure of social capital is an imperfect proxy for trust.

3 Conceptual Framework

To understand how social capital is developed over time, we focus on Blau's (1977) theory based on the homophily principle. The homophily principle states that people who are similar across socio-demographic dimensions are more likely to trust one another and develop social capital. RGF's (2006) definition of organizations suggests that members of organizations tend to exist in similar socio-demographic space. This theoretical space is defined across characteristics such as race, income, education, age, and other demographic factors. Figure (1) reveals that initial socio-demographic space is made up of homogenous member across race, income, education and other socio-economic characteristics. McPherson, Popielarz, and Drobiic (1992), hereafter (MPD), use Blau's model in a dynamic context to examine the impact of social networks on organizational behavior over time. In particular, MPD's model examines how socio-demographic changes over time impact the missions and goals of the organization as revealed in Figure (1). Over time as new members are added the socio-demographic space could retain its basic characteristic i.e., homogenous and cohesive, if new members generally have the same or similar socioeconomic characteristic. However, if new members have different socioeconomic characteristics the socio-demographic space could become more heterogenous and less cohesive leading to the models first discussed in Allport (1954) and empirically examined in Putnam (2007).

While the aforementioned conceptual framework explains the basic structure of groups, an evolutionary model is needed to explain how groups change over time as members enter and leave. MPD (1992) suggests there are three factors to consider: *variation*, *retention*, and *selection*. *Variation* focuses on differences in socio-demographic characteristics such as race, income, and age. *Retention* examines how groups recruit members to keep their socio-demographic space constant. A group able to keep its socio-demographic space constant will exhibit more "group-like" behavior. *Selection* involves the recruitment of new members and is of concern because group members tend to bring in new members through the homophilous network ties among existing members.

Since it is the selection process that governs the evolution of the group over time, it is important to identify whether new entrants are *stabilizing* or *disruptive* selections. *Stabilizing* selection occurs when the recruitment/attrition ratio is smaller for members further away from the center in socio-demographic

[11] Like most indexes, the RGF social capital measure has limitations. We discuss some of these in more detail in section 7.

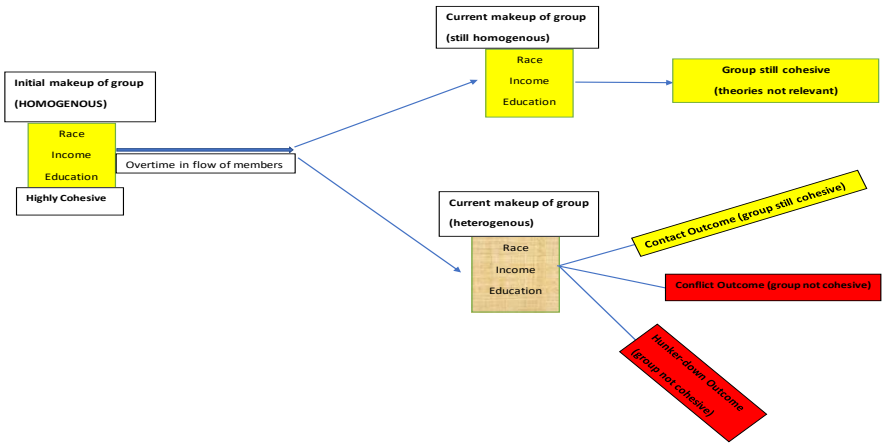

Figure 1: Representation of Conceptual Framework

space. This would make the group members more similar over time. *Disruptive* selection occurs when the recruitment/attribution ratio is greater for members away from the center in demographic space, causing the group to generalize and increase the variance over its socio-demographic dimensions. Selection can clearly impact the behavior of the group in its effort to achieve goals.

The focus of this paper is the relationship between variation in a demographic variable associated with group members and trust across the group's membership. More specifically, we examine how variations in the race and ethnicity of group members impact the production of social capital. Essentially testing whether increases in racial diversity in a county induces *disruptive* selection which over time reduces social capital. To investigate the three aforementioned theses in a more formal fashion we adopt an empirical model similar to that of RGF(2006).

4 Detailed Data Description

The data used to examine our questions of interest is U.S. Census data for 1990, 1995 and 2005 along with the social capital index used in RGF (2006). The data used by RGF to construct this index was first developed by Anil Rupasingha and Stephan J. Goetz in 1992 and is housed in the Northeast Regional Center for Rural Development at the College of Agricultural Sciences at Penn-State.[12]

[12]It is important to mention that apart from the share of democratic votes in a recent presidential election in a county and the share of votes for the leading candidate in the most recent presidential election, all control variables both dependent and independent used in the analysis are measured in 1990, 1995 and 2005 respectively or adjacent years. Details of exact year variable is measured can be found in the appendix. Also, all the variables that are used to construct the social capital index apart from voting fraction from nearest Presidential election

The Northeast Regional Center's data repository contains social capital data for the years 1990, 1997, 2005 and 2009 for each county in the continental U.S.[13] We make use of 1990, 1997 and 2005 data solely for consistency because the calculation of the index is different for the 2009 survey. Specifically, the method used to create the social capital index for 2009 is incompatible with the data available for 1990.[14]

Table (1) provides the components used by RGF (2006) to build the social capital index. This is a composite index created using principle components analysis. The specific variables used in the index are: the aggregate total of listed organizations and non-profit organizations as well as the the census response rate and voting fraction of the population.[15] Summary statistics for the variables used to create the social capital index are presented in Table (2). Because the social capital index can range from negative to positive, the mean values in each panel are generally near zero.[16]

Table 1: Components of the Social Capital Index

Name	Description
Bowling centers	Total number of bowling centers
Civic associations	Total number of civic and social associations
Gyms	Total number of physical fitness facilities
Golf courses	Total number of golf courses
Religious organizations	Total number of religious organizations
Sports clubs	Total number of sports clubs
Recreation clubs	Total number of recreation clubs
Political groups	Total number of political organizations
Professional groups	Total number of professional organizations
Business groups	Total number of business organizations
Labor groups	Total number of labor organizations
Other groups	Total number of other membership organizations
Total organizations	Aggregate sum of organizations listed above
Census responses	Census response rate
Non-profit organizations	Total number of not-for-profit organizations
Population	Total residential population
Voting fraction	Voting fraction from nearest Presidential election

The definitions for the independent variables used in our empirical analysis are found in Table (3). This data is derived from the CenStats Databases of the United States Census Bureau (2010). The independent variables are constrained

are measured in the year of analysis.

[13]Note that this data is restricted to 48 out of 50 states. Hawaii and Alaska are excluded.

[14]It is important to note that the 1997 and 2005 surveys can be recalculated to be compatible with the method used in the 2009 panel. However we choose not to consider 1997-2009 period given the recession happened during this period and could serve as a confounding factor in our econometric analysis. Moreover, given the changes in social interaction and networking from the mid 2000s through the increased availability of the internet and social media, our measure of social capital is likely to be less precise for more recent data. Hence, it is preferable to consider the period 1990-2005 rather than 1997-2009.

[15]It should be noted that the voting fraction is taken from the nearest Presidential election. For the 1997 and 2005 panels, these are the 2000 and 2004 Presidential elections, respectively. The 1990 panel averages data from the 1988 and 1992 Presidential election.

[16]The social capital index for the year 2005 is calculated excluding recreation clubs and "other groups".

Table 2: Descriptive Data (Dependent Variable and Components)

	(1990)	(1997)	(2005)
Social capital	0.001	0.001	0.000***
	(1.350)	(1.298)	(1.646)
Bowling centers	1.910	1.766	1.490
	(4.671)	(3.954)	(3.106)
Civic associations	12.513	11.489	10.097
	(34.927)	(30.602)	(24.765)
Gyms	2.444	3.541	10.161
	(7.980)	(11.874)	(30.806)
Golf courses	0.888	1.949	3.825
	(2.270)	(3.854)	(7.269)
Religious organizations	40.979	50.711	54.982
	(87.040)	(106.072)	(116.530)
Sports clubs	0.288	0.623	0.244
	(1.662)	(2.877)	(1.038)
Recreation clubs	4.172	4.779	
	(10.119)	(11.758)	
Political groups	0.495	0.578	0.950
	(2.112)	(2.650)	(4.442)
Professional groups	1.670	2.499	2.407
	(7.268)	(10.465)	(11.259)
Business groups	3.879	4.396	5.572
	(13.131)	(14.004)	(17.548)
Labor groups	6.071	6.008	5.132
	(19.293)	(18.067)	(15.991)
Other groups	3.012	2.365916	
	(8.028)	(5.448)	
Total organizations	78.320	90.795	94.859
	(188.472)	(209.379)	(218.882)
Census responses	66.666*	62.519**	64.248
	(8.213)	(8.831)	(8.881)
Non-profit organizations	40.674	42.131	446.417***
	(152.390)	(157.681)	(1357.03)
Population	78832.7	84821.72	93837.78***
	(262630.9)	(275600.5)	(304866.5)
Voting fraction	54.001	52.888	58.256***
	(10.851)	(9.805)	(9.577)
Voting fraction (2nd)	60.265		
	(10.210)		

All variables have a sample size of 3110 unless otherwise indicated by asterisks.
Includes standard deviation in parenthesis.
* N = 2434, ** N = 3066, *** N = 3107
*** RGF's data source for non-profit organizations is identical for 1990 and 1997 but different for 2005.

Table 3: Description of Independent Variables

Name	Description/Definition
Ethnic Fractionalization	Level of Diversity
Black Proportion	Fraction of population that identifies as black
Income Per Capita	Personal income per capita
Education	Percent of the population that is over 25 and has a Bachelor's Degree
Family Households	Fraction of total households that are family households
Median Age	Median age
Square of Median Age	Square of median age variable
Urban	1 if county population is greater than 2500, 0 otherwise
Population Density	Population per square mile
Residence Time	Time resident spends in a given county
Working Women	Percent of population made up of women in the labor force

Table 4: Descriptive Data (Independent Variables)

	(1990)	(1997)	(2005)
Ethnic Fractionalization	0.183	0.208^{***}	0.225^{***}
	(0.177)	(0.177)	(0.177)
Racial Dominance	0.876	0.861^{***}	0.850^{***}
	(0.141)	(0.141)	(0.141)
Black Proportion	0.086	0.088^{***}	0.088^{***}
	(0.144)	(0.145)	(0.142)
Latino Proportion	0.039	0.051^{***}	0.060^{***}
	(0.083)	(0.081)	(0.084)
White Proportion	0.869	0.854^{***}	0.843^{***}
	(0.158)	(0.159)	(0.159)
Asian Proportion	0.006	0.008^{***}	0.009^{***}
	(0.012)	(0.014)	(0.016)
Income Per Capita	11122.81	17476.74	-263735
	(2681.445)	(3956.538)	(1.59e07)
Education	13.475	16.493	18.647
	(6.577)	(7.809)	(8.528)
Family Households	73.580	70.659^{**}	68.417^{***}
	(4.664)	(4.801)	(5.339)
Median Age	34.407	37.354	40.329
	(3.611)	(4.078)	(5.178)
Square of Median Age	1196.867	1411.925	1653.246
	(253.259)	(296.384)	(407.369)
Urban	0.963	0.965	0.960
	(0.189)	(0.185)	(0.196)
Population Density	221.022	238.955^{**}	253.655^{***}
	(1438.214)	(1649.197)	(1735.426)
Residence Time	-33.381^{*}	-44.520^{***}	282.662^{****}
	(6522.952)	(3829.71)	(8409.171)
Working Women	20.451^{**}	22.139^{**}	22.612^{***}
	(2.954)	(3.010)	(3.105)

All variables have a sample size of 3110 unless otherwise indicated by the asterisks.
Includes standard deviation in parenthesis
* N = 3109 , ** N = 3108 , *** N = 3107 , **** N = 3106

to all counties in the continental United States. Summary statistics of these variables can be found in Table (4) and the description of how these variables are calculated is highlighted in the appendix. This research uses control variables found in most of the work examining ethnicity and social capital including Alesina and La Ferrara (2004), RFG(2006), and Gundelach (2014). Control variables other than the ethnic fractionalization measure(s) include income per capital, education, fraction of family households, median age and median age squared, a dummy variable for urban relative to rural population per square mile, population density per square mile, time the average resident has spent in a given county and the percent of the population made-up of working women.[17] High wages tend to be associated with employment that requires less hours, leaving more time for civic and social engagement. Putnam (1995), Helliwell and Putnam (1999) as well as Glaeser et al., (2002) have all documented the positive relationship between social and human capital. RGF(2006) argues that at the community level, education attainment is highly correlated in a positive manner with civic engagement. Putnam (1995) cites the breakdown of traditional family structure as being negatively related to social capital formation. However, the family is usually associated with more involvement in local schools and other social initiative suggesting that the increase in the proportion of intact family units should be associated with increase social capital. Since Putnam (1995) suggests that high levels of social involvement tend be positively correlated with age we include a variable that captures age as well as following Glaeser et. al (2000) we include an age squared term capturing the life cycle relationship of social capital. Because we believe that social interaction in rural areas differ from that in more metropolitan areas we follow Putnam (1995), Glaeser et. al (2000) and Brown (2001) and include a dummy variable for rural relative to urban counties. We also include variables covering population density per square mile and the average time spent by a resident of each county in an effort to capture the behavior of city dwellers relative to those living in more rural situations. Finally, to capture the changing role of working women in the social structure we include a variable that capture the percent of the county made-up of working women.

Figure (2) highlights the distribution of the social capital index across the three time periods. This figure shows a slight leftward shift in the distribution by 2005. This leftward shift is suggestive evidence of a reduction in social capital over time. Figure (3) shows the distribution of ethnic diversity (our primary variable of interest) across counties in the three survey periods. This figure suggests that the density of counties with little or no ethnic fractionalization has declined overtime. Conversely, there is an increase in the density of counties with more ethnic diversity.

[17]Some variables included in RGF(2006) are not included in our empirical model because of insufficient theoretical justification for their inclusion in the production of social capital. For example share Black in a county.

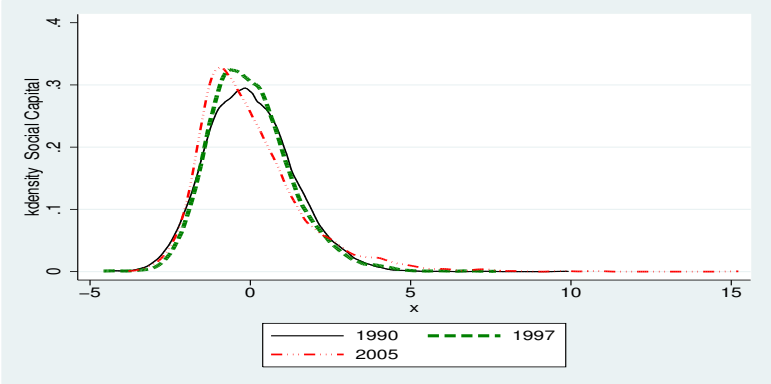

Figure 2: Distribution of Social Capital Index Over Time

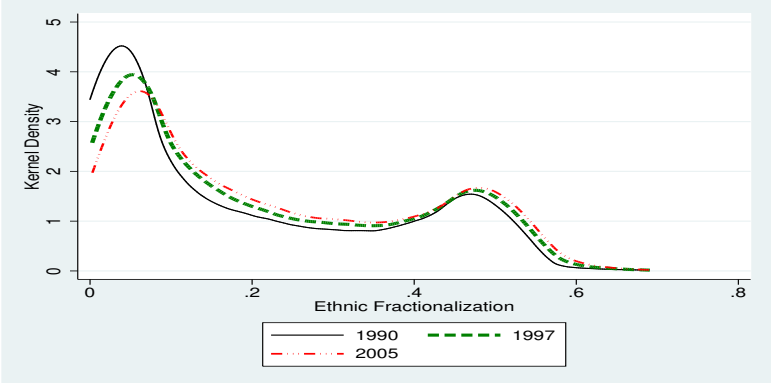

Figure 3: Distribution of Ethnic Fractionalization across counties Over Time

5 Methodology

Econometric Model for Question 1

We address the issue of time differences in the impact of ethnic diversity on social capital by initially using a simple OLS model with standard controls as depicted in equation (1).

$$S_{c,t} = \rho_{0t} + \rho_{1,t} e_{c,t} + \sum_{j=1}^{J} \delta_{j,t} x_{j,c,t} + \epsilon_{c,t} \tag{1}$$

In equation (1), S is social capital in county c in year t, e is ethnic fractionalization in county c in year t, $x_{j,c,t}$ are additional explanatory/control variables ($j = 1, ..., J$) that could affect social capital at the county level and ϵ is the error term. Our control variables include income per capita, the percentage of the population over 25 with a Bachelor's degree, the fraction of family households over total households, a dummy variable differentiating urban or rural counties, the median age, a quadratic variable on median age, and population density.[18] We also include state fixed effects (not included in RGF, 2006) because a simple OLS estimation of equation (1) can lead to biased estimates due to the potentially endogenous nature of ethnic fractionalization. Specifically, our parameter of interest could suffer from omitted variable bias if there is an unobserved determinant of social capital at the state level, correlated with ethnic fractionalization. For example, differences in institutions and/or economic polices could be correlated with ethnic fractionalization. At the same time these institutions and/or polices could produce lower levels of trust across the population and hence lower levels of social capital. We reduce the likelihood of this potential source of bias in the estimation of ρ_1 by introducing state fixed effects. In this instance, the state fixed effects controls for time invariant institutions unique to a state that potentially impacts social capital. Hence, the coefficients in the model are estimated using variation across counties within a state at a given period of time.

We derive estimates of ρ_1 and other parameters for t= 1990, 1997 and 2005. These estimates are then examined for statistical differences using T-tests.[19] Although we are confident that the inclusion of state fixed effects attenuates bias in our estimates, it is still possible to argue that even within states, there exists an omitted variable that is correlated with ethnic fractionalization across counties within a state and also correlated with social capital accumulation. While we are unable to identify such a variable, we cannot rule out this possibility. Hence, we address this potential source of bias using instrumental variables (IV) and a

[18] Apart from population density, the independent variables used in estimation are a subset of the variables included in RGF (2006) social capital empirical model. We choose to exclude some variables RGF(2006) include in their model to reduce potential multicollinearity. For example we do not include share of family households with children because of its potential collinearity with share of family households. We do not include a nonlinear relationship between ethnic fractionalization and social capital because when we graph social capital and ethnic fractionalization using a median spline or a lowess we observe a linear relationship.

[19] It is important to note that we could have also tested for time differences in the impact of ethnic fractionalization by pooling all the data and interacting ethnic fractionalization with time dummies. However, given the length of time that has elapsed between each survey year, assuming other parameters in the social capital model do not change over time might not be appropriate.

2-stage least squares (2SLS) estimator. We estimate our coefficients of interests separately for the three time periods considered. In the IV specification just as in the OLS specification, we estimate the coefficient on ethnic fractionalization for each period and test for significant differences in the estimates in each period using T-tests.

In this analysis we make use of two instruments which we refer to as our preferred instruments. These instruments are level of inequality in the county (Gini) and the share of Blacks in a county.[20] As a robustness check, later in the paper, we introduce two alternate instruments: the share of democratic votes in a recent presidential election in a county and the share of votes for the leading candidate in the most recent presidential election. We combine the Gini and Black population proportion instruments in our estimation because they satisfy the J-Hansen over identification restriction test.[21] In contrast, any other pair of instruments or combination of three instruments fail this test suggesting that at least one instrument is not valid. The fact that only Gini and Black share satisfy the over-identification restriction in most instances suggests that at least one of these instruments is valid. While our other alternate instruments appear exogenous, the Gini and Black share are preferred given their relevance (correlation with ethnic heterogeneity). Moreover, both performed well on standard weak instrument tests and condition on the controls we include in our analysis, these instrument should satisfy exclusion restrictions.[22]

The biggest concern in the IV analysis is whether relevant instruments satisfy exclusion restrictions. We argue that our instruments are valid as we do not expect a direct effect of each of these variables on social capital in a county. The only way these variables affect social capital is through their impact on ethnic fractionalization. For example, higher democratic vote can be associated with more racial heterogeneity but having higher democratic votes in a county does not affect the accumulation of social capital. Both Republican and Democrat leaning counties form social groups and accumulate social capital. An additional example is the instrument, share of Blacks in a county. Just as with democratic vote, there is no logical reason to expect that higher proportions of Blacks will have an independent effect on levels of social capital. There is no evidence that Blacks form social groups or accumulate social capital differently than do other ethnic groups. However, the increase in Blacks or any other ethnic group in a county could increase ethnic fractionalization which could lead to lower levels of trust and social capital as predicted by the conflict theory. Similarly, the level of inequality in a county does not inherently reduce or increase the accumulation of social capital. However, higher levels of inequality in a county can be associated with ethnic fractionalization through past historical institutions.[23]

[20]It is possible to argue that using inequality within a county may not satisfy exclusion restrictions, given Alesina and La Ferrara suggested an association between social capital and inequality. However robust empirical evidence of this relationship is lacking. They find that when appropriate controls are included inequality does not have statistically significant effect on social capital. However inequality is correlated with ethnic fractionalization- making it a possible IV choice.

[21]The exception occurs in the estimation using the 1997 survey. In this instance using the Black share IV alone is preferred but estimates using both instruments and just Black shares produces estimates that are not statistically different (-0.437 using Black share alone vs -0.462 using both.).

[22]Examples of weak instrument tests we conduct include "first-stage F-statistic", Shea partial R^2, Stock-Yogo (2005) bias method.

[23]Later on in the paper, we provide evidence of this lack of independent effect of our preferred

Econometric Model for Question 2

To address our second question regarding evidence in support of Putnam's claims, we pool the data from the three periods and re-estimate the model using four estimation strategies. First as a benchmark, we estimate β_1 in equation (2) using OLS.

$$S_{ct} = \beta_0 + \beta_1 e_{ct} + \sum_{j=1}^{J} \phi_j x_{jct} + \epsilon_{ct} \tag{2}$$

Second, we exploit the panel nature of our data and estimate a panel model as highlighted in equation (3) making use of a FE estimation method.

$$S_{ct} = \beta_1 e_{ct} + X'_{ct}\lambda + \theta_c + \epsilon_{ct} \tag{3}$$

In equation (3), θ_c is the individual county level effect and X is a vector of all the control variables that vary across county and time. We also include time dummies in X. By exploiting a fixed effect panel model specification, our effects are identified over variation within a county over time. The fixed effect model attenuates the potential of deriving biased estimates which can occur when using an OLS strategy to estimate effects even with time and state level dummy variables. While it is still possible to argue that the fixed effects model is threatened by within-county time varying unobservables, the existence of such variables is unlikely. Hence, conditional on the variables included in our analysis,it is difficult to provide a compelling story that there is still an omitted variable within a county that is time varying, affects social capital, and is correlated with ethnic fractionalization in the county.[24]

Our third estimation strategy addresses the potential limitation of the fixed effects model by making use of IV/ 2SLS estimation methods. The only difference here compared to our prior IV analysis, is that here we pool all the data and include year dummies. Similar to our previous IV specification, the main assumption for consistent estimation of β_1 are valid instruments ($E(\mu_{c,t}|z_{ct} = 0)$.

Our fourth strategy exploits the panel nature of the data along with the use of an IV approach. This approach refereed to as FE-IV is useful in light of the probable existence of county level fixed effects and the distinct possibility that $e_{c,t}$ is correlated with $\epsilon_{c,t}$. In this case, we need to make use of instruments $z_{c,t}$ which are correlated with e_{ct}, but are not correlated with ϵ_{ct}. The FE-IV estimation strategy given its advantages, is our preferred specification for deriving consistent estimates.[25]

Strategy to Test for the Relevance of Hypotheses

The *contact hypothesis* suggests that greater levels of diversity lead to increased social capital in communities. To empirically test this hypothesis we focus on the estimated β_1 in our IV and fixed effect models. Our null hypothesis is $\beta_1 \leq 0$.

instruments on social capital by estimating simple linear regressions of social capital on our control variables, ethnic fractionalization and our instruments.

[24]One potential drawback of using the fixed effect model with few time periods is the limited variation over which potential effects are being identified.

[25]It is important to mention that in both the panel estimation and the models using pooled data, we cluster our standard errors given that our variable of interest varies at the county level.

We reject our null hypothesis if $\beta_1 > 0$, meaning contact claim is relevant. Failing to reject the null implies: (1) If $\beta_1 = 0$ then our results suggests that none of the theories are consistent with trends in U.S.data. (2) Alternatively, if $\beta_1 < 0$ then either the *conflict hypothesis* or *hunker-down hypothesis* is relevant. It is important to mention that our estimation strategy does not allow specific differentiation between the conflict and hunker-down hypotheses. To provide evidence for or against either of these notions, we look at trends at the bottom and the top of the ethnic fractionalization distribution. We discuss this *prima facie* evidence in the results section.

6 Results

6.1 Summary of Estimated Effects for Basic Control Variables

Before addressing our two key questions, it is useful to examine the estimated coefficients for the control variables in the estimated models summarized in Tables (5-9). While some of the control variables generally hold intuitive relationships with social capital throughout all of the different specifications, the signs of some control variables vary across specifications. In instances where estimates of controls change significantly across specification, we focus on estimated relationships in our preferred model.[26]

The level of education shows a positive relationship with social capital across time. In effect, a county that holds proportionally more undergraduate degrees tends to have a greater level of social capital. Educated people are more likely to be employed in workplaces that place greater importance on teamwork and building social connections. We can also argue that skilled occupations are scalable and have more flexible working conditions, which allow more time to engage in the production of social capital.

Whether a community is urban shows a significant, negative relationship with social capital across time. Urban environments are generally perceived as less social, possibly due to higher crime or less social space. Residence time, the approximate total time an individual spends in a city or town, shows a negative, but insignificant relationship with social capital. A possible reason for this is the difficulty in approximating residence time accurately. The percentage of working women shows a significant positive relationship with social capital. This result suggests that more women in the workplace increases social capital, which corroborates the reasoning behind the relationships for level of education.

The variables income per capita, percentage of family households, and median age show different relationships depending on the model specification. Our preferred specification (IV-FE) summarized in Table (8) reveals a significant positive relationship between income per capita in a county and social capital. Further, the percentage of family households also shows a positive, but non-significant relationship with social capital. Population density shows a negative relationship with social capital. Median age shows a negative and significant

[26]While the control variables we include in our model are similar to RGF (2006), we only include a subset of the controls used in RFG (2000) and as a result we do not directly compare our estimates to theirs. However, the direction of the estimated coefficients of the subset of variables common to both papers, are similar in most cases although the magnitudes differ.

relationship, while the square of median age shows a positive significant relationship. This result suggests a nonlinear relationship between social capital formation and median age of individuals in a community.[27]

6.2 Testing for time differences

Columns (1)-(3) of Table (5) summarize OLS results across our three time periods 1990, 1997, and 2005 and provide a benchmark for possible estimated effects. In contrast, columns (4)-(6) summarizes the results across the three time periods from the second stage of the IV estimation. The first stage estimates for the cross-sectional IV analysis in columns (4)-(6) of Table (5) are summarized in Table (A1) of the appendix columns(1)-(3).

We present results using our preferred instruments, the combination of the Gini coefficient and the share of Blacks in a county. We posit that consistent estimates can be derived given these instruments satisfy the exclusion restriction conditional on the controls we have included in the analysis. See appendix A for a detailed discussion on evidence in support of the instruments. Furthermore, we do not worry too much about the consistency of our estimates when testing for time differences. This is because our test for time differences will be valid as long as any potential bias in our OLS or IV estimates are time invariant. The first stage results (Table (A1) in our appendix columns(1)-(3)) reveals a strong positive relationship between ethnic fractionalization and both the Black share of the population in the county and the Gini coefficient in the county.[28]

Our variable of interest in Table (5) is the estimated impact of ethnic fractionalization on social capital for each survey year. Notice across the three periods the estimated impact is negative using both OLS and IV. This result suggests that an increase in ethnic fractionalization is negatively correlated with social capital. Our IV estimates are slightly larger than our OLS estimates. This difference suggests that our OLS estimates are slightly downward biased. Given that we can infer causal relationship using the IV model, these results suggest that higher levels of ethnic fractionalization at the county level leads to a decrease in social capital at the county level in the three periods we analyze (1990, 1997 and 2005).

While these yearly IV estimates of the impact of ethnic fractionalization are useful, our first question is focused on investigating the existence of time differences in the impact of ethnic fractionalization on social capital. To address this question, as noted in our empirical section, we test for statistical difference in the estimated impact of ethnic fractionalization on social capital across the three survey periods using T-tests.

Formal testing for statistical difference is important as we cannot conclude the existence of time difference based on simply observing changes in coefficient magnitudes. Specifically, in both the OLS and IV specifications the magnitude of the negative impact of ethnic fractionalization on social capital is increasing over time. For example, the estimated negative impact rose by 0.07 from 1990 to 2005 using OLS. Alternatively, in the IV model the negative impact rose by 0.247 from 1990 to 2005. However, these increases in the estimated impacts

[27] If we do not include a quadratic term for median age the estimate on median age in the IV-FE is positive and more consistent with other specifications.

[28] Notice in three of the four cases, the sign on the estimated impact of Gini and the share of Blacks in a county is positive and significant at least at the 10 percent level.

across survey years does not provide evidence of statistical difference. Results from a two-sample mean-comparison test suggests that estimates of the impact of ethnic fractionalization are not statistically different i.e., that there is no statistical difference between estimates in 1990 and 1997, 1990 and 2005, and 1997 and 2005.

The conclusion from this analysis is that while ethnic fractionalization has a negative impact on social capital and this effect has persisted over time, this effect has not changed in magnitude over our survey period (1990-2005). This finding suggests that there are no time differences or time heterogeneity in the impact of ethnic fractionalization on social capital between 1990 and 2005.

6.3 Testing for Evidence of Contact, Conflict or Hunker-down

To provide evidence for or against the contact hypothesis, we focus on the estimated effects using the pooled data. Table (6) provides a summary of the estimated impact of ethnic fractionalization on social capital using the different estimation strategies highlighted above. Columns (1) provides OLS estimates, column (2) provides estimates using the FE method, and in columns (3) -(5) we summarize the estimates using alternative instruments. Column (3) presents results using our preferred instruments (the Gini coefficient and the Black share in the county). Columns (4)-(5), in contrast, provide estimated effects using democratic vote in a county and share of lead vote in the closest presidential election as instruments. Table (A1) in the appendix columns (4) to (6) provide a summary of the first stage results of the three IV estimations in Table (6).[29]

The strength of the fixed effect model is the identification of the relationship between ethnic fractionalization and social capital using within county variation. However, the possibility that ethnic fractionalization in a county is correlated with both an unobservable within the county and with social capital is our rationale for also exploring an IV strategy. All the estimates of β in Table (6) are negative and significant. Given this negative relationship in all cases, we can clearly reject the *contact hypothesis*. Therefore, in the case of the U.S., the *conflict* or *hunker down hypotheses* are more relevant in explaining what happens to social capital with increased diversity. It is worth noting that though magnitudes differ across the OLS, FE and IV (1) specifications, estimates are not statistically different.[30]

Another observation from Table (6) is that estimated effects when we use either the democratic vote share or the lead vote share as instruments are statistically different from the estimated effects using the OLS, FE and the preferred instruments. Though it appears that these instruments satisfy exclusion restrictions both are more likely to suffer from weak instrument issues given the low correlation between these instruments and ethnic fractionalization.[31]. It is

[29]In this pooled analysis just as in the individual year IV analysis, only a combination of the Gini and Black Share satisfies the over-identification restriction test. This is why we use the other two instruments individually.

[30]We test to see if estimated effects are statically different using T-tests and fail to reject our null of equal means across OLS vs FE estimates, OLS vs preferred IV estimate and FE vs preferred IV estimates.

[31]See first stage results in Table (A1) columns (5) and (6) in the appendix for estimated effects.Though these instruments have F stats greater than 10, the Shea Partial R^2 is less than 0.1 when either alternative IVs are employed which is in contrast to 0.45 for our preferred IV

also worth mentioning that even if the estimates of β are consistent using the two alternative instruments, we are not overly concerned with the difference between these estimates and our preferred IV estimates, given that instrumental variables estimators generally estimate local average treatment effects (LATE), with the specific average depending on the choice of instruments.[32].

In Table (7) columns (1) to (3), we present the results using an IV-FE estimator. Column (1) provides results using the Gini coefficient and the share of Blacks in a county as instruments. Columns (2) and (3) provides estimates when we instrument for ethnic fractionalization using the percent of people voting democratic in a county and the percent of the county voting for the leading party as instruments, respectively. The IV-FE specification allows us to combine the benefits of IV and fixed effects. However, it is important to mention that the potential for weaker correlation between the instrument and ethnic fractionalization is higher here given we are identifying effects using variation within a county. The benefit is that the instruments are more likely to satisfy exclusion restrictions.[33] We concentrate our discussion on our preferred IV-FE results in column (1) of Table (7) but also present the estimated effects using the alternative instruments for completeness.[34]

Table (7) shows that the estimate of β using IV-FE and our preferred instruments is negative and similar to the estimates summarized in Table (6) using other estimation methods. These results suggest that increases in ethnic fractionalization leads to a decline in social capital. Hence, we reject *contact hypothesis* and posit that our findings are consistent with the *conflict* and/or *hunker down hypotheses*.[35]

Note that the estimated effect using the IV-FE is much larger than using the IV or FE models: -3.318 versus 0.453 and 0.891.[36] This larger magnitude is not a surprise given identification of effects are different using both models. In the regular IV, we identify effects using variation in the instrument within a state, but in the case of the IV-FE we are identifying effects using variation in the instrument within a county over time.

---

specification. When instruments are weak, IV estimates may not be consistent as the 2SLS estimator with weak instruments is biased in small samples. The aforementioned issues make us more cautious about the large estimated effects with the alternative instruments.

[32]Angrist and Imbens, (1995)

[33]It is worth mentioning that first stage estimates (not included in the paper) using our preferred IV-FE method are still very strong and we do not worry about weak instrument issues. Also the Hansen J-test for over-identification restrictions still suggests that at least one of these preferred IVs is valid.

[34]As noted with the IV estimation, although these alternative instruments pass the first stage $F > 10$ rule of thumb, we are concerned that estimates using these instruments may be inconsistent given low Shea partial R^2s and the weak correlation between these two alternative instruments and the endogenous variable in the first stage.

[35]It is useful to mention that although we do not focus on the estimates using our alternative instruments, the democratic share instrument in (column 2) provides results that suggest no significant relationship between ethnic fractionalization and social capital. In contrast, the lead share IV (column 3) suggests a negative relationship which is consistent with our preferred IV-FE model and our other specification results.

[36]This effect is also statistically different from the estimates derived using the OLS, IV and FE.

6.4 Robustness Checks

So far our different model specifications have all led to the rejection of contact hypothesis. However, one of the challenges in considering social capital is how it is measured. In this paper we make use of an index that is readily available and put together by researchers with a lot of knowledge in this field. However, this index measure is not without its limitations. One limitation mentioned earlier, is the changing way people connect and potential generational differences. In addition, it is possible to argue that our results may be tied to our use of this particular index of social capital. While we do not have data on other potential alternative measures of social capital at the county level or proxies for social capital for our data periods, we can investigate separately the possible relationship between some of the indicators used to construct the index and our variable of interest, ethnic fractionalization.

Table (8) summarizes the results of estimating our preferred model IV-FE using instead of social capital index, six variables that potentially could be used separately to proxy for social capital. These variables are all a part of the social capital index we use in the above analysis. In Panel A the two proxies for social capital are the number of bowling centers in a county and the number of civic and social associations in a county controlling for population size. In Panel B we proxy for social capital using the total number of golf courses and the total number of religious organizations divided by population size. Finally, in Panel C we proxy using the total number of professional groups divided by population and the response rate to the census in a county.

The results in Table (8) suggest that even with these proxies for social capital which may be less precise, most coefficients on ethnic fractionalization are negative, though many are insignificant. It is worth noting that the effect of ethnic fractionalization is significant in the specification using membership of social associations and provides inferences consistent with our highlighted finding. The results of our robustness checks are consistent with Putnam's conflict or hunker down hypotheses. The consistence of our estimates using component parts of the social capital index suggests that the social capital index, though not perfect, is a good measure and our core findings are not specifically dependent on the derivation of the social capital index.

6.5 Conflict versus Hunker Down Hypotheses

Our findings highlighted above suggest that U.S data is not consistent with the contact hypothesis but is compatible with both the conflict and hunker down thesis. For the conflict hypothesis increases in ethnic diversity leads to lack of trust of those joining the group and a decreased investment in social capital. In contrast the hunker-down notion suggests that an increase in ethnic diversity causes members of a group to distrust not just out-group members, but also in-group members. While our data does not allow direct testing to determine which hypothesis is more relevant in a regression framework,[37] we explore other ways of providing suggestive evidence for or against either of these hypotheses. First, we isolate counties that fall into the lowest quartile of ethnic fractionalization

[37]In an earlier version of our paper (Belton, Huq and Uwaifo Oyelere (2014)), we attempt to test for hunker down hypothesis and find evidence against its relevance within a regression framework.

in the 1990 survey and remain in this quartile in the other survey years. Next, we isolate counties that remain in the top quartile for ethnic fractionalization in the three survey periods. Using simple bar graphs and basic summary statistics, we compare and contrast the trends in social capital in these counties over time. [38] Initially, we note that mean social capital values in the counties with low fractionalization is significantly higher in each year than for counties with high fractionalization. This observation similar to our econometric analysis is not consistent with contact theory but rather consistent with the alternatives. To provide evidence for or against hunker down theory, we compare our 1990 to 2005 values. We do not compare 1990 to 1998 because there is not a substantial change in mean ethnic fractionalization over these two periods. However, there is significant change in mean ethnic fractionalization between 1990 and 2005. In counties with low fractionalization, small increases in ethnic fractionalization cannot significantly impact a communities' cohesion and mean social capital negatively unless this increase leads to a change in interaction among those of the dominant ethnic group. This is because the entry of few people of a different ethnicity into a homogeneous community can only create a limited number of new interactions. Even if these new interactions between the majority group and the new entrants do not lead to trust, they cannot negatively alter the level of social capital within the community in a statistically significant way as long as trust between those of the dominant ethnicity within this community does not change. Hence in counties with low fractionalization, evidence for the hunker-down hypothesis can only be found if small increases in ethnic fractionalization is associated with a statistically significant decrease in mean social capital. Figure (4) indicates that mean social capital increased in these low fractionalization counties between 1990 and 2005 despite entrance of out-group members. This finding is inconsistent with what the hunker down hypothesis will predict. It is important to mention that this increase in social capital was not just a trend effect because when we consider counties with ethnic fractionalization in the top quartile (Figure 5), we find that mean social capital decreased between 1990 and 2005 while ethnic fractionalization increased. One possible argument that could be made is that comparing figure (4) and (5) is not appropriate because counties with high diversity in 1990 could be poorer and should have lower levels of social capital ceteris paribus. It is this idea of a correlation between ethnic fractionalization and income that led to our including income as one of the control variables in our regression analysis and also guided identifying effects using within county versus across county variation.

Figure (6) suggests that the marked difference in social capital between figures (4) and (5) cannot be explained by income. Figure (6) shows mean real per capital income (RPCI) in the counties highlighted in figures (4) and (5). Notice that mean RPCI was actually higher in the most diverse counties in 1990 but increased at a slower rate than the counties which are less diverse. In addition while the least diverse counties now have a higher mean RPCI, the difference in means across both is relatively small and cannot explain the huge gap in means of the social capital index for these two sub-samples.

What do these three figures suggest? While we cannot use a regression framework to test the conflict relative to hunker down thesis, the conflict thesis appears more consistent with U.S data between 1990 and 2005. The evaluation

[38] See Figures (4) and (5).

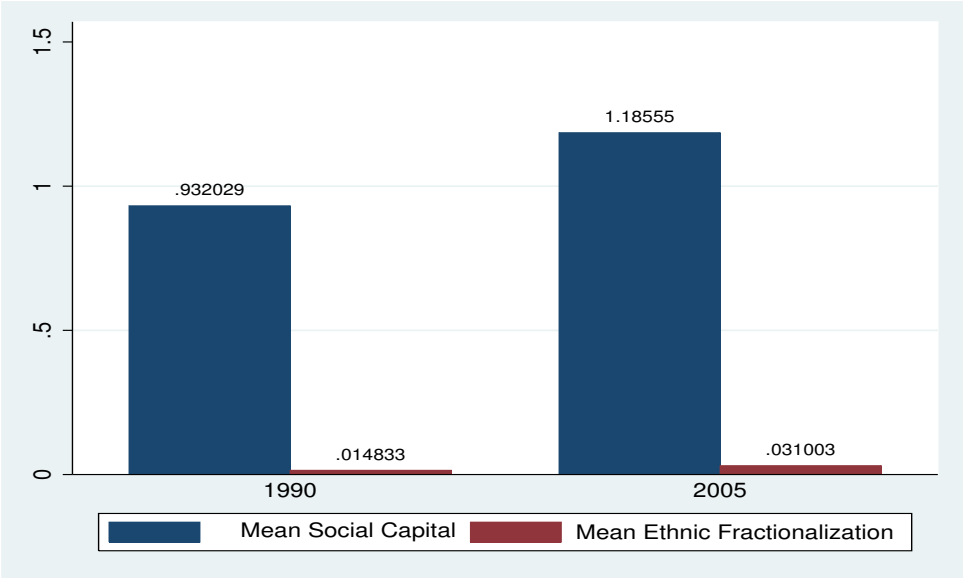

Figure 4: Mean Social Capital vs Ethnic Diversity in counties in the lowest quartile of Ethnic Fractionalization Overtime

of the least and most diverse counties across the 1990 to 2005 time period provides prima facie evidence against Putnam's hunker-down notion. These results suggest that conflict hypothesis seems more consistent with the patterns in our data i.e., as communities become more diverse, there is less trust within the community stemming from less trust for those in the out-group or those of other ethnicities. This lack of trust leads to less willingness to invest in social goods leading to a decline in social capital.

7 Conclusion and Inferences

In this paper we first test for time differences in the impact of ethnic fractionalization on social capital accumulation using OLS and IV methods. Subsequently, we check to see if any of theories of social interaction discussed in Putnam (2007) is consistent with our data using FE, IV and IV-FE methods. Our results suggest that though ethnic fractionalization has increased over time in the U.S., there is no significant change in the negative impact of ethnic fractionalization on social capital. This result suggests that the relationship between social capital and ethnic fractionalization is stable over the evaluation period (1990-2005).

What does our results suggest about the claims highlighted in Putnam (2007)? We find evidence against the *contact hypothesis* given the identification of a negative relationship between ethnic diversity and social capital. Our results are consistent with the *conflict hypothesis* which suggests that an increase in diversity leads to a decline in social capital. However, the magnitude of the estimated impact of ethnic fractionalization is small. Specifically, our preferred model (IV-FE) [Table (8) columns (2)], reveals a coefficient of (-3.318) implying

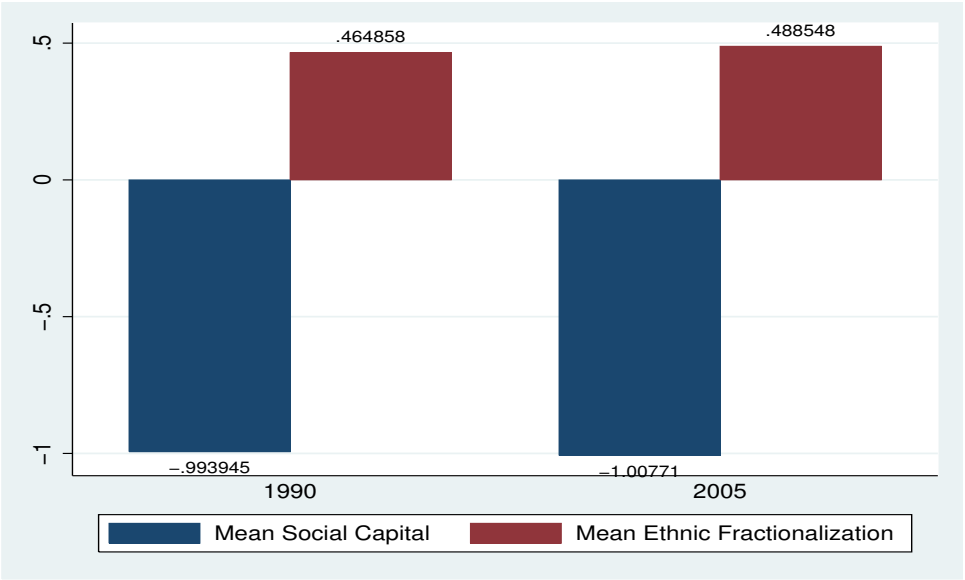

Figure 5: Mean Social Capital vs Ethnic Diversity in counties in the highest quartile of Ethnic Fractionalization Overtime

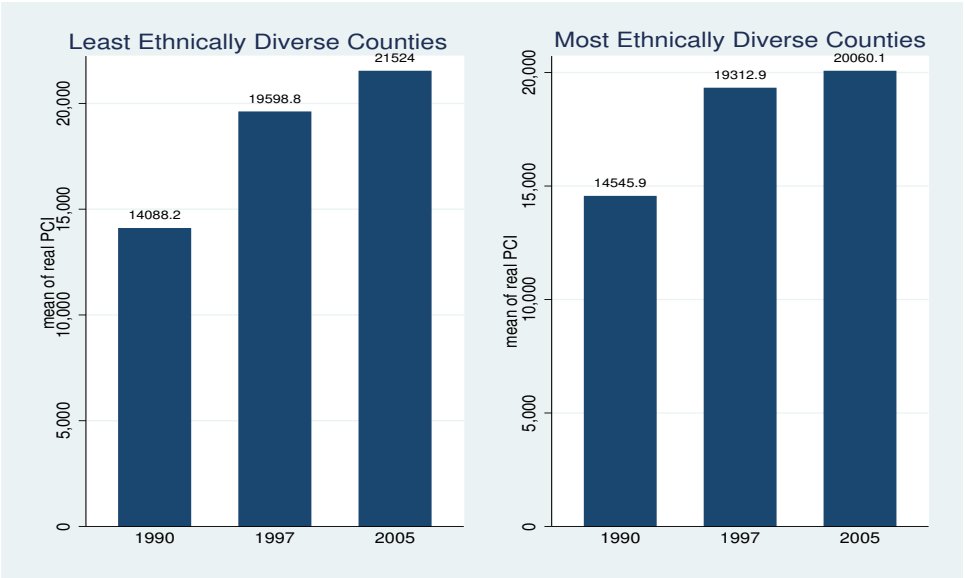

Figure 6: Mean Real Per Capita Income in top and bottom quartiles overtime

Table 5: Testing for Time Differences: OLS Regression and IV Results

	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
	1990	1997	2005	1990	1997	2005
Ethnic Fractionalization	-0.427***	-0.457***	-0.497**	-0.344*	-0.462**	-0.591**
	(0.136)	(0.163)	(0.196)	(0.191)	(0.218)	(0.244)
Income Per Capita	-0.000	0.000	-0.000***	-0.000	0.000	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education	0.033***	0.038***	0.022***	0.033***	0.038***	0.022***
	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)	(0.005)
Family Households	-0.045***	-0.055***	-0.051***	-0.044***	-0.055***	-0.051***
	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
Median Age	0.477***	0.236***	0.132**	0.478***	0.236***	0.131**
	(0.055)	(0.055)	(0.059)	(0.054)	(0.054)	(0.058)
Square of Median Age	-0.005***	-0.002***	-0.000	-0.005***	-0.002***	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Urban	-0.570***	-0.435***	-0.871***	-0.571***	-0.435***	-0.873***
	(0.138)	(0.156)	(0.231)	(0.137)	(0.155)	(0.229)
Population Density	-0.000***	-0.000**	-0.000**	-0.000***	-0.000**	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Residence Time	0.000	-0.000	-0.000	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Working Women	0.028**	0.033***	0.049***	0.028**	0.033**	0.048***
	(0.012)	(0.013)	(0.016)	(0.012)	(0.013)	(0.016)
Constant	-7.855***	-3.320**	-2.342*			
	(1.160)	(1.297)	(1.384)			
Observations	3108	3106	3106	3108	3106	3106
Adjusted R^2	0.708	0.653	0.605	0.300	0.301	0.241

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Testing Contact Hypothesis: OLS, FE and 2nd stage IV Results

Dependent Variable: Social Capital	Other Specifications (1) OLS	(2) Fixed Effect	2nd Stage: IV Analysis (3) Preferred Instruments	(4) IV 2 Democratic Vote	(5) IV 3 Lead Vote
Ethnic Fractionalization	-0.424***	-0.891**	-0.453**	-3.334***	-9.029***
	(0.140)	(0.391)	(0.186)	(0.537)	(1.516)
Income Per Capita	-0.000***	0.000***	-0.000***	-0.000	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education	0.028***	0.011	0.028***	0.029***	0.031***
	(0.003)	(0.007)	(0.003)	(0.003)	(0.004)
Family Households	-0.049***	-0.005	-0.049***	-0.058***	-0.074***
	(0.004)	(0.008)	(0.004)	(0.004)	(0.007)
Median Age	0.155***	-0.301***	0.154***	0.129***	0.079
	(0.035)	(0.044)	(0.035)	(0.036)	(0.051)
Square of Median Age	-0.001*	0.004***	-0.001*	-0.001**	-0.001**
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)
Urban	-0.646***	-1.003***	-0.646***	-0.641***	-0.632***
	(0.140)	(0.273)	(0.139)	(0.141)	(0.153)
Population Density	-0.000**	-0.000	-0.000**	0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Residence Time	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Working Women	0.040***	0.029**	0.040***	0.026**	-0.003
	(0.009)	(0.013)	(0.010)	(0.010)	(0.013)
Year (1997)	-0.561***	-0.077*	-0.559***	-0.384***	-0.036
	(0.026)	(0.041)	(0.028)	(0.043)	(0.098)
Year (2005)	-1.006***	-0.226***	-1.003***	-0.683***	-0.049
	(0.040)	(0.064)	(0.043)	(0.074)	(0.178)
Constant	-1.911**	6.176***	-1.882**	0.897	6.392***
	(0.893)	(1.271)	(0.901)	(1.047)	(1.889)
Observations	9320	9320	9320	9320	9320
Adjusted R^2	0.627	0.070	0.627	0.586	0.271

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Testing for Evidence of Putnams's Theories

	IV-Fixed Effects Estimates 2nd Stage Estimates			Evidence of Instrument Validity		
Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)
Social Capital	IV 1	IV-2	IV-3	OLS	OLS	OLS
Ethnic Fractionalization	-3.318***	3.834	-26.393***	-0.402**	-0.400**	-0.270
	(1.257)	(3.420)	(6.228)	(0.190)	(0.191)	(0.195)
Black Proportion				-0.036	-0.028	0.184
				(0.130)	(0.131)	(0.140)
Gini Coefficient					-0.278	-0.258
					(0.632)	(0.643)
Democratic Vote						-0.007***
						(0.001)
Leading Party Vote						0.011***
						(0.001)
Income Per Capita	0.000***	0.000***	0.000	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education	0.014**	0.003	0.050***	0.028***	0.028***	0.027***
	(0.007)	(0.008)	(0.013)	(0.003)	(0.003)	(0.003)
Family Households	-0.003	-0.010	0.020*	-0.049***	-0.050***	-0.056***
	(0.008)	(0.008)	(0.012)	(0.004)	(0.004)	(0.005)
Median Age	-0.319***	-0.266***	-0.489***	0.155***	0.154***	0.166***
	(0.044)	(0.051)	(0.075)	(0.035)	(0.035)	(0.035)
Square of Median Age	0.004***	0.004***	0.005***	-0.001*	-0.001*	-0.001**
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Urban	-1.012***	-0.985***	-1.102***	-0.646***	-0.647***	-0.560***
	(0.274)	(0.274)	(0.332)	(0.140)	(0.140)	(0.142)
Population Density	-0.000	-0.000	-0.000	-0.000**	-0.000**	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Residence Time	-0.000*	-0.000	-0.000*	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Working Women	0.022	0.042**	-0.041**	0.040***	0.040***	0.036***
	(0.014)	(0.017)	(0.020)	(0.009)	(0.010)	(0.010)
Year (1997)	0.032	-0.289*	1.066***	-0.562***	-0.557***	-0.491***
	(0.069)	(0.157)	(0.278)	(0.026)	(0.029)	(0.030)
Year (2005)	-0.036	-0.595**	1.769***	-1.008***	-1.001***	-1.100***
	(0.118)	(0.271)	(0.487)	(0.040)	(0.043)	(0.043)
Constant				-1.902**	-1.717*	-2.014**
				(0.894)	(0.992)	(1.016)
Observations	9319	9319	9319	9320	9320	9320
Adjusted R^2	-0.409	-0.446	-1.856	0.627	0.627	0.633

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Robustness Check: Alternative Measures of Social Capital

	(1) (IV-FE)	(2) (IV-FE)
	Panel A	
	Number of civic groups	Number of bowling places
Ethnic Fractionalization	-0.000183* (0.000)	-0.000037 (0.000)
Observations	9319	9319
	Panel B	
	Number of religious organizations	Number of golf courses
Ethnic Fractionalization	-0.000802*** (0.000)	-0.000054 (0.000)
Observations	9319	9319
	Panel C	
	Census response rate	Number of Professional organizations
Ethnic Fractionalization	-0.003005 (0.005)	-0.000045 (0.000)
Observations	8563	9319

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Note: This table provides a summary of 8 FE-IV estimations.

that a 100% change in ethnic fractionalization would lead to a 3.318 decline in the social capital index. A reasonable change in ethnic fractionalization is about 4% (which is the change in the mean between 1990 and 2005) and this would lead to a change of approximately 0.13 in social capital.[39] Our evaluation of means for counties in the lowest and highest quartile of ethnic fractionalization suggests that Putnam's *hunker-down*thesis does not appear consistent with our data.

It is important to mention that this analysis has one main limitations. First, we consider social capital using pre-existing measures. Though consistent with the way social capital was formed in the past, it is possible to argue that these measures may not be as relevant for social capital formation for younger generations. Those born in the 70s and the Millennials have grown-up with a huge internet presence, advanced communication technologies, and social media. They are networking and forming groups in ways that could be quite different from older generations and it is reasonable to assume these groups will also create social capital. Hence, our measure of social capital could be downward biased for the younger cohorts. However since the period of data covered in our study is 1990-2005 which is before the explosion of social media, this argument may not be as relevant in our study. Notice in Table 8 column (1) that the effect of age is negative not positive which goes against what one would expect if younger cohorts were less likely to accumulate social capital than older cohorts, in ways captured by our data.

In summary, our results suggest that between 1990 to 2005, as communities evolved in the U.S, and ethnic fractionalization increased, individuals' trusted new entrants less which lead to less social capital accumulation. However, this cost of increased immigration and migration, which are the primary sources of increase in ethnic diversity in communities, could be far out weighed by the documented private benefits of immigrant and migrant increases within communities in the U.S. Hence, though diversity may lead to a decrease in social capital as captured by our measures, our results do not provide evidence of the

[39] A change of 0.13 is a small change in our social capital measure.

impact of diversity on other welfare outcomes or alternative measures of social capital. Given the nature of our finding, further studies are needed to determine whether the positive effects of diversity and immigration outweighs the potential negative effects on social capital. In addition, studies understanding why increased diversity seems to lead to decreased trust are useful. Moreover, coming up with initiatives and policies to increase social capital is useful for future societal health. There is also a need for a more comprehensive measure of social capital that includes in its calculation, the types of groups and networks to which the younger cohorts belong. Finally, it is important to reiterate that our results capture the relationship between social capital and ethnic fractionalization between 1990 and 2005. A lot has changed in the last 12 years that could affect the way different groups view each other and interact. In particular U.S elected its first non-White President in 2008 which could be an indicator of a change in interactions across group. In future research, we hope to reexamine the relationship between social capital and ethnic fractionalization post 2006.

References

[1] Allport, G. W. (1954). The nature of prejudice. Addison-Wesley, Cambridge, Massachusett

[2] Angrist, J. D., and G. W. Imbens,(1995). Two–Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity. Journal of the American Statistical Association, 90, 430, 431-442.

[3] Alesina, Alberto and Eliana LaFerrara. (2004). Ethnic Diversity and Economic Performance. NBER Working Paper 10313, National Bureau of Economic Research.

[4] Alesina, Alberto and Eliana LaFerrara. (2000). Participation in Heterogeneous Communities. Quarterly Journal of Economics, 115, 847-904.

[5] Arrow, Kenneth J. (2000). Observations on Social Capital. Social Capital: A Multifaceted Perspective. Edited by Partha Dasgupta and Ismail Serageldin. The International Bank of Reconstruction.

[6] Belton, W., Huq, Y., & Uwaifo Oyelere, R. (2014). Diversity and Social Capital in the US: A Tale of Conflict, Contact or Total Mistrust? (No. 8384). IZA Discussion Papers.

[7] Blau, Peter M. (1977). *Inequality and Heterogeneity: A Primitive Theory of Social Structure*. New York: Free Press.

[8] Blummer H. (1958). *Race Prejudice as a of Group Position*, Pacific Sociological Review, Spring, Vol.1, Number 1.

[9] Borjas, George, (2013). Immigration and the American worker. Center for Immigration Studies, Washington, DC . Backgrounder.

[10] Bourdieu P. (1987). Distinction. A Social Critique of the Judgement of Taste. London, Routledge

[11] Browne, W. (2001). The Failure of National Rural Policy. Georgetown University Press, Washington D.C.

[12] Card, David. (2005). Is the New Immigration Really So Bad? *Economic Journal*, 115, 300–323.

[13] Chhibber, Ajay, et al. (1997). The State in a Changing World. *World Development Report*, 1997. Oxford University Press, Inc., 2001 Evans Rd., Cary, NC 27513, 1997.

[14] Colman J.S. (1988). The Creation and Destruction of Social Capital: Implications for the Law. Notre Dame *Journal of Law, Ethics, and Public Policy* 3:375-404.

[15] Colman J.S. (1990). *Foundations of Social Theory*, Cambridge: Belknap Press of Harvard University Press (1990)

[16] Dasgupta, Patha (2009). A Matter of Trust: Social Capital and Economic Development. Manuscript May 2009. Presented June 2009, Annual Bank Conference of Development Economics

[17] Dasgupta, Patha (1988). Trust as a Commodity. D. Gambetta ed. *Trust : Making and Breaking Cooperative Relations* (Oxford:Basil Blackwell)

[18] Diette, Timothy M. and Ruth Uwaifo Oyelere, (2014). Gender and Race Heterogeneity: The Impact of Students with Limited English on Native Students' Performance, *American Economic Review*, American Economic Association, 104(5), 412-17, May.

[19] DeFilippis, J. (2001). The Myth of Social Capital in Community Development *Housing Policy Debate* 12, 781-806.

[20] Durlauf, S.N. (2002). Symposium on social capital: introduction. *Economic Journal*, 112, 437-458.

[21] Glaeser E.L., Laibson, D., Scheinkman, J.A. (2002). The Economic Approach to Social Capital. *Economic Journal*, 115, 417–18.

[22] Gundelach, Birte. (2014). In diversity we trust: the positive effect of ethnic diversity on outgroup trust. Political behavior, 36.1, 125-142.

[23] Helliwell, John F., and Robert D. Putnam. (2004). The social context of well-being. Philosophical transactions-royal society of London series B biological sciences: 1435-1446.

[24] Knack, S. and Keefer, P. (1997). Does Social Capital have an Economic Payoff? A Cross-Country Investigation. *Quarterly Journal of Economics* 112, 1251-1288.

[25] McPherson, J., Popielarz, P. A., and Drobnic, S. (1992). Social Networks and Organizational Dynamics. *American Sociological Review*, 57(2), 153-170.

[26] Meer, T. van der and Tolsma, J. (2014). Ethnic diversity and its effects on social cohesion. Annual Review of Sociology, 40, 459-478.

[27] Narayan, Deepa, and Lant Pritchett. (1997). Cents and Sociability: Household Income and Social Capital in Rural Tanzania.

[28] North, Douglas C. and Robert P. Thomas. (1973). The Rise of the Western World: A New Economic History. Cambridge: Cambridge University Press.

[29] Olson, Mancur. (1971). The Logic of Collective Action: Public Goods and the Theory of Groups. (Revised edition ed.). Harvard University Press. 1965.

[30] Ottaviano, Gianmarco IP, and Giovanni Peri. (2012). Rethinking the effect of immigration on wages. *Journal of the European Economic Association* 10.1 (2012): 152-197.

[31] Pichler, F. Wallace (2007) Patterns of Formal and Informal Social Capital. *European Sociological Review*, 24, 423-35.

[32] Portes, Alejandro.(1998) Social Capital: Its Origins and Applications in Modern Sociology, *Annual Review of Sociology*, 24, 1-24

[33] Putnam, Robert (1993). Making Democracy Work: Civic Traditions in Modern Italy. *Princeton University Press*

[34] Putnam, Robert (1995). Bowling Alone: America's Declining Social Capital. *Journal of Democracy* 6 (1): 65–78.

[35] Putnam, Robert (2000). Bowling Alone: The Collapse and Revival of American Community. New York,Simon and Schuster.

[36] Putnam, Robert E. Pluribus Unum: (2007). Diversity and Community in the Twenty-first Century The 2006 Johan Skytte Prize Lecture. *The Scandinavian Political Studies*, 30, 2: 137-174.

[37] Rupasingha, Anil, Stephan J. Goetz, and David Freshwater. (2006). The production of social capital in US counties. *The Journal of Socio-economics* 35.1, 83-101.

[38] Quillian, L. (1995). Prejudice as a Response to Perceived Group Threat: Population Composition and Anti-immigrant and Racial Prejudice in Europe. *The American Sociological Review* 60: 586-611.

[39] Rupasingha, Anil and Stephan J. Goetz. US County-Level Social Capital Data, 1990-2005. The Northeast Regional Center for Rural Development, Penn State University, University Park, PA, 2008.

[40] Solow, Robert. (2000). Notes on Social capital and Economic Performance. Social Capital: A Multifaceted Perspective. Edited by Partha Dasgupta and Ismail Serageldin. The International Bank of Reconstruction.

[41] Solow, Robert M. (February 1956). A contribution to the theory of economic growth. Quarterly Journal of Economics. Oxford Journals. 70 (1): 65–94.

[42] Swan, Trevor W. (November 1956). Economic growth and capital accumulation. Economic Record. Wiley. 32 (2): 334–361.

[43] Stock, J. H., and Yogo, M. (2005). Testing for weak instruments in linear IV regression. Identification and inference for econometric models: Essays in honor of Thomas Rothenberg.

[44] U.S. Census Bureau. (2010). USA Counties- Demographic data. [Table]. Retrieved from http://www.censtats.census.gov/usa/usa.shtml

Appendix

Appendix A:

Table A1: First Stage IV Results (Dependent Variable -Ethnic Fractionalization)

	(1) 1990 IV 1	(2) 1997 IV 1	(3) 2005 IV 1	(4) Pooled IV 1	(5) Pooled IV 2	(6) Pooled IV 3
Black Proportion	0.759***	0.693***	0.711***	0.726***		
	(0.030)	(0.030)	(0.029)	(0.017)		
Gini Coefficient	0.393***	0.167*	0.040	0.128***		
	(0.108)	(0.086)	(0.083)	(0.049)		
Democratic Vote					0.003***	
					(0.000)	
Percent of Presidential Vote for Leading Party						-0.002***
						(0.000)
Income Per Capita	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education	-0.001**	-0.001**	0.002***	0.002***	0.000**	0.000**
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Family Households	-0.000	-0.001	0.002***	0.001***	-0.001***	-0.002***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Urban	0.012	0.005	-0.006	0.007*	-0.011**	-0.007
	(0.008)	(0.008)	(0.008)	(0.004)	(0.005)	(0.005)
Population Density	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	3108	3106	3106	9320	9320	9320
P-value Hansen test	0.2549	0.0141	0.6335	0.8711	NA	NA
Adjusted R^2	0.835	0.828	0.823	0.827	0.710	0.690

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
IV 1 Are our preferred Instruments while IV2 and IV3 are alternative instruments.
Note* We also controlled for median age, square of median age, residence time and share of working women (coefficients not included).

Appendix B: Building a Case for the Validity of the Instruments

First, evidence of the relevance of both instruments can be found in our first stage estimates in Table (A1) in our appendix columns(1)-(3). These results reveals a strong positive relationship between ethnic fractionalization and both the Black share of the population in the county and the Gini coefficient in the county. Notice in three of the four cases, the sign on the estimated impact of Gini and the share of Blacks in a county is positive and significant at least at the 10 percent level.

We are also confident in our choice of instruments as our preferred instruments Black share and Gini coefficient pass the Hansen J test for overidentification restriction in the 1990 and 2005 surveys (suggesting that at least one of them is valid). Our instruments also performed well on other tests for weak instruments. We mention some these tests earlier in the paper but do not present

weak instrument test statistics to restrict table length to a page. These test results are available on request.

Although we posit that these instruments satisfy exclusion restrictions, a possible source of concern is whether both instruments are valid. While we argue that our instruments are valid, the over-identification test only suggests that at least one instrument is valid. We provide evidence of the lack of independent effect of our instruments on social capital by estimating a pooled regression of social capital on ethnic fractionalization, while including our preferred instruments and other controls. We provide these results in columns (4) and (5) of Table (7). Notice that both the share of Blacks and Gini coefficient are not significant which provides evidence they have no direct effect on social capital. Instead, the only way our instruments impact social capital is through their effect on ethnic fractionalization. Notice in column (4) of Table (7) that Black share has no significant impact on social capital once ethnic fractionalization is included in the model. In column (5) we add the Gini coefficient and again notice that neither Black share nor Gini have a significant impact on social capital.If we exclude ethnic fractionalization from the regression and only included Black share or Gini we find a significant effect of these variables which is evidence of the correlation between these variables and ethnic fractionalization. While these results do not prove validity, they provide strong evidence in favor of our instruments.

Appendix C: Description of Independent Variables used in Analysis

The ethnic fractionalization variable measures the level of diversity within a county.

$$Ethnic\ Fractionalization = 1 - \sum_i (Race\ Proportion_i)^2 \mid i = Black, White, Asian, Latino$$

where the race proportion is the proportion of a population that identifies with race (i), which can be Black, White, Asian or Latino. This is the same method used in Rupasingha et al (2006).

The racial dominance variable identifies the proportion of the majority race in a given community. Larger values imply a more dominant race within a county. Racial dominance is defined as follows:

$$Racial\ Dominance = max(Race\ Variable_i \mid i = Black, White, Asian, Latino).$$

Racial dominance is defined across four racial groups, Black, White, Latino and Asian.

The race proportions is calculated by dividing the population of race i by the total number of Blacks, Whites, Latinos, and Asians in a particular county.

$$Race\ Proportion_i = \frac{Race\ Population_i}{Total\ Population}$$

Ethnic fractionalization, racial dominance, and racial proportion data were all derived from 1990, 2000, and 2005 for the 1990, 1997, and 2005 panels,

respectively. Income per capita was collected from 1989, 1999, and an average from 2005-2009 for the 1990, 1997, and 2005 panels, respectively. Education level was collected from 1990, 2000, and an average from 2005-2009 for the 1990, 1997, and 2005 panels, respectively.

The proportion of family households was derived from components gathered from 1990, 2000, and an average from 2005-2009 for the 1990, 1997, and 2005 panels, respectively. The calculation for this variable is shown below:

$$Family\ Households = \frac{Total\ Number\ of\ Family\ Households}{Total\ Number\ of\ Households}$$

This represents the proportion of households that are family units.

Median age data was taken from the U.S. census of 1990, 2000, and 2010 for the 1990, 1997, and 2005 panels, respectively. We are forced to use median age data from 2010 rather than using 2000 census data for both 2000 and 2005 panels.

The determination of whether a county is urban was based on population data collected from the 1990, 1997, and 2005 time periods for their respective panels.

The calculation for population density is as follows:

$$Population\ Density = \frac{Total\ Population}{Area}$$

Total population data was collected from 1990, 1997, and 2005 time periods for the respective panels. Land area was collected from 1990, 2000, and 2010 for the 1990, 1997, and 2005 panels, respectively.

Residence time was approximated using an engineering equation to calculate the time a particle spends in a specified vessel. An assumption behind this equation is that a given town has the capacity for a set amount of people, which is more accurate over a short period of time. The equation is as follows:

$$Residence\ Time = 10 * \frac{Total\ Population}{\Delta Population_{10\ years}}$$

The 10 year change in population was collected from change from 1980-1990, 1990-2000, and 2000-2010 for the the 1990, 1997, and 2005 panels, respectively.

The percentage of women working in the overall population was calculated with the following equation:

$$Working\ Women = 100 * \frac{Total\ Number\ of\ Working\ Women}{Total\ Population}$$

Data on the total number of working women in a county was collected from components gathered from 1990, 2000, and an average from 2005-2009 for the 1990, 1997, and 2005 panels, respectively.