

ICPSR BULLETIN

Volume XXXI, No. 4

Special Edition 2011



2010-2011 ICPSR Undergraduate Research Paper Competition First Place Winner

Do More Parties Make for Happier Voters?

**Democratic satisfaction and party representation across thirty-six
democracies**

Tommaso Pavone, University of Michigan ... p. 1

2010-2011 ICPSR Master's Research Paper Competition First Place Winner

The Long-Term Effects of Colonial Land Tenancy:

Micro Evidence from India

Sayon Deb, Boston University ... p. 23

2010-2011 Resource Center for Minority Data Research Paper Competition First Place Winner

**Educational Outcomes for Latino Immigrants in Los Angeles County:
Importance of Gender, Immigrant Generation and Mother's Education**

Whitney Boyer, Washington University in St. Louis ... p. 51

INTER-UNIVERSITY CONSORTIUM FOR POLITICAL AND SOCIAL RESEARCH

George Alter, Acting Director
Mary Vardigan, Assistant Director
Diane Winter, Assistant Director for
Administration

COUNCIL MEMBERS

Ann Wolpert, Chair
Massachusetts Institute of Technology

Francine Berman
Rensselaer Polytechnic Institute

G. Sayeed Choudhury
Johns Hopkins University

Paul N. Courant
University of Michigan

Catherine A. Fitch
University of Minnesota

Thomas LaVeist
Johns Hopkins University

Jeffrey Moon
Queen's University

Gregory N. Price
Morehouse College

Rogelio Saenz
Texas A&M University

Barbara Schneider
Michigan State University

Lori M. Weber
California State University, Chico

Christopher Zorn
Pennsylvania State University

Aletha C. Huston, Past Chair
University of Texas, Austin

BULLETIN STAFF

Editor: Dan Meisler
dmeisler@umich.edu
(734) 615-7904

ICPSR is a center within the Institute
for Social Research at the University of
Michigan.

From Acting Director George Alter

We are very proud to present the first-place winners of the 2011 ICPSR Research Paper Competitions in this special edition of the *Bulletin*. The submissions this year showed a great deal of creativity and a high level of scholarship.



Tommaso Pavone of the University of Michigan submitted the winning paper in the Undergraduate Competition with an analysis of whether more political parties result in greater satisfaction among voters. He used data from the Comparative Study of Electoral Systems held at ICPSR in his paper.

Sayon Deb of Boston University won first place in the Master's Competition for a paper examining the long-term effects of colonial land tenancy in India. The analysis used data from the 2005 Indian Human Development Survey.

The winner of the Resource Center for Minority Data Competition was Whitney Boyer of Washington University in St. Louis, who used data from the Los Angeles Family and Neighborhood Study to analyze factors affecting the educational outcomes of Latino immigrants.

Congratulations to the winners, and thanks to all entrants for their hard work!

About ICPSR

The Inter-university Consortium for Political and Social Research (ICPSR) is an international consortium of about 700 academic institutions and research organizations. ICPSR provides leadership and training in data access, curation, and methods of analysis for the social science research community.

ICPSR maintains a data archive of more than 500,000 files of research in the social sciences.

ICPSR's educational activities include the Summer Program in Quantitative Methods of Social Research, a comprehensive curriculum of intensive courses in research design, statistics, data analysis, and social methodology. ICPSR also leads several initiatives that encourage use of data in teaching, particularly for undergraduate instruction.

ICPSR-sponsored research focuses on the emerging challenges of digital curation and data science. ICPSR researchers also examine substantive issues related to our collections, with an emphasis on historical demography and the environment.

For more information, please visit www.icpsr.umich.edu.

Do More Parties Make for Happier Voters?

Democratic satisfaction and party representation across thirty-six democracies

Tommaso Pavone¹

B.A. Public Policy, Dec. 2010

Professor Allen Hicken

December 7th, 2010

Submitted for consideration for the 2011 ICPSR Undergraduate Research Paper Competition

Abstract

Extensions of Downsian theory of party competition imply that the greater the number of political parties, the greater the perception of representation and satisfaction with the democratic process. However, this logic has been subjected to increasing scrutiny. This paper conducts a cross-national analysis of public opinion data from thirty-six democracies to assess whether a) feeling represented by a party increases democratic satisfaction, and b) whether more parties induce a greater sense of party representation. Multivariate regression results find that feeling represented by a party correlates with greater satisfaction with the democratic process. The more striking results emerge when testing the relationship between the number of parties (as measured by the effective number of electoral parties (ENEP) and parliamentary (ENPP) parties) and party representation. Namely, although more parliamentary parties correlate with greater perceived representation, the opposite is true for the number of electoral parties. This implies that more parties do not necessarily make for happier voters.

¹ Contact info: tpavone@umich.edu. I am grateful to Allen Hicken for his feedback and thoughts regarding this paper. The usual conventions apply, and all errors are my own.

1. Introduction and paper overview

Citizens in modern democracies are represented through and by parties. This is inevitable.
Sartori (1968: 471).

This paper analyzes one of the most fundamental *raison d'être* of political parties; the function of representing people's preferences. The analysis is also extended to the party system level to see whether some types of party systems, namely pluralist ones, are more representative of the people's preferences than others. This is assessed through public opinion data, largely because, whether parties represent the electorate or not, the *perception* of representativeness in the eyes of the public can seriously effect support for democracy. The analysis conducted uses a large N dataset spanning across thirty-six democracies and thirty-eight elections held between 2001 and 2006. Two primary questions are explored. First, are voters who feel represented by a political party more likely to be satisfied with the democratic process? Finally, do more parties engender a greater sense of party representativeness amongst voters?

The paper is divided into six sections. Section 2 provides both a literature review and some hypotheses to test the above questions. It first provides a cogent overview of the literature on party representation and democracy, developing a hypothesis that voters who feel represented are more likely to feel satisfied with the democratic process. It then moves to consider, and strongly scrutinize, Downsian theory and its implications for party system representativeness, deriving the second hypothesis, namely that there is no relationship between the number of parties and the perception of party representativeness. Section 3 moves to discuss the dataset compilation and coding processes. Section 4 conducts multivariate regression analysis to test the hypotheses in the paper, finding support for the first hypothesis and mixed support for the second. More specifically, the results strikingly find that increasing the number of electoral parties decreases the sense of party representation amongst voters, whereas increasing the

number of parliamentary parties has the opposite effect. Finally, section 5 provides some possible interpretations for the regression results and section 6 concludes.

2.1 Party representation and democracy

Political parties were not always viewed as beneficial or necessary by political scholars. Early political philosophers, including such notables as Machiavelli, Madison, and Montesquieu, could not conceptualize of a party as being anything more than a faction. Parties were seen, in the words of Bolingbroke, as “a political evil, and faction is the worst of all parties” (Sartori 1976: 5). Yet with time, political parties gradually began to be viewed “not as a weed but as a necessary microbe lodged deep in the digestive tract – not pretty, but vital to keeping the body politic in good health” (Stokes 1999:244). Indeed, many contemporary scholars now argue for the necessity of parties. To this end, Schattschneider (1942) famously argued that “political parties created democracy...modern democracy is unthinkable save in terms of the parties.”

One of the most oft-cited reasons why political parties are increasingly perceived as necessary for democracy is their representation function. Sartori (1976) writes that “parties are channels of expression. That is to say, parties belong, first and foremost, to the means of representation: They are an instrument, or agency, for representing the people by expressing their demands.” Similarly, Hirshman (1970) characterizes parties as providing a “voice function” for voters. Nuemann (1956) discusses how parties are able to organize the public’s preferences and to then channel such preferences to the governmental level. Key (1961) considers parties to be the primary tools to translate public opinion into public policy, and Lipset and Rokkan (1990) discuss how parties represent different social cleavages through their expressive function. These

arguments highlight that parties are conditioned by society, serving a “bottom up” function that diffuses popular preferences to the policymaking process (Sartori 1969; Stokes 1999).

In short, political parties create an institutionalized linkage between government and citizens, engendering a government that is responsive to the preferences of the people (Dix 1992; Stokes 1999). This is due to political parties’ ability to aggregate the preferences of voters, to codify them into political platforms, to advocate for them on the campaign trail, and to implement them once in office (Sartori 1976). This matters because governmental responsiveness is often considered the primary characteristic that sets democracy apart from alternative forms of government (Key 1961; Dahl 1971; Sartori 1976; Dalton 1985).

The foregoing clearly supports the notion that parties’ representation function is beneficial. Therefore, voters who feel represented by a political party should be more satisfied with the democratic process. This inference has been subjected to relatively little skepticism, especially amongst contemporary scholars, and is thus incorporated in the first hypothesis:

Hypothesis I: *Voters who feel represented by a political party are more likely to be satisfied with the democratic process than voters who do not feel represented by a political party.*

2.2 Party representation and the number of parties

If there is empirical evidence supporting Hypothesis I, then the natural question to ask is “what factors contribute to feeling represented by a political party?” The literature linking party representation to party system characteristics is strongly grounded in the spatial theory resulting from the work of Anthony Downs. Extending theories of spatial economic competition first developed by Hotelling (1929), Downs (1957) modeled the preferences of voters in one-dimensional space and showed that, given uniformly or normally distributed voter preferences, vote-seeking political parties have an incentive to converge upon the ideal point espoused by the

median voter. Downs' spatial theory was later extended to a multidimensional Euclidean space and to cases of multiparty competition (see Davis et al. 1970; McKelvey 1986).

The basic notions of the spatial model of competition are that individuals and parties are rational, that voters hold complete and transitive preferences, and that they will vote for the party whose platform is closest to their policy ideal point (Ferejohn 1995; Stokes 1999; Tsebelis 2002). The incentive for the vote-seeking party is, therefore, to adopt policy positions that minimize the distance to the policy ideal points of the largest group of voters possible (Stokes 1999). Note that what matters is not the direction (i.e. whether the party platform is more conservative or more liberal than the voters' preferences), but simply the proximity between the respective policy positions of party and voters (Achen 1978; Dalton 1985: 280).

By extending this logic to the party system level, spatial theory implies that the greater the number of parties, the greater the sense of representation. This is because, *ceteris paribus*, increasing the number of parties should reduce the proximity between voters and party platforms. This appears both logical and theoretically parsimonious. However, scholars have directly challenged many of the simplifying assumptions that underlie this logic.

Spatial theory assumes that parties are unitary actors which espouse a single platform. Indeed, both Downs (1957) and Black (1958) characterized political parties as "teams" (Stokes 1999). Further, these unitary actors are able to move freely across policy space, minimizing their platform's proximity to the greatest amount of voters without cost (Stokes 1999). Finally, citizens are also assumed to have access to perfect information regarding the policy positions of parties, to be aware of their own policy preferences, to be able to compare the platforms of parties to their own ideal points. But clearly, these assumptions are far from realistic.

Przeworski and Sprague (1986) have criticized the notion that voters hold exogenous policy preferences and that parties always seek to minimize their proximity to the greatest number of voters. They argue that by adopting more extreme positions than voters, parties can shape both their political identities and values, thereby pulling public opinion towards their respective party platforms. This occurs because activists can often drive the party agenda (Aldrich 1983; Tsebelis 1990). The implications are that parties do not always minimize proximity to the greatest number of voters and that there exists a ‘tug of war’ of sorts between both ideological party activists and party pragmatists and political parties and the electorate (Aldrich 1983; Strom and Mueller 1995). This fundamentally challenges the notion that parties are unitary actors that can move through ‘frictionless’ policy space. Rabinowitz and MacDonald (1989) also reach similar conclusions. They argue that because voters do not have fully formed preferences, they are attracted to clear and somewhat more extreme messages. Thus, parties have another incentive to adopt policy positions that are more extreme than the voters they target.

There exists empirical evidence supporting the foregoing notions. Inglehart (1984) and Dalton (1985) conducted separate analyses of the first wave of elections to the European Parliament, and both found that candidates generally adopted more extreme economic policies than the majority of voters. Similar patterns surfaced through studies of Sweden (Holmberg 1989), Norway and the United States (Listhaug et al. 1990) and cross-country analyses (Rabinowitz et al. 1991; Iversen 1994). This supports the assertion that parties tend to adopt policy positions that are more extreme than the majority of the electorate.

However, even this conclusion may not be robust. In the words of Lipset and Rokkan (1990), “no party can hope to gain decisive influence...without some willingness to cut across existing cleavages to establish common fronts with potential enemies.” Indeed, in societies

where social cleavages are prominent, political parties must be more moderate than their bases in order to avoid deadlock in government. This argument is found in the consociational literature (Lijphart 1984; Tsebelis 1990; Lijphart 1999). This exposes another flaw in traditional spatial assumptions, namely that parties are solely vote-seeking. As Tsebelis (1990) argues, political parties are constantly balancing their strategies in *both* the electoral and the policymaking arenas.

In short, there are reasons to be skeptical of the simplistic logic that more parties engender a greater sense of representation. If Rabinowitz and MacDonald (1989) find that parties in two-party systems, such as the United States, have incentives to be somewhat more extreme than voters, whereas Lijphart (1984) and Tsebelis (1990) argue that parties in a multiparty system, such as Belgium, often moderate over the heads of the electorate, is it obvious that the Belgian voter should feel more represented than the American voter? The answer is unclear, and due to the insufficient evidence regarding the relationship between the number of parties and political representation, the second, and final, hypothesis takes the form of a null hypothesis:

***Hypothesis II:** There is no relationship between the number of parties in a party system and voters feeling represented by a political party.*

3 Data compilation and coding

To test the foregoing hypotheses, data were gathered primarily from the Comparative Study of Electoral Systems accessed through the online database of the Inter-university Consortium for Political and Social Research (Philips and Sapiro 2007). The dataset used in this analysis covers thirty-eight elections between 2001 and 2006 across thirty-six democracies (voters in Taiwan and Portugal were polled on two separate election cycles). The sample size approaches nearly 60,000 voters. For all variables, responses such as “not available,” or “refused” were coded as missing.

For Hypothesis I, the dependent and independent variables of interest are *democratic satisfaction* and *representation by a political party*, respectively. The democratic satisfaction variable is compiled from the question “On the whole, are you very satisfied, fairly satisfied, not very satisfied, or not at all satisfied with the way democracy works in [country]?” with answers coded on a Likert scale ranging from 1 (not at all satisfied) to 4 (very satisfied). The representation by a political party variable is compiled by aggregating two dichotomous variables (“Do you usually think of yourself as close to any particular political party?” and “Would you say that any of the parties in [country] represents your views reasonably well?” with responses to both questions coded as 0 for a negative response and 1 for a positive response). The resulting variable thus ranges from 0 (“no” to both questions) to 2 (“yes” to both questions).

For Hypothesis II, the independent variable of interest is the *number of parties*. There has been much debate over how to best count the number of relevant parties in a party system (see Rae 1967; Sartori 1976; Sartori 1994; Lijphart 1999). This debate need not be repeated here, except for mentioning that any counting measure will possess some degree of arbitrariness and will be sensitive to differing interpretations of what “relevant” means. The measure for this analysis, and which has seen extensive use in empirical studies, is the effective number of parties (ENP) formula by Laakso and Taagepera (1979):

$$ENP = \frac{1}{\sum_{i=1}^n p_i^2}$$

Where n is the number of parties receiving at least one vote or parliamentary seat and p_i^2 is the proportion of seats or votes squared. This formula is used to calculate both the effective number of parliamentary parties (ENPP) and the effective number of electoral parties (ENEP) in this analysis. This helps clarify whether voters refer to electoral parties or parliamentary parties when they consider party representation, and allows for analyzing whether the significance and

directionality of the partial regression coefficients of the two measures differ. Note that I use the *log of the ENPP and the ENEP* in the analysis, the logic being that the effect of adding a political party decreases as more parties are present in a party system.

Other variables in the analysis are used as controls. These include demographic variables, namely *age*, *gender* (dummy coded as 1 for male and 0 for female), *education* (ranging from 1 for “no formal education” to 8 for “college degree attained”), *race* (coded as 0 for “European-Caucasian” and 1 for “other”), *marital status* (dummy coded as 1 for “married” and 0 for “divorced, single, or other”), *employment status* (where the coding is 1 for “employed” and 0 for retired, student, or other), *union, business, or farmers’ association membership* (coded as 1 for being a member and 0 for not being a member), *residence* (ranging from 1, denoting a rural residence, to 4, specifying residence in a large town or city), *self-reported religiosity* (on a scale from 1 to 4, with 4 denoting “very religious”), and *self-reported political ideology* (ranging from 1 for left-wing to 10 for right-wing).

For Hypothesis I, additional control variables are included if there are reasons to suspect that they may correlate with democratic satisfaction. These variables include *leader representation* (“Would you say that any of the individual party leaders/presidential candidates at the last election represent your views reasonably well?”), *election representativeness* (“How well do elections ensure that the views of voters are represented by MPs?”), *corruption perception* (“How widespread do you think corruption such as bribe taking is amongst politicians in [country]?”), *government performance* (“how good or bad a job do you think the government/president in [capital] has done over the past [number of years between the previous and the present election OR change in government] years?”), and *support for democracy*

("Please tell me how strongly you agree or disagree with the following statement: "Democracy may have problems but it's better than any other form of government").

For Hypothesis II, measures of corruption, government performance, and support for democracy are removed, since there is little theoretical reason why they should correlate with feeling represented by a political party. In their place, a control variable for the *average age of political parties* is used, compiled from the 2009 version of the Database of Political Institutions (Beck, Keefer, and Clarke 2009). The measure averages the age of the three largest parliamentary political parties. The logic for including this measure is that party attachment may be greater in institutionalized and longstanding party systems. All other control variables, namely demographic variables and representation variables (i.e. election and leader representation) are also retained in the analysis.

4 Results

Multivariate regression analysis was used to test both Hypothesis I and Hypothesis II. Cases were excluded pairwise in order to maximize the number of observations; nevertheless, the case count was reduced to 13,283 for all analyses, largely due to lower response rates for some variables and the exclusion of responses such as "don't know" or "refused" from the analysis. Final valid responses per variable, along with descriptive statistics, are displayed in Table 1.

For all of the analyses provided, I chose a significance level of $\alpha=.05$, and this is denoted with a double asterisk in the regression results. Coefficients that are significant at an $\alpha=.10$ level are considered to be approaching significance, and are denoted with a single asterisk to render them easier to discern. Significance levels of $\alpha=.01$ and $\alpha=.001$ are marked with three and four asterisks, respectively. The results for the first hypothesis test are displayed in Table 2 in the

column titled “Model 1,” displaying the correlation of the independent variables with the democratic satisfaction (dependent) variable. Partial regression coefficients are provided with their significance levels, followed by standard errors in parentheses. Finally, model summary statistics, such as R square values and F statistics, appear on the bottom of the regression table.

Table 1: Descriptive statistics for variables of interest.

	Mean	Std. Deviation	Minimum	Maximum	Value Labels	N Valid
Party representation	1.09	0.82	0	2	0=not close to party & no party represents my views; 2=close to party & party represents my views	49,710
Log (ENPP)	0.58	0.16	0.3	0.93	none	59,278
Log (ENEP)	0.66	0.19	0.34	1.45	none	59,278
Average party age	48.85	31.59	5	152.5	none	58,698
Leader representation	0.59	0.49	0	1	0=no; 1=yes	49,909
Election representativeness	2.44	0.74	1	4	1=not well at all; 4=very well	48,569
Corruption perception	2.96	0.91	1	4	1=it hardly happens at all; 4=very widespread	55,062
Democracy support	3.27	0.69	1	4	1=disagree that it is best form of gov't; 4=agree that it is best form of gov't	54,405
Government performance	2.45	0.76	1	4	1=very bad job; 4=very good job	51,368
Age	45.89	16.94	16	102	none	58,910
Gender	0.48	0.5	0	1	0=female; 1=male	59,176
Education	5.01	1.84	1	8	1=none; 8=undergraduate college degree completed	58,719
Marital status	0.66	0.47	0	1	0=other; 1=married or living together as married	56,936
Race	0.44	0.5	0	1	0=other; 1=European (Caucasoid)	20,326
Religiosity	2.57	0.97	1	4	1=have no religious beliefs; 4=very religious	35,344
Political ideology	5.15	2.56	0	10	0=left; 10=right	47,158
Union membership	0.23	0.42	0	1	0=no; 1=yes	51,512
Business association membership	0.04	0.19	0	1	0=no; 1=yes	35,685
Farmers' association membership	0.02	0.15	0	1	0=no; 1=yes	34,913
Employment status	0.06	0.24	0	1	0=other; 1=unemployed	55,860
Urban/rural residence	2.52	1.2	1	4	1=rural area/village; 4=large town or city	51,758

Note: Responses such as “refused,” “not available,” “see codebook,” and “other” were coded as missing and removed from the analysis.

Table 2: Regression results: democratic satisfaction (Model 1) and party representation (Models 2a-2c).

	Model 1	Model 2a	Model 2b	Model 2c
Party representation	.042**** (.008)			
Log (ENPP)			-.190**** (.039)	.470**** (.063)
Log (ENEP)		-.400**** (.034)		-.723**** (.055)
Average party age		.000 (.000)	.000 (.000)	.000 (.000)
Leader representation	.020 (.014)	.810**** (.013)	.813**** (.013)	.804**** (.013)
Election representativeness	.186**** (.008)	.135**** (.008)	.137**** (.008)	.135**** (.008)
Corruption perception	-.139**** (.007)			
Democracy support	.188**** (.009)			
Government performance	.289**** (.008)			
Age	.001**** (.000)	.004**** (.000)	.004**** (.000)	.004**** (.000)
Gender	.008 (.012)	.042**** (.012)	.041**** (.012)	.043**** (.012)
Education	.019**** (.003)	.020**** (.004)	.023**** (.004)	.021**** (.009)
Marital status	-.008 (.013)	.007 (.013)	.005 (.013)	.009 (.013)
Race	.006 (.013)	-.031** (.013)	-.001 (.013)	-.035*** (.013)
Religiosity	-.009 (.006)	-.006 (.006)	-.010 (.006)	-.001 (.006)
Political ideology	.009**** (.002)	.006** (.002)	.006** (.002)	.006** (.002)
Union membership	.038*** (.014)	.041*** (.015)	.035** (.015)	.024* (.015)
Business association membership	.003 (.031)	.021 (.031)	.023 (.031)	.026 (.031)
Farmers' association membership	.003 (.040)	.111*** (.041)	.104** (.041)	.134*** (.041)
Employment status	-.057** (.025)	-.023 (.025)	-.028 (.025)	-.017 (.025)
Urban/rural residence	-.019**** (.005)	.016*** (.005)	.017*** (.005)	.018*** (.005)
Constant	1.006****	.179****	-.018	.108**
R square	0.280	0.314	0.308	0.317
Adjusted R square	0.279	0.313	0.307	0.316
ANOVA F statistic	287.006****	380.076****	369.499****	362.392****
N	13,283	13,283	13,283	13,283

Note: Unstandardized partial regression coefficients provided, with standard errors in parentheses. Highest VIF value=3.063 for log (ENEP) in Model 2c (denoting little collinearity). *p<.1, **p<.05, ***p<.01, ****p<.001.

The regression results for Model 1 support Hypothesis I, namely that voters who feel represented by a political party are more likely to be satisfied with the democratic process. The partial slope regression coefficient for party representation is .042, meaning that, on average, for every unit increase in the party representation variable (for example, moving from a respondent who does not feel close or represented by a political party to one who does feel represented, but not close, to a political party) there is a corresponding .042 increase in the democratic satisfaction variable (recall that this variable ranges from 1, an unsatisfied response, to 4, a fully satisfied response). This result is statistically significant at the 99% level ($p=.000$). This outcome is largely unsurprising, and supports the literature arguing that voters gain some utility, and grow more satisfied with the democratic process, if they feel that a party represents their views.

These results justify moving to an analysis of the relationship between the number of parties and party representation; since feeling represented by a political party appears to improve satisfaction with the democratic process, it is important to scrutinize the spatially-derived assumption that increasing the number of parties should engender a greater sense of representation. Recall that Hypothesis II predicted that there would be no relationship in this case. This hypothesis is tested through Models 2a, 2b, and 2c in Table 2. The first model includes the $\log(\text{ENEP})$ as the independent variable of interest and omits the $\log(\text{ENPP})$, the second model does the reverse, including the $\log(\text{ENPP})$ and omitting the $\log(\text{ENEP})$, and the third model includes both. The reason for including three models in the regression table becomes clear when the significance level and directionality of the partial slope regression coefficients of both measures are analyzed.

Prima facie, the regression results appear to reveal a perplexing picture. In Model 2a, the $\log(\text{ENEP})$ partial slope regression coefficient is both negative ($b=-.400$) and highly statistically

significant ($p=.000$), suggesting that, on average, the greater the number of effective electoral parties, the less voters feel represented by a political party. Further, a similar results surfaces in Model 2b, with the $\log(\text{ENPP})$ partial slope regression coefficient being both negative ($-.190$) and statistically significant ($p=.000$). If the analysis was truncated here, it would appear that, against all intuition, the greater the number of both electoral and parliamentary parties, the less voters feel represented by a party. But Model 2c returns a bit of sanity to the analysis.

When both the $\log(\text{ENEP})$ and the $\log(\text{ENPP})$ are included in the analysis, and thereby each variable controls for the effects of the other, a striking shift occurs: the partial slope regression coefficient for the $\log(\text{ENEP})$ becomes even more negative ($b=-.723$) and statistically significant ($p=.000$), whereas the corresponding coefficient for the $\log(\text{ENPP})$ switches to becoming positive ($.470$) while also achieving very high statistical significance ($p=.000$). This shift in the more specified model might be explained away as a manifestation of multicollinearity, however the Variance Inflation Factor (VIF) does not rise above 3.063 for either measure². While clearly there is a correlation between the $\log(\text{ENEP})$ and the $\log(\text{ENPP})$ ($R^2=.642$) and thus some detectable collinearity between the two variables, such VIF values are far from concerning. Additionally, the high statistical significance of both variables in Model 2c suggests that it would take a significant amount of multicollinearity to warrant placing their coefficients' directionality and statistical significance under scrutiny.

A more plausible interpretation is that, because of the moderate correlation between the $\log(\text{ENEP})$ and the $\log(\text{ENPP})$, each variable was capturing the very significant, and directionally opposed, effect of the other variable in the first two Models. In other words, Models 2a and 2b suffer from omitted variable bias. In Model 2a, the coefficient for $\log(\text{ENEP})$ is less

² Note that a VIF value of 10 is often considered, albeit somewhat arbitrarily, as the point when multicollinearity can become a significant concern. VIF values of 2 or 3 usually denote minor amounts of multicollinearity.

negative than in Model 2c because it is likely detecting some of the positive correlation between the log(ENPP) and the dependent variable (party representation). Extending this logic, in Model 2b the coefficient for the log(ENPP) is pulled into negative territory because it is capturing some of the very negative correlation between the log(ENPP) and party representation. Once the model is fully specified, and both variables are included and allowed to control for each other's effects, their true correlations manifest themselves. These correlations are, vis-à-vis feeling represented by a political party, highly negative for the log(ENEP) and highly positive for the log(ENPP).

The striking story that Model 2c appears to support is that, on average, increasing the number of *electoral* parties makes voters feel *less* represented by parties, whereas increasing the number of *parliamentary* parties makes voters feel *more* represented by political parties. Therefore, rather than refuting or supporting Hypothesis II, the regression results provide mixed support at best and suggest that Hypothesis II lacked specificity; in order to determine whether the number of parties impacts party representation, one must first define *what kind of party* is being analyzed. But beyond this conclusion, it is important to interpret these results by providing some possible theoretic explanations, which is the focus of the next section.

5 Electoral versus parliamentary party representation

Given that the result supporting Hypothesis I, or that party representation engenders greater democratic satisfaction, finds widespread theoretic backing and is of little surprise, it is worthwhile to move to interpreting the more interesting results from the second hypothesis test. Two questions seem to warrant consideration. First, how can having *more* choice with respect to electoral parties cause voters to feel *less* represented? Second, why would the effect (vis-à-vis party representation) of increasing the number of electoral parties be opposite the effect of

increasing the number of parliamentary parties? The results certainly seem to contradict spatial logic and the rational-choice assumptions upon which it stands. Two plausible explanations are that a) voters are less happy when they have more parties to choose from during elections, and b) the conflict-inducing effect of more electoral parties disillusion voters, whereas the compromise-inducing effects of more parliamentary parties attracts voters.

The first explanation is not as counterintuitive as it might first appear. It is worthwhile to mention that there exists significant scholarly research, primarily in the behavioral economics literature, which finds that sometimes utility is increased by *reducing* the number of choices, that is that less choice sometimes makes people better off. For example, Sunstein and Thaler (2008) argue that a reduction in the number of options can often increase individual utility by overcoming uncertainty in the face of a complex choice. In the context of voters' decision over which parties to choose from, a plethora of parties may be a source of frustration and confusion for voters. This is due to two primary factors. First, as the number of parties increases, it becomes increasingly difficult to distinguish amongst parties. Second, making an informed decision when faced with many parties to choose from may require a greater amount of information and research on the part of the voter compared to a simpler binary choice. This requires time and is costly to the voter. In this light, it is plausible that voters actually prefer having fewer parties competing in elections because it simplifies their voting decision.

A strength of this interpretation is that it clearly distinguishes between electoral parties and parliamentary parties. Since parliamentary parties are parties in government, they do not necessitate voters having to 'choose' amongst them. Contrariwise, electoral parties compete for votes in elections, and voters ultimately have to create a party preference ordering to finalize which party will receive their support. In short, increasing the number of electoral parties does

increase the proximity between party and voter, but it also complexifies the decision voters must make regarding which party to support. The regression results suggest that, in this case, the latter effect may dominate the former, and thus voters' utility is decreased the more parties compete in elections. In the case of parliamentary parties, the second effect disappears, thereby allowing the benefits of greater representation in government to manifest themselves.

The second possible explanation is that voters garner a greater sense of representation when there are more parties in the parliamentary arena but not in the electoral arena. This could be caused by a number of factors. First, voters may care more about their policy positions being represented by parties in government, where they can be turned into policy, instead of parties competing in elections, where, at best, voters' preferences can be incorporated in a political platform. Further, an important issue to consider is that of inter-party divisiveness. Linz (2002) finds that voters across countries dislike when parties become a source of division within society, especially when such division is coupled with negative campaigning. This divisiveness is engendered by inter-party competition. Importantly, however, this competition amongst parties must be balanced by compromise in the parliamentary arena in order to avoid deadlock in government. This need for compromise is increased when there is a greater number of parliamentary parties, given that none of them commands a majority of seats. In short, a greater number of parties in government may cause its associated compromise-inducing effect to trump the parties' divisive effect. This argument is explicitly made by Lijphart (1984) and is one of the foundational assumptions of the consociational literature (or of "consensus" democracy, as Lijphart refers to it). Voters may thus appreciate that a greater number of parliamentary parties

not only increases the likelihood of their preferences being included in the policymaking agenda, but that the policymaking process is characterized by compromise and consensus³.

Conversely, a greater number of parties in the electoral arena may spark more conflict and engender greater divisiveness. This is because the “prize” in elections is not policy (as in the parliamentary arena), but electoral victory. The associated strategy, therefore, may move from one of compromise to one of negative campaigning, because there is a lesser incentive for electoral parties to cooperate vis-à-vis parliamentary parties. This is especially true if there are a significant amount of parties competing for votes, because each party has an incentive to distinguish itself from its competitors. Given the research finding that inter-party conflict is less than appreciated by voters, it would make sense that they would distance themselves from this culture of negativity by affirming that the increased conflict they engender leads to a disillusionment with the entire process.

To recapitulate, there exist plausible interpretations of the regression results. The first interpretation is that more choice renders voters worse off because it complexifies their voting decision. This effect only manifests itself in the electoral arena because voters do not have to choose amongst parliamentary parties. The second possibility is that a greater number of electoral parties engenders conflict and negative campaigning, thus causing disillusionment amongst voters, whereas the presence of more parliamentary parties provides an impetus for compromise and consensus in the policymaking process, which attracts voters instead. Voters may thus feel more represented by the compromise-stimulating effect of more parliamentary parties than by the divisiveness-stimulating effect of more electoral parties. These theories are summarized in Figure 1:

³ This is in clear contrast to the British two-party system case, for example, where the parliamentary party in power holds a monopoly on both parliamentary seats and agenda-setting power (Tsebelis 2002).

Figure 1: Selected impacts of increasing the effective number of parties.

More electoral parties	• <i>More choice, thus more complex voting decision</i> • <i>More inter-party conflict and division</i>
More parliamentary parties	• <i>No impact on decision-making process</i> • <i>Greater incentive for compromise and consensus</i>

While both of these conceptualizations may not fully explain the regression results, they do offer a plausible interpretation that can be validated, refuted, or revised through future research.

6 Conclusion

The possible interpretations offered in Section 5 are, stylized, developed *a posteriori*, and the analysis is tentative at best. However, their purpose is to show that there exist plausible interpretations of the regression results. Indeed, the analysis in this paper may have surfaced a puzzle meriting future research. Namely, the results suggest that while party representation increases satisfaction with the democratic process, more parties do not necessarily engender a greater sense of party representation; more electoral parties tend to decrease perceived representativeness, whereas a greater number of parliamentary parties tend to increase a sense of representation. To the author's knowledge, no prior study has found similar results. Further, given both the large size of the sample used for the analysis and the high statistical significance of the regression results, there is reason to conclude that the regression output is robust. Thus, while the results should not be free from skepticism, two conclusions can be drawn with confidence. First, when assessing party representation, both the type of representation and the type of party being scrutinized matter. Finally, while party representation appears to improve voters' satisfaction with democracy, more parties do not necessarily make for happier voters.

References

- Achen, C. 1978. "Measuring representation." *American Journal of Political Science*, 22: 475-510.
- Aldrich, John. 1983. "A Downsian spatial model with party activism." *American Political Science Review*, 82: 405-422.
- Beck, Thorsten, Keefer, Philip E., and George R. Clarke. 2009. Original 2001 publication: "New tools in comparative political economy: The Database of Political Institutions." *World Bank Economic Review*, 15(1): 165-176. 2009 data retrieved from: <<http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/0,,contntMDK:20649465~pagePK:64214825~piPK:64214943~theSitePK:469382,00.html>>.
- Dahl, R. 1971. *Polyarchy: Participation and Opposition*. New Haven, CT: Yale University Press.
- Dalton, R. J. 1985. "Political parties and political representation: party supports and party elites in nine nations." *Comparative Political Studies*, 18: 267-299.
- Davis, O. A., Hinich, M. J., and P. C. Ordeshook. 1970. "An expository development of a mathematical model of the electoral process." *American Political Science Review*, 64: 426-448.
- Dix, R. H. 1992. "Democratization and the institutionalization of Latin American political parties." *Comparative Political Studies*, 24: 488-511.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.
- Duverger, Maurice. 1951. *Political Parties*. New York: Wiley.
- Ferejohn J. 1995. "The spatial model and elections" in B. Grofman (ed.), *Information, Participation, and Choice: An Economic Theory of Democracy in Perspective*. Ann Arbor: University of Michigan Press.
- Hirschman, Albert O. 1970. *Exit, Voice, and Loyalty*, Cambridge, MA: Harvard University Press.
- Hotelling, Harold. 1929. "Stability in Competition" *Economic Journal*.
- Inglehart, Ronald. 1984. "The changing structure of political cleavages in Western society." In R.J. Dalton, Sc. Flanagan, and P.A. Beck (eds.), *Electoral change in advanced industrial democracies- Realignment or dealignment?* Princeton, NJ: Princeton University Press.
- Iversen, Torben. 1994. "The Logics of Electoral Politics: Spatial, Directional, and Mobilizational Effects." *Comparative Political Studies*, 27(2): 155-189.
- Key Jr, V. O. 1961. *Public Opinion in American Democracy*, Knopf, p. 433.
- Laakso, M., and R. Taagepera. 1979. "Effective Number of Parties: A Measure with Application to West Europe." *Comparative Political Studies*, 12: 3-27.
- Lijphart, Arend. 1984. *Democracies: Patterns of Majoritarian and Consensus Government in Twenty-One Countries*. New Haven, CT: Yale University Press.
- Lijphart, Arend. 1999. *Patterns of Democracy: Government Forms and Performance in Thirty Six Countries*. New Haven, CT: Yale University Press.
- Linz, Juan. 2002. "Parties in Contemporary Democracies: Problems and Paradoxes," in Richard Gunther and Jose Ramon Montero and Juan Linz (eds.), *Political Parties: Old Concepts and New Challenges*. Oxford: Oxford University Press.
- Lipset, Seymour M., and Stein Rokkan. 1990. "Cleavage Structures, Party Systems, and Voter Alignments," in Peter Mair (ed.), *The West European Party System*, pp. 99-138.
- Listhaug, O., Macdonald, S. E., and G. Rabinowitz. 1990. "A comparative spatial analysis of European party systems." *Scandinavian Political Studies*, 13: 227-254.

- McKelvey, R. D. 1986. "Covering, dominance, and institution free properties of social choice." *American Journal of Political Science*, 30: 283-314.
- Neumann, Sigmund. 1956. *Modern Political Parties*. Chicago: University of Chicago Press, p. 397.
- Przeworski, A., and J. Sprague. 1986. *Paper stones: a history of electoral socialism*. Chicago: University of Chicago Press.
- Rabinowitz, G. and S. E. McDonald. 1989. "A directional theory of issue voting." *American Political Science Review*, 83: 93-121.
- Rabinowitz, G., Macdonald, S. E., and O. Listhaug. 1991. New players in an old game: Party strategy in multiparty systems. *Comparative Political Studies*, 24: 147-185.
- Rae, Douglas. 1967. *The Political Consequences of Electoral Laws*. New Haven: Yale University Press.
- Sapiro, Virginia, and W. Philips Shively. COMPARATIVE STUDY OF ELECTORAL SYSTEMS, 2001-2006 [Computer file]. ICPSR03808-v2. 2007. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2008-07-01. doi:10.3886/ICPSR03808
- Sartori, Giovanni. 1968. "Representational systems." In *International Encyclopedia of the Social Sciences*, vol. 13: 470-475. New York: Macmillan.
- Sartori, Giovanni. 1969. "From the sociology of politics to political sociology." In S.M. Lipset (ed.), *Politics and the social sciences*. Oxford: Oxford University Press.
- Sartori, Giovanni. 1976. *Parties and Party Systems: A Framework for Analysis*. Vol.1, Cambridge: Cambridge University Press.
- Sartori, Giovanni. 1994. *Comparative Constitutional Engineering: An Inquiry into Structures, Incentives, and Outcomes*. New York: New York University Press.
- Schattschneider, E.E. 1942. *Party Government*. New York: Rinehart.
- Simmel, Georg. 1903. "Soziologie der Konkurrenz." *Neue Deutsche Rundschau* 14(4): 1009-1023.
- Stokes, Susan. 1999. "Political Parties and Democracy." *Annual Review of Political Science*, 2: 243-267.
- Strom, Kaare, and Wolfgang C. Muller (eds.). 1995. *Policy, Office, or Votes*. Cambridge: Cambridge University Press.
- Sunstein, Cass R., and Richard H. Thaler. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.
- Tsebelis, George. 1990. *Nested Games: Rational Choice in Comparative Politics*. Berkeley, CA: University of California Press.
- Tsebelis, George. 2002. *Veto Players: How Political Institutions Work*. Princeton, NJ: Princeton University Press.

THE LONG-TERM EFFECTS OF COLONIAL LAND TENANCY: MICRO EVIDENCE FROM INDIA

Sayon Deb
sdeb@bu.edu

Faculty Advisor:
Robert A. Margo
margora@bu.edu

Institutional Affiliation:
Boston University
M.A., Expected May 2011

THE LONG-TERM EFFECTS OF COLONIAL LAND TENANCY: MICRO EVIDENCE FROM INDIA

ABSTRACT

This paper uses household survey data from India to examine the impact of historic land tenure institutions on economic and social outcomes for households today. It offers evidence on specific channels through which the structure and quality of land tenure (i.e. revenue collection) systems could persist today. We find that districts where land ownership was dominated by landlords, today have lower annual income, per capita consumption, and cumulative household asset levels than districts which were characterized by non-landlord tenure systems. Households in landlord districts are more likely to have narrower social networks and lower levels of memberships in community organizations, weaker propensity to work collectively in solving communal problems, and are more likely to be subject to crime than district where non-landlord systems were prevalent. Our results are significant and robust to a diverse set of controls.

I. Introduction

A series of recent studies find that India's colonial past had persistent effects on current regional performance. The seminal work on this topic is that by Abhijit Banerjee and Lakshmi Iyer (2005). They show that differential colonial land tenure institutions gave rise to distinct patterns of development, the effects of which persist even today. Parts of India where proprietary rights were concentrated among landlords (i.e. elites) have significantly lower levels of agricultural investments and lower productivity compared to areas where the rights were much more widespread (i.e. individual cultivators or village bodies). They also find significant differences in health and educational infrastructure. They argue that serious differences in collective action due to the political antagonism, lower levels of investments in agricultural technologies, and lower levels of investment in human capital account for such a divergence.

Various complementary studies have added onto the empirical examination of the structure of land tenure systems and public goods availability in India. Banerjee, Iyer, and Somanathan (2005) have explored the role played by colonial power, landowner-peasant relations, and social fragmentation based on Hindu caste system on the availability of public goods. They find a negative relationship between concentration of power among the elites and caste fragmentation with availability of public goods. Banerjee and Somanathan (2006) study the mechanism of allocation of public goods by a centralized state using the example of India in the 1970s and 1980s. A model outlined by Banerjee, Iyer, and Somanathan (2006) argue the efficacies of a "top-down" intervention in solving collective action problems, in contrast to the popular "bottom-up" model, in improving access to public goods.

Separately, Iyer (2010) compares economic outcomes between areas in India governed directly under British administration and those under indirect rule. She finds significantly lower

levels of access to public goods in the post-colonial period for areas that were under direct rule. Taking the opposite direction in history, Chaudhary (2008, 2010) finds that there were considerably different patterns of investments in public education during the colonial period. While investments show a significant positive effect in increasing literacy, large inter-regional differences which existed can be attributed to the differences in the amount of land revenue collected, and indirectly the method it was collected in as well. Yet another study by Pandey (2010) finds that local governance and education outcomes in Oudh (in present-day Uttar Pradesh) are worse in villages that belonged to the landlord-based system. The effect is attributed to a history of concentration of political power among the landed elite.

This study is an extension of that agenda, but where it differs is in offering evidence at the micro-level. Inspecting at a finer disaggregation than previous related studies, we are able to examine carefully the exact channels through which historical institutions may persist. In particular, we estimate the differences between landlord and non-landlord areas in household income, consumption per capita, and cumulative asset levels. We find large, positive, and significant differences for households in non-landlord areas. We examine the social capital of households to test channels which proxy for collection action. We find significant differences in the extent of their social networks, civic participation, confidence in institutions, and propensity to work collectively to solve communal problems. This potentially provides some evidence on the channels through which the effect of land tenure institutions continue to persist.

The paper is structured as follows. In the following section, we provide a brief historical background, a description of the various land tenure systems, and the conceptual framework and motivations for undertaking the study. Section 3 describes the sources of data, and in Section 4 the empirical approach and estimates are discussed. Section 5 concludes.

II. Historical Background and Conceptual Framework

A. Company Raj

The British East India Company's dominion began rather humbly with securing rights to construct a trading post in the port of Surat in 1612 from the Mughal emperor Jahangir. The Company expanded over the remainder of the 17th century most notably with a land grant in 1639 to establish trading posts on the Coromandel Coast (present-day Chennai) and the leasing of Bombay island in 1668 and Kalikata (present-day Kolkata) in 1698. International trade under the Company flourished as the markets for spices, silk, indigo, and cotton expanded. In 1717, recognizing the growing economic role the Company played, emperor Farrukhsiyar granted the British exemption, by *firman* or decree, from paying custom duties in Bengal.

The Company's victories against *Nawabs* under Robert Clive at Plassey in 1757 and at Buxar in 1764 led to the Treaty of Allahabad in 1765 granting it *diwani* or sole revenue collection rights in Bihar, Bengal, and Orissa. Majority of these newly acquired areas were placed under landlord systems which were eventually subject to nominally fixed amounts of revenue under the Permanent Settlement Act.

Victories during the Anglo-Mysore Wars (1766 to 1799) and Anglo-Maratha Wars (1772-1818) enabled the expansion of the British Empire over much of southern and western India. Unlike the wars in eastern India, the turmoil lasted nearly half a century. The resulting environment of conflict might have played a role in delaying and in some cases preventing the rise of a landlord or landed elite class in these areas (Peers 2006). Over the next half century, the Company asserted its growing power through annexation or treaties over the remaining parts of the Indian subcontinent, allowing for a handful of princely states to maintain administrative

autonomy. As the Company's share of territory increased, land revenue became one of the principal sources of funds.

Company *Raj* came to an end in 1858 following the Sepoy Rebellion and the administration and military in India was reorganized and governed directly by the British government, marking the beginning of British *Raj*. India would not become an independent nation for another 90 years.

B. Land Revenue Systems

The British installed three forms of land revenue systems in India: (1) landlord (*zamindari* or *malguzari* or *talukdari* in North-West Provinces), (2) village-body (*mahalwari*), or (3) individual cultivator (*raiyatwari*). The landlord system was implemented in parts of present day states of Bihar, West Bengal, Orissa, Madhya Pradesh, Tamil Nadu, Andhra Pradesh, Uttar Pradesh (i.e. Oudh) and Assam. Virtually all agricultural land in these areas were under the proprietorship of a single landlord. They were responsible for all revenue collection and directly responsible for paying a portion of that revenue to the British. Although in some parts the amount of revenue owed to the British was permanently or temporarily settled, the amount collected by the landlord was self-determined and variable over time allowing for arbitrary extraction of rents from landless peasants.

The village-body based system was implemented in parts of present day states of Punjab, Haryana, Himachal Pradesh, and Uttar Pradesh. The village-body's jurisdiction over area, its size, and composition varied from place to place. In some places, the village-body consisted in essence of a landlord family, and in others, it was composed of many members with each member responsible for a fixed share of the revenue (Banerjee and Iyer 2005).

Unlike the previous two systems, under the individual cultivator system, cultivators were given legal rights over their land and were directly responsible for paying the British revenue

demand. Due to this direct relationship with state and cultivator, the revenue amount was responsive to changes in annual yields of the cultivator. The present day states of Maharashtra, Gujarat, Karnataka, and Tamil Nadu comprised most of the historical districts under the individual cultivator system.

Banerjee and Iyer (2005) argue that decisions to install one form of land revenue system over another were often based on the influence of individual administrators, political events, the date of annexation, and the pre-annexation or pre-colonial presence of a landlord class (i.e. Bengal). It is noteworthy that districts annexed at a later date had political precedent for favoring non-landlord arrangements whereas in districts annexed at an earlier date, the cost of delegating the role of revenue collection to an intermediary was low leading to the empowerment of those privileged individuals. The argument used to defend a non-landlord system were along the lines of prevailing principle or ideology of the time¹ than factors directly related to agricultural productivity. They conclude that the choice of land revenue system can thus be treated as largely exogenous.

To address remaining concerns of exogeneity of the choice of land revenue systems they pursue two alternative strategies: (1) by comparing contiguous non-landlord and landlord districts and (2) comparing districts that were brought under British revenue control between 1820 and 1856 with districts outside of this period. The support for the strategy is derived from the fact that districts annexed after 1820 were for the most part installed with non-landlord system due to political influences of individual administrators (i.e. Holt Mackenzie in North-West Provinces, Sir Thomas Munro in Madras Province, and Lord Elphinstone in the Bombay Presidency). They find similar results from both strategies as from assuming that the choice of revenue collection system was exogenous.

¹ James Mill's *Elements of Political Economy* was influential in advocating the "ryotwari" or individual cultivator system (Brown 1994).

C. Conceptual Framework

Can contemporary household-level data on political behavior and economic conditions of individuals shed light on the mechanisms through which institutional quality persists? Can we infer from such an analysis what role the “long hand of history” plays in determining growth patterns and social climate? These are the primary questions which drive the motivation for this paper. Given both the political and economic aspects of the question, separate angles of analysis are needed.

We look at the economic aspect using a framework similar to that established in Banerjee and Iyer (2005), but the fineness of our data allows us to extend the analysis much further. While there is some overlap between the approach and interpretation of many of our results, we are able to examine some of the channels that result in different productivity outcomes in aggregate. In their paper, they find that states with a higher landlord population have higher Gini coefficients of land inequality in 1885 than non-landlord population and the difference persisted as recently as 1947. As the landed elite grew richer, the inequality in those areas grew higher. While inequality grew in the non-landlord areas also, it happened at a much slower pace and at a later date. One of the channels through which the effect of revenue collection institutions could persist today is the inequality in the distribution of wealth. The theoretical link between income distribution and long-run development is well established in the empirical literature. One example is Banerjee and Newman (1993), which finds that the initial distribution of wealth is a crucial determinant of prosperity or stagnation of the economy in the long-run. Growth in economies with high initial inequality in wealth distribution flatten over time to a low employment, log wage equilibrium whereas economies with low initial inequality (although poor in aggregate) prosper to a high employment, high wage steady state. We look for the effects of

inequality resulting from the landlord system in current households' income, consumption per capita, and asset accumulation. While income and consumption per capita provide a short-term measure of economic wellbeing, it is likely that they are determined by the long-term institutional environment established in those areas. This is not to suggest that short-term shocks or reforms are unimportant. India's rapid growth over the last decade may as well ameliorate much of the institutional deficiencies in the future. Asset accumulation is thought of in terms of the household's accumulated level of consumer goods and housing (i.e. investments made to improve quality). This enables us to look at the long-term economic levels of each household and its potential as asset accumulation requires sustained employment and/or income.

We take up the issue of differences in political environment by looking at the culture of participation through household memberships in social organizations and voting, levels of confidence in institutions, and levels of social capital (particularly social networks and collective action). The impact on individuals' beliefs is another channel through which history can affect current outcomes (Hoff and Stiglitz 2001). The landed elite's control over the terms and collection of revenue naturally increased its political clout. Exploitation and oppression by landlords of the landless peasants created an environment of insecurity in which political participation was not only discouraged but wholly prevented. Bose (1993) argues that the political atmosphere created by class-based conflict shaped much of policies and politics in pre- and post-independent landlord communities. While class-based conflict was not entirely absent from non-landlord areas, their total magnitude was much smaller. In non-landlord areas, particularly where individual cultivators held titles to the land, the incentives to participate and have aligned political interests were highly beneficial in negotiations with the state.

In contrast, the landed elites and peasants were misaligned in their political agendas resulting in a culture of mistrust. Banerjee and Iyer (2005) argue that collective action most likely failed in landlord areas due to this misalignment. Additionally in India, class-based conflict has most often been along the lines of the caste system². Almost all of the landed elites were comprised of *Brahmins*, who are high-caste individuals, while the landless peasants were mostly lower caste and tribal members. Hoff, Kshetramade, and Fehr (2009) find that mistrust is more prevalent in lower caste members who are less likely to take punitive action against violators of cooperation. The cultural legacy of caste system on individual trust levels is also documented in Hoff and Pandey (2005). Tabellini (2008) models the persistence of mistrust using norms of cooperation that are passed on from generation to generation. These norms in turn determine the quality of institutions created by individuals in a society. Individuals with lower levels of trust tend to choose weaker institutions which sanctions uncooperative behavior less often, thus setting the stage for the persistence of mistrust in following generations. Nunn and Wantchekon (2010) is another such study which finds individuals whose ethnic groups' bore the brunt of the slave trade today display significantly lower levels of trust. They conclude that the slave trade through its effects on individual trust levels is partly to blame for Africa's current underperformance. While we do not have exact measures of trust like that used in their study, we do have a weaker correlated measure of confidence in various institutions. Newton and Norris (1999) argue that while there is a relationship between trust and confidence, it is most likely indirect. If performance of political institutions are the outcome of the levels of social capital (partly determined by trust levels), better performance will be positively correlated with higher levels of confidence. This essentially means that while we are able to gauge the level of confidence

² Castes can be described either as *varnas* or *jatis*. *Varnas* are derived from ancient Hindu society which classifies individuals in terms of occupations into five groups: *Brahmins*, *Kshatriya*, *Vaisya*, *Sudras*, and *Ati Sudras*. The *jati* system fits the contemporary social structure better and classifies individuals into five groups: *Brahmins*, *Scheduled Castes* (SC), *Scheduled Tribes* (ST), and *Other Backward Castes* (OBC).

households have in political institutions, we are unable to disentangle the added effect from individual's perception of institutional performance.

The hypothesis that we test is whether non-landlord areas which were characterized by clusters of “good” institutions perform better in political and economic outcomes at the household level. The aim of our analysis and intended contribution is to estimate empirically the extent to which the political behavior and economic structure of households continue to be impacted by the colonial land tenure systems.

3. Data

Our data is a combination of historical data at the district-level with recent survey data at the household-level. In 2001, India had 593,731 inhabited villages, 384 urban agglomerations, and 5,161 towns making up in aggregate 593 districts³. An average district is about 5,200 square kilometers in size with 1.73 million residents. The total number of households in India was 193,579,954, with the mean household size being 5.3 individuals. While the 2005 statistics are most likely higher in some categories, we can assume that the increase is not substantially different for this analysis.

*A. Colonial Land Tenure Data*⁴

The measure of the proportion of non-landlord area in each district is taken as constructed in Banerjee and Iyer (2005). They matched modern districts to older British districts that were under direct administrative control. For each of these districts, they use historical evidence of land revenue systems between 1870 and 1890 about villages, estates, and/or land area to compute the proportion of the district under non-landlord control during the colonial period. For districts

³ In India, a district or *Zill-Parishad* is the level of disaggregation finer than state. A village (*Gram Panchayat*) or urban neighborhood (*Wards*) comprised of municipal corporations (*Mahanagar-Palika*), municipalities (*Nagar-Palika*) or town councils (*Nagar-Panchayats*) are the lowest level of disaggregation.

⁴ The land tenure dataset is publicly available at http://www.aeaweb.org/aer/data/sept05_data_banerjee.zip.

without exact information on district-wise non-landlord control, they assign values of 0 (completely landlord) or 1 (completely non-landlord) depending on the prevailing colonial land revenue system. See Web Appendix tables 1, 2, and 3 of Banerjee and Iyer (2005) for detailed information on historical sources and construction of the measure for each district. We use the measure unedited for this analysis.

B. Indian Human Development Survey (IHDS 2005)⁵

The household level data is from the 2005 Indian Human Development Survey (IHDS) conducted during 2004 and 2005. The nationally representative survey was designed and conducted by the University of Maryland and the National Council of Applied Economic Research covering 1503 villages and 971 urban neighborhoods in 383 (of 626) districts in India. Local interviewers were employed in translating and administering the survey in 13 Indian languages. The survey yields a potential sample of 41,554 Indian households. The IHDS (also known as HDPI-2) is a continuation of the 1994-95 Human Development Profile of India (HDPI-1) survey and thus includes 13,900 re-interview households. The dataset is constructed from two one-hour interviews in each household covering education, employment, economic status, health, political behavior, and social capital. From the eight sections⁶ of the survey, only the household section is used for this paper. The summary statistics are available in table 1.

Household.—The IHDS asked each household questions on a wide range of topics, of which we focus on three economic measures and four measures of social capital.

The three measures of a household's economic levels included in the IHDS are (a) income, (b) monthly consumption expenditures, and (c) assets (i.e. consumer goods owned) and housing quality. The income measure was constructed from a rich array of questions based on 50

⁵ The IHDS dataset is available through ICPSR at <http://www.icpsr.umich.edu/icpsrweb/DSDR/studies/22626>.

⁶ Eight sections include individuals, households, medical facility, non-resident, primary school, birth history, village, and crops.

Table 1
Summary Statistics

Variable	Mean	Std. dev.	Min	Max	N
<i>Economic indicators</i>					
ln income	10.327	0.978	6.908	15.691	24904
ln consumption per capita	6.515	0.685	4.605	10.578	25337
Household asset index	11.651	6.173	0	30	25407
<i>Social indicators</i>					
Social network position	1.131	1.217	0	4	24893
Community memberships	0.565	1.048	0	9	25345
Vote in 2004 election	0.916	0.278	0	1	25368
Conflict in village	1.375	0.739	0	2	25265
Conflict between <i>jati</i>	1.664	0.555	0	2	25267
Confidence in institutions	23.313	3.406	10	30	22486
Local crime	0.073	0.319	0	3	25340
<i>Village infrastructure</i>					
Access to road	1.602	0.592	0	2	16605
Access to electricity	0.692	0.324	0	1	16605
Access to telephones	0.579	0.494	0	1	25407
Access to public programs	15.358	4.922	1	25	16605
Access to community groups	3.427	2.155	0	9	16597
Access to efficient fuel	0.061	0.238	0	1	16605
State	Mean altitude (Std. dev)	Mean latitude (Std. dev)	Mean rainfall (Std. dev)	Mean non-landlord proportion (Std. dev)	
Andhra Pradesh	291.68 (90.21)	16.77 (1.60)	877.10 (160.40)	0.664 (0.288)	
Bihar	295.35 (105.69)	24.89 (1.17)	1733.58 (402.13)	0.000 (0.000)	
Gujarat	212.94 (92.61)	22.33 (1.02)	618.89 (257.79)	1.000 (0.000)	
Haryana	535.16 (76.27)	29.03 (0.89)	1149.21 (260.01)	0.852 (0.154)	
Karnataka	462.63 (35.92)	14.31 (1.71)	1171.28 (683.59)	1.000 (0.000)	
Madhya Pradesh	369.6 (102.29)	23.33 (1.46)	1125.85 (145.37)	0.097 (0.261)	
Maharashtra	401.84 (44.77)	19.38 (1.40)	1117.83 (417.04)	0.778 (0.416)	
Orissa	177.07 (51.98)	20.69 (0.91)	1402.22 (176.31)	0.320 (0.404)	
Punjab	661.41 (98.44)	30.95 (0.64)	1119.92 (177.10)	0.872 (0.178)	
Rajasthan	327.46 (53.06)	26.16 (1.60)	664.09 (201.54)	0.000 (0.000)	
Tamil Nadu	318.39 (70.64)	10.79 (1.39)	895.85 (246.03)	0.751 (0.217)	
Uttar Pradesh	384.54 (139.39)	26.97 (1.66)	1377.66 (246.54)	0.424 (0.317)	
West Bengal	213.61 (147.39)	24.11 (1.54)	2284.38 (1090.90)	0.000 (0.000)	

different sources of income. These were classified into eight categories: (1) farm income, (2) agricultural wages, (3) non-agricultural wages, (4) monthly salaries and/or daily wages, (5) net business income, (6) household remittance received from non-residents, (7) government benefits, (8) unearned sources of income like property, pensions, etc. The annual totals were estimated using the number of days worked as reported by individuals in the household. For our sample, the household mean income was Rs. 49,936 (in 2004-05 Rupees), with a standard deviation of Rs. 79,906. We restrict our sample to incomes greater than or equal to Rs. 1000.

The consumption per capita measure was constructed using 47 questions⁷ about household consumption. The first 30 items were basic household items which are frequently purchased on a monthly basis, and the remaining seventeen items used an annual frame and included items which are more expensive. The household mean consumption per capita for our sample was Rs. 880, with a standard deviation of Rs. 909. We restrict our sample to consumption greater than or equal to Rs. 100. The household assets measure is a sum of 30 dichotomous items⁸ which measure the household possessions and housing quality. Each household's index ranges from 0 to 30. The household mean asset level was 11.86, with a standard deviation of 6.1.

Questions regarding social capital consist of (a) the extent of social network in medical, education, and government institutions, (b) memberships in groups or organizations and political activity, (c) local conflict and collective action, and (d) confidence in institutions. For all questions on social capital, the log of household income is included as part of the controls. This

⁷ Rice, wheat, sugar, kerosene, other cereals, cereal products, pulses, meat, sweeteners, edible oil, eggs, milk, milk products, vegetables, salt/spices, other food, paan or tobacco, fruits/nuts, eating out, fuel, entertainment, telephone, personal care, toiletries, other household items, conveyance, housing/other rent, consumer taxes/fees, domestic services, medical out-patient, medical in-patient, school fees, school books, clothing/bedding, footwear, furniture, crockery, household appliances, recreation goods, jewelry, transport equipment, therapeutic, other personal, repair/maintenance, insurance premiums, vacations, and social functions.

⁸ Bicycle/scooter, sewing machine, mixer/grinder, motor vehicle, b&w tv, color tv, air cooler/conditioner, clock/watch, electric fan, chair/table, cot, telephone, cell phone, refrigerator, pressure cooker, car, washing machine, computer, credit cards, 2 clothes, footwear, piped indoor water, separate kitchen, flush toilet, LPG, electricity, pucca wall, pucca roof, pucca floor.

is necessary as to make sure we are not simply repeating the results of table 1 and thus interpreting spurious correlations.

Social networks of households are gauged by asking whether “among [their] acquaintances and relatives, are there any who are” in medical institutions (i.e. doctors, nurses, or hospital/clinic staff), educational institutions (i.e. teachers, school officials, or school staff), or government (i.e. officers or above, clerks, or staff). In addition, the survey also asks whether anyone in the household or someone close to the household is an official of the village council or ward committee. Following the strategy used by Vanneman et. al. (2007) for the same dataset, we construct a positional generator of social networks which indexes how many ties a household has to the four categories mentioned above. The index ranges from 0 (no connections) to 4 (connected to all four). The household mean for our sample was 1.19 and the median was 1. About 20% of the households had 3 or more connections whereas almost 41% had no connections.

Political participation by households are measured using questions on memberships in groups or organizations and whether they voted in the 2004 national election. Households are inquired about memberships in nine organizations: (1) *mahila mandal* (women’s empowerment group), (2) youth club, sports group, or reading room, (3) self help groups, (4) credit or savings group, (5) religious or social group or festival [organizing] society, (6) caste associations, (7) development group or NGO, and (8) agricultural, milk, or other co-operative. We generate a cumulative index from 0 to 9 to account for gender-specific, age-specific, and occupation-specific differences in memberships within a household. In our sample 32% of households had memberships in 1 or more organization, whereas 68% of the households in our sample did not have any affiliation with the groups or organization surveyed.

Direct questions on trust were not asked in this survey. Instead households' opinions on the level of conflict, collective action, and confidence in institutions were gauged. Two questions were asked on local (i.e. village or neighborhood) conflict: (1) "do people generally get along with each other or is there some conflict or a lot of conflict?", and (2) "how much conflict would you say there is among communities/*jatis* that live here?". The answers for the second question followed the same format as the first, in which respondents chose among "a lot of conflict", "some conflict", or "not much conflict". Responses for both questions were assigned a value of 0, 1, or 2, respectively. In our sample, 14% reported that there is a lot of conflict in their village and only 4% reported there is a lot of conflict among *jatis*. The third question on a different but related topic was whether "people bond together to solve [a community] problem ... or take care of their own families individually". The respondents chose between "bond together to solve problem" or "each family solve individually". The responses were assigned a value of 1 or 0, respectively. 57% of our sample reported that they bond together to solve communal problems.

The last measure of social capital concerned the confidence in institutions. Respondents were asked whether they had a "great deal of confidence" (a value equal to 1), "only some confidence" (=2), or "hardly any confidence at all" (=3) in ten institutions: (1) politicians, (2) military, (3) police, (4) state government, (5) newspapers, (6) village councils, (7) schools, (8) hospitals and doctors, (9) courts, and (10) banks. We reversed the scales to indicate greater confidence, and summed values over the nine questions to create an index with a 30-point maximum. The mean value of our index was 23.4, with a standard deviation of 3.3.

Despite the breadth of the IHDS data, the analysis is constrained to 271 potential districts for which the non-landlord proportion measure exists. In addition, 69 districts for which the non-landlord measure exists there is no corresponding IHDS data. After removing these observations,

a final sample of 202 districts from 13 states⁹, with 971 primary sampling units (576 rural villages and 395 urban neighborhoods) are cleanly matched with districts for which the non-landlord proportion is constructed.

4. Empirical Approach and Estimates

A. OLS Estimation

We estimate the relationship between a district's historic non-landlord control and households' political and economic outcomes today in that district. Our baseline estimating equation is:

$$(1) \quad y_{ijk} = \alpha + \beta PNL_k + \gamma FRAC_j + X'_{ijk} \varphi + X'_{jk} \phi + \varepsilon_{ijk}$$

where i indexes individual households, j villages/neighborhoods, and k districts. PNL_k measures the proportion of district k under non-landlord control during colonial India. This estimating equation is similar to that used by Banerjee and Iyer (2005), but we include controls for household characteristics and village caste fractionalization. Neither district-fixed nor state-fixed effects are used in our baseline estimation for two reasons: (1) the proportion of a district under non-landlord control is historically determined for each district; and (2) the proportional measure also varies more between-states than within-states thus adding state-fixed effects would effectively drop states that were either completely under landlord or non-landlord systems. For each village/neighborhood there are observations for at least two or more households and we account for variation by clustering at that level.

The variable $FRAC_j$ is intended to capture the sub-caste or *jati* composition of the household's village/neighborhood. The motivation to measure sub-caste heterogeneity arises

⁹ Andhra Pradesh, Bihar (including Jharkhand), Gujarat, Haryana, Karnataka, Madhya Pradesh (including Chhattisgarh), Maharashtra, Orissa, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh (including Uttaranchal), and West Bengal. States formed after 1991 are included as part of old states.

from empirical research on population heterogeneity and its consequences for citizenship behavior. Easterly and Levine (1997) find cross-country evidence in Sub-Saharan Africa which suggests that high ethnic fragmentation may explain a significant part of depressed economic growth. Alesina and La Ferrara (2002, 2005) find that in more heterogeneous communities, individual trust levels and social participation are significantly lower. We calculate fractionalization in a village or neighborhood of our sample using the Hirschman-Herfindahl index:

$$(2) \quad FRAC_j = 1 - \sum_{c=1}^n p_c^2$$

where p_c is the share of sub-caste c over the total population. We construct the index using percent share of sub-castes in each village/neighborhood from the IHDS Village Questionnaire. Measure (2) ranges from perfect homogeneity (equals 0) where every individual is from the same *jati* to perfect heterogeneity (equals 1) where every individual is from a different *jati*. The measure assumes a shared identity among individuals which can be thought of as an objective category but does not make value assumptions about caste standing (Anderson and Paskeviciute 2006). For our sample, the average fractionalization index is 0.63.

The vector X'_{ijk} denotes a set of household-level covariates, which include a gender indicator variable that equals one if the head of the household is a female, an indicator variable that equals one if the household is in an urban location, five caste fixed effects, twelve fixed effects for principal source of income for the household or the dominant occupation of the household, a native indicator variable that equals one if the household has been residing there longer than 70

years¹⁰, the number of individuals within a household, and nineteen fixed effects for the highest level of education completed by an adult (defined as 21 or years older)¹¹.

The vector X'_{jk} denotes the set of district-level geographic controls and an additional variable that equals the year when the district came under British revenue control. Geographic controls include climatic, topographic, soil, and a coastal indicator variable which equals one if the district is on the coast. The date of British land revenue control is intended to capture the length of British rule in the district. Both the timing of the revenue control and the duration of revenue control often determine unique characteristics and the nature of the land revenue system in some districts. These independent effects are picked up using this control. The vector is taken from Banerjee and Iyer (2005).

Estimates of equation (1), for income, consumption per capita, and asset levels of households, are reported in table 1. The difference in all three measures are large and significant between non-landlord and landlord areas. In column (1) we test for differences using the full sample of districts. Using our base specification, we find that households in non-landlord areas have 25% higher income, 39% higher per consumption levels, and 24% higher levels of asset accumulation. The linear relationship is also shown in figure 1. We follow a second strategy in column (2) by restricting our sample to only districts where there was a historic mixture of non-landlord and landlord revenue systems. While the sign of the coefficient on household income stays positive, it decreases in effect and loses its significance alluding to the fact that our results are perhaps partly driven by the districts in our sample where the revenue system was wholly

¹⁰ The original variable records the number of years the household has resided in that village. The values range from 1 to 90. We intentionally use the 70-year cutoff to place the household's first appearance in the village around 1935, a little more than a decade before the formal abolition of the landlord system.

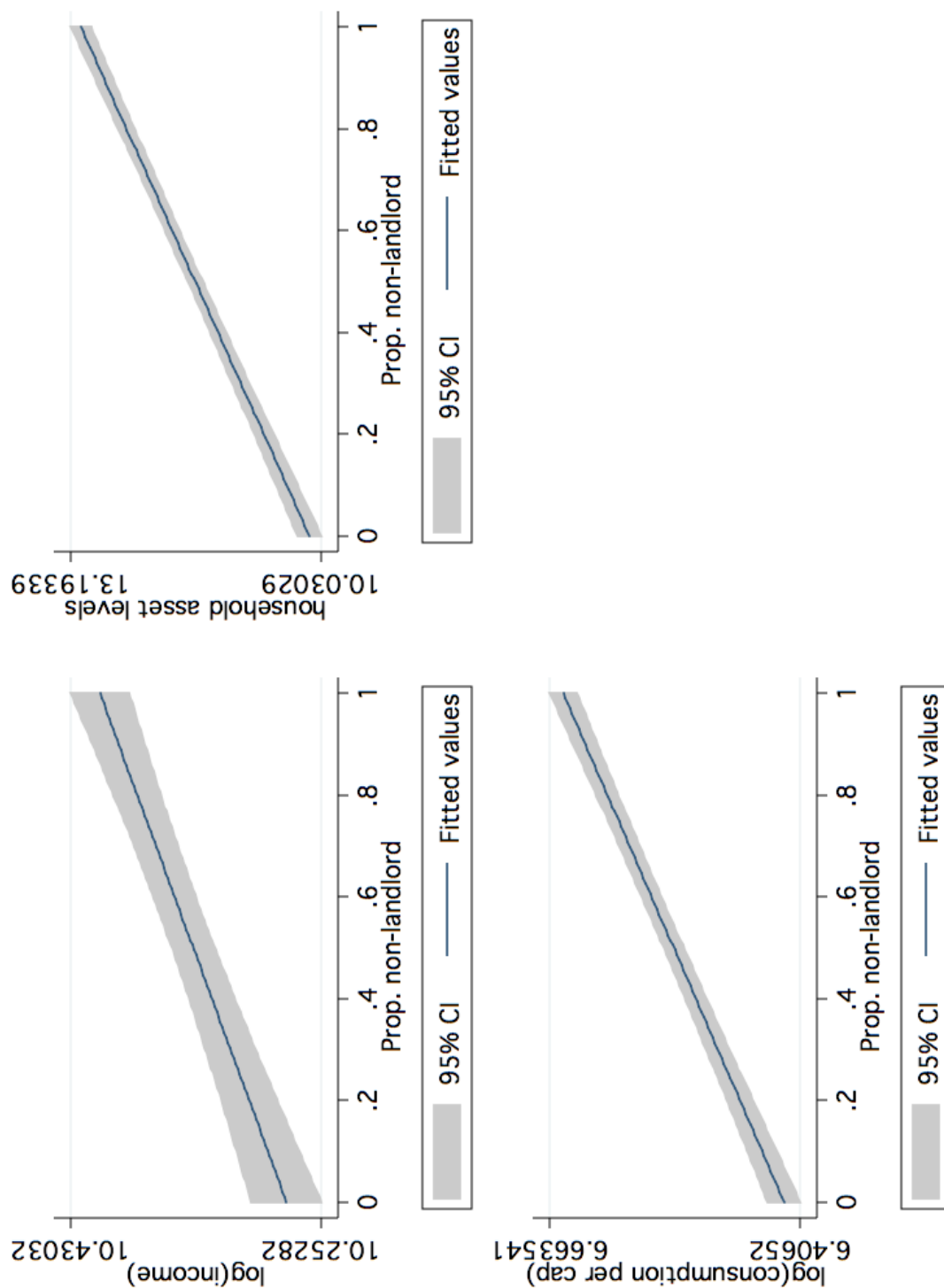
¹¹ Alternative measures of education attained include highest education completed by a male who is 21 or years older and highest education completed by a female who is 21 years or older. The difference in estimates using either of these alternate measures are negligible. We prefer to use the highest education completed by an adult who is 21 years or older to capture the effects of the most educated person in that household without disaggregating into gender differences.

Table 1: OLS Estimates of Household Economic Outcomes
Mean non-landlord proportion = 0.5802, standard deviation = 0.4393

Dependent variable	Mean of dependent variable	Coefficient on non-landlord proportion			Coefficient on non-landlord indicator		
		OLS Full Sample (1)	OLS Restricted Sample (2)	OLS Excluding West Bengal, Orissa, and Bihar (3)	OLS Full Sample (4)	OLS Individual Cultivator Only (5)	
log (household income)	10.327	0.201*** (0.047)	0.031 (0.093)	0.232*** (0.045)	0.186*** (0.040)	0.038 (0.063)	
log (household consumption per capita)	6.515	0.329*** (0.041)	0.198*** (0.067)	0.354*** (0.042)	0.167*** (0.037)	0.257*** (0.058)	
Household asset index	11.65	2.846*** (0.308)	4.974*** (0.669)	3.078*** (0.304)	1.733*** (0.274)	1.314*** (0.362)	
Number of districts		91	32	75	91	61	
Number of villages (primary sampling unit)		524	171	436	525	361	
Geographic controls		Yes	Yes	Yes	Yes	Yes	
Base controls		Yes	Yes	Yes	Yes	Yes	
Date of British revenue control		Yes	Yes	Yes	Yes	Yes	
Fractionalization index		Yes	Yes	Yes	Yes	Yes	

Notes: Standard errors in parentheses, corrected for village/neighborhood-level clustering. * Significant at 10-percent level; ** significant at 5-percent level; *** significant at 1-percent level. The restricted sample in column (2) only includes districts where there is variation in the non-landlord proportion (i.e. greater than 0 but less than 1). The non-landlord indicator variable equals 0 for landlord dominated districts and 1 otherwise. The individual cultivator indicator equals 1 if the district is largely individual cultivator dominated. Base controls include a gender indicator variable that equals one if the head of the household is a female, an indicator variable that equals one if the household is in an urban location, five caste fixed effects, twelve fixed effects for principal source of income for the household or the dominant occupation of the household, a native indicator variable that equals one if the household has been residing there longer than 70 years, the number of individuals within a household, and nineteen fixed effects for the highest level of education completed by an adult (defined as 21 or years older). Geographic controls include climatic, topographic, soil, and a coastal indicator variable which equals one if the district is on the coast. The date of British land revenue control is intended to capture the length of British rule in the district.

Figure 1: Relationship between current economic performance and proportion of non-landlord areas



Relationship between log of income (top-left), log of consumption per capita (bottom-left), household asset index (top-right) and the non-landlord proportion per district. Data for economic indicators are from 2005 IHDS and historic measures are from 1870s and 1880s from Banerjee and Iyer (2005).

landlord or non-landlord. We still find that consumption per capita in this sample is 22% higher, and household asset levels are 43% higher in non-landlord areas. We pursue yet a third strategy in column (3) by restricting our sample by excluding West Bengal, Bihar, and Orissa as these were the first areas to come under British revenue control and were almost entirely landlord areas. We want to be sure that this particular historic characteristic is not the larger part of our result. Banerjee and Iyer (2005) pursue a similar strategy but only exclude West Bengal and Bihar from their sample. However, the Treaty of Allahabad in 1765 states that the British East India Company was given rights to collect revenue from areas in present-day Orissa as well. We find that all three effects grow given this exclusion. Household income is 29% higher, household consumption levels are 42% higher, and household asset levels are 26% higher for non-landlord areas in this second restricted sample compared to our base estimation.

We conduct two additional robustness checks of our results by using indicator variables instead of proportional measures to account for the extent of non-landlord areas. The indicator is constructed following the method outlined in Banerjee and Iyer (2005): a district is non-landlord if it was never under a landlord system including changes in revenue system. Alternatively, districts with higher proportions of non-landlord systems converge to the value of 1. This method is coarser than our proportion method as many districts where the non-landlord indicator takes on the value of 1, in fact had had landlord-based revenue systems as well. Our results remain significant in all three measures, but their effect is much smaller. Household income is 19% higher, consumption per capita is 18% higher, and asset levels are 15% higher. In the next check, we compare between individual cultivator areas and non-individual cultivator areas. The income measure loses its significance but maintains a positive sign. The consumption per capita (29%

higher) and asset levels (11% higher) both remain highly significant and positive following this restriction.

These results are consistent with our framework. In non-landlord areas, the historic distribution of wealth is less unequal than landlord areas. We see the manifestation of this inequality in the long-term growth of those areas. Despite various restrictions on our sample, the difference between non-landlord areas and landlord areas are large and highly significant in both short-term and long-term economic wellbeing of households. Income and consumption, being short-term measures, show between 18% and 42% increase for households in districts that were historically under non-landlord systems. There is a large and significant effect on asset level accumulation as well, reflecting long-term depressed economic wellbeing, as non-landlord areas have between 11% and 43% higher amount of accumulated assets. While we do not pursue an instrumental variable (IV) specification as Banerjee and Iyer (2005), it is important to keep in mind that the OLS estimates are likely biased downwards and not upward due to omitted variables. In their IV results, they find that the IV coefficients are larger than OLS estimates, and conclude that while the IV estimates could be artificially higher, the estimates are still larger than OLS.

We estimate equation (1) for various measures of social capital and report them in table 2. Using our base specification, we find that there are significant positive effects in network position, community memberships, cooperation between *jatis*, collection action, confidence in institutions, and lack of local crime in non-landlord areas when compared to landlord areas. More specifically, households in non-landlord areas have 26% more connections with individuals in the medical, educational, and governmental institutions. They are also likely to have 40% more memberships. As discussed before, while we do not have exact measures of trust

Table 2: OLS Estimates of Social Capital

Dependent variable	Mean of dependent variable	Coefficient on non-landlord proportion
		OLS Full Sample (1)
Network position	1.13	0.299*** (0.084)
Membership in community groups	0.56	0.222** (0.093)
Vote in 2004 election	0.92	0.029 (0.024)
Lack of conflict in village	1.37	-0.032 (0.060)
Lack of conflict between jati	1.66	0.115** (0.039)
Collective action	0.57	0.116** (0.040)
Confidence in institutions	23.3	0.755*** (0.240)
Local crime	0.07	-0.068*** (0.014)
Geographic controls		Yes
Base controls		Yes
Date of British revenue control		Yes
Fractionlization index		Yes

Notes: Standard errors in parentheses, corrected for village/neighborhood-level clustering. * Significant at 10-percent level; ** significant at 5-percent level; *** significant at 1-percent level. Base controls include a gender indicator variable that equals one if the head of the household is a female, an indicator variable that equals one if the household is in an urban location, five caste fixed effects, twelve fixed effects for principal source of income for the household or the dominant occupation of the household, a native indicator variable that equals one if the household has been residing there longer than 70 years, the number of individuals within a household, and nineteen fixed effects for the highest level of education completed by an adult (defined as 21 or years older). Geographic controls include climatic, topographic, soil, and a coastal indicator variable which equals one if the district is on the coast. The date of British land revenue control is intended to capture the length of British rule in the district.

for districts, using an array of proxies, we can infer that individuals in non-landlord areas likely have higher levels of trust as well. Some proof can be found in the likelihood of local crime and inter-caste conflict, confidence in institutions, and propensity towards collective action. Collective action is a channel that Banerjee and Iyer (2005) hint towards as a determinant for lower levels of public goods availability and poorer economic performance. We find that collective action is indeed significantly higher in non-landlord areas. This channel, along with membership levels, local crime and conflict, describe the environment that is likely a result of the differential environment of cooperation and political control created by the land revenue systems.

Robust historic institutions create an environment of trust and efficiency which lead to robust institutions in the future. In non-landlord districts individuals were more likely to be vigilant of the quality of local institutions and the participation of individuals was likely a necessary condition for this outcome.

5. Conclusion

This paper adds to the important work by Banerjee and Iyer (2005) and other related studies that seek to better understand the role that historic institutions play on the political structure and economic wellbeing of individuals in the long-run.

We have shown that in India, depressed levels of economic performance at the micro level and low levels of social capital can be traced back to the legacy of the colonial land tenure system. Income, consumption per capita, and asset outcomes are worse for households in districts that were historically under landlord systems. In addition, in landlord areas we observe anemic social participation and low levels of other forms of social capital, while levels of victimization and communal violence are likely to be higher.

The answer likely lies in the political environment developed as a result of class-based antagonism in landlord areas. We can infer that misaligned interests between groups exacerbated collective action problems. In non-landlord areas, the need to work collectively in representing mutual interests helped cultivate clusters of “good” institutions. An empirical examination of the channel of collective action is overdue in the literature. While isolated case studies exist at the district or state-level, there is to-date no all-India study which examines the effect of historic land tenure systems on collective action. The management of common property resources for districts is good candidate for this examination. The study would have to answer the question of how rules on the usage of common property resources like watersheds, forests, and grasslands are created and enforced. While such a study might be limited in its influence on public policy or interventions in alleviating collective action problems, it will serve to inform their source. Since states in India have undergone a wide array of land reforms in the post-independence era as outlined in Besley and Burgess (2000), future research should examine specific state policy impacts on channels of collective action, and more specifically trust.

References

- [1] Alesina, Alberto, and Eliana La Ferrara, "Who trusts others?," *Journal of Public Economics*, 85 (2002), 207-234.
- [2] _____, "Ethnic diversity and economic performance," *Journal of Economic Literature*, 43 (2005), 762-800.
- [3] Anderson, Christopher and Aida Paskeviciute, "How Ethnic and Linguistic Heterogeneity Influence the Prospects for Civil Society: A Comparative Study of Citizenship Behavior," *The Journal of Politics*, 68 (2006), 783-802.
- [4] Banerjee, Abhijit and Lakshmi Iyer, "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India," *American Economic Review*, 95 (2005), 1190-1213. (Accompanying dataset is downloadable from the American Economic Association website: <http://www.aeaweb.org/articles.php?doi=10.1257/0002828054825574>)
- [5] _____, Iyer, Lakshmi and Rohini Somanathan, "History, Social Divisions, and Public Goods in Rural India," *Journal of the European Economic Association*, 3 (2005), 639-647.
- [6] _____, Iyer, Lakshmi and Rohini Somanathan, "Public Action for Public Goods," NBER Working Papers: 12911, NBER, (2006).
- [7] _____ and Rohini Somanathan, "The Political Economy of Public Goods: Some Evidence From India," *Journal of Development Economics*, 82 (2006), 287-314.
- [8] _____ and Andrew Newman, "Occupational Choice and the Process of Development," *Journal of Political Economy*, 101 (1993), 274-298.
- [9] Besley, Timothy and Robin Burgess, "Land Reform, Poverty Reduction, and Growth: Evidence from India," *Quarterly Journal of Economics*, 115 (2000), 389-430.
- [10] Bose, Sugato, *Peasant Labour and Colonial Capital: Rural Bengal since 1770*, Cambridge University Press, Cambridge, 1993).
- [11] Brown, Judith, *Modern India: The Origins of an Asian Democracy*, (Oxford University Press, New York 1994).
- [12] Chaudhary, Latika, "Determinants of Primary Schooling in British India," *Journal of Economic History*, 69 (2008), 269-302.
- [13] _____, "Taxation and Education Development: Evidence from British India," *Explorations in Economic History*, 47 (2010), 279-293.
- [14] Desai, Sonalde, Reeve Vanneman, and National Council of Applied Economic Research, New Delhi. India Human Development Survey (IHDS), 2005 [Computer file]. ICPSR22626-v8. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010-06-29. doi:10.3886/ICPSR22626.
- [15] Easterly, William, and Ross Levine, "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics*, 112 (1997), 1203-1250.
- [16] Hoff, Karla, Kshetramade, Mayuresh, and Ernst Fehr, "Caste and Punishment: The Legacy of Caste Culture in Norm Enforcement," IZA Discussion Papers: 4343, IZA, (2009).
- [17] _____ and Priyanka Pandey, "Belief Systems and Durable Inequalities: An Experimental Investigation of Indian Caste," Policy Research Working Paper Series: 3351, World Bank, (2005).
- [18] _____ and Joseph Stiglitz, "Modern Economic Theory and Development," In *Frontiers of Development Economics: The Future in Perspective*, ed. Gerald M. Meier and Joseph E. Stiglitz, (Oxford University Press, New York 2001), 389-459.
- [19] Iyer, Lakshmi, "Direct versus Indirect Colonial Rule in India: Long-term Consequences," *The Review of Economics and Statistics*, 92 (2010).
- [20] Newton, Kenneth and Pippa Norris, "Confidence in Public Institutions: Faith, Culture, or Performance," Paper presented at Annual Meeting of the American Political Science Association, (1999).

- [21] Nunn, Nathan and Leonard Wantchekon, "The Slave Trade and the Origins of Mistrust in Africa," NBER Working Papers: 14783, NBER, (2010).
- [22] Pandey, Priyanka, "Service Delivery and Corruption in Public Services: How Does History Matter," *American Economic Journal: Applied Economics*, 2 (2010), 190-204.
- [23] Peers, Douglas, *India under Colonial Rule 1700–1885*, (Harlow and London: Pearson Longmans 2006).
- [24] Robb, Peter, *A History of India*, (Palgrave Macmillan, Houndmills, Hampshire 2004).
- [25] Tabellini, Guido, "The Scope of Cooperation: Norms and Incentives," *Quarterly Journal of Economics*, 123 (2008), 905–950.
- [26] Vanneman, Reeve, Noon, James, Sen, Mitali, Desai, Sonalde, and Abusaleh Shariff, "Social Networks in India: Caste, Tribe, and Religious Variation," IHDS Working Paper: 3, IHDS, (2007)

Educational Outcomes for Latino Immigrants in Los Angeles County:
The Importance of Gender, Immigrant Generation, and Mother's Education Level

Whitney Boyer

Faculty Advisor: Jenine Harris

Washington University in St. Louis:

George Warren Brown School of Social Work

Master of Social Work Candidate

Expected Graduation Date: December 2011

Email: whitneyboyer@wustl.edu

Abstract

In the United States, the proportion of Latinos is growing at a rate faster than any other minority group; the Pew Research Center reports that Latinos have accounted for 50% of the United States population growth since the year 2000. Research since the 1960s has consistently identified a gap between Latinos and Whites in educational outcomes. In order to expand on this research, this study uses a recent data set from the L.A.FANS (2001) survey to explore the effects of theoretically supported variables on educational attainment. The final logistic regression model was significant ($\chi^2 (8)=90.27, p<0.001$), with gender, mother's educational attainment, and father's nativity as predictors of high school graduation. Contrary to prior studies, income and household language were not significant predictors of educational outcomes in this sample.

Keywords: educational outcomes, high school graduation, Latino immigrants

Educational Outcomes for Latino Immigrants in Los Angeles County:

The Importance of Gender, Immigrant Generation, and Mother's Education Level

Many research studies since the 1960s suggest that there is an educational gap between Latinos and Whites. However, this gap is not solely due to immigration status, and it threatens to increase as the Latino population rises. Iyigun (1999) writes that “schooling is a primary determinant of individuals’ earnings as well as their socioeconomic classes” (p. 707), and that “immigrant children are much more likely to be poor than their native peers” (Mistry, Biesanz, Chien, Howes, & Benner, 2008, p. 195). Yet other immigrant groups do not suffer these gaps as severely as Latinos; Schmid notes that Asian immigrants tend to have a higher socioeconomic status than Latinos, and that the education gap between these two groups is large (2001, p. 72). In fact, Latinos have higher dropout rates than most other racial groups, at 21.1% (Cheeseman Day & Jamieson, 2003, p. 6).

Furthermore, Latinos are growing at a rate faster than any other minority group in the United States (Stamps, 2006, p. 1225); the Pew Research Center reports that Latinos have accounted for 50% of the United States population growth since the year 2000 (Fry, 2008). This population growth has triggered a national immigration debate regarding the fate of Latino immigrants in the U.S. One measure that will bring economic returns to the country is ensuring adequate education to all immigrant children. In order to do this, it is essential to determine what factors have the greatest impact on Latino educational outcomes.

Background

Critical Race Theory

Critical race theory argues that a difference in the definition of success between Latinos and other Americans can lead Latinos to find legitimate reasons to drop out of high school (Villenas & Deyhle, 1999, p. 424). Since Latinos define success in terms of others, and particularly in terms of the family, they may see themselves as being more successful if they drop out of school to be able to contribute to the family. In fact, Cauce and Domenech-Rodriguez (2002) note that *Personalismo* and *Familismo* are characteristics held higher than individual ability and material success, and are part of a collective rather than individualistic culture that values family solidarity, obligation, and parental authority (p. 12). On the other hand, many other Americans define success individually, and see themselves as being more successful if they graduate from high school. In this sense, because of the way the school and educational systems are set up, critical race theory argues that Latinos may already be behind other White Americans when entering high school. However, if educators know which level of intervention most influences educational outcomes, they will be better prepared to help Latinos overcome the barriers described by critical race theory. In the end, this means Latino students may be less likely to drop out.

Language

In contrast to critical race theory, human capital theory provides a more tangible way to measure characteristics important to educational attainment. Human capital is commonly defined through concepts such as knowledge, skills, motivation, and health, which is expected to have eventual payoffs or returns in the future (Schrivver, 2004, p. 503). In this study, human capital is measured through language, mother's education level, immigrant generation, and gender. Bleakley and Chin (2008) propose that a child's English-language skills predict educational outcomes (p. 13). Reese and Goldenberg (2008) then studied the effect of community socio-

demographic characteristics and language and literacy resources on children's literacy development, and found that the language of literacy (reading in Spanish vs. reading in English), rather than literacy itself, is associated with community influences, family influences, and child income (p. 136). Similarly, Ream (2005) writes that teachers who value bi-literacy can help increase bilingual students' social capital by promoting "student engagement and leadership positions" (p. 205). Stanton-Salzar and Dornbusch (1995) also found that "Bilingualism may play a prominent role in determining access to social capital" (p. 127). More specifically, bilinguals may gain the cultural capital needed to share resources with others in the dominant culture, while at the same time maintain enough trust that the education system will provide cultural accommodation that "will ultimately produce desired returns" (p. 132).

Parental Education

Besides language, parental educational attainment is another important measure of human capital. Iyigun (1999) notes that "educated parents create a better learning environment at home, directly influencing the academic potential of their children" (p. 698). Rumberger (1983) reports that children may look at their same-sex parent as a role model for how much education they should have (p. 207). This may be why mother's education is an important predictor of educational outcomes; traditionally, males have more education, so it is easier for boys to see from their fathers that education is important. However, since women traditionally have less education, a mother's education may be an important predictor of how much education a girl receives, especially if she does look to her mother as a role model. Furthermore, Mistry, et al. describe the fact that socioeconomic status affects the cognitive ability of a child in the future, partly through the mother's ability to provide an educationally stimulating environment in the home (2008, p. 207).

Immigrant Generation and Gender

Additionally, immigrant generation is a further component of human capital. Schmid (2001) found that “those whose two parents were foreign born had significantly higher chances of completing high school than did those whose parents were U.S. Born” (p. 74). Wojtkiewicz and Donato (1995) discovered that foreign birth, family structure, and parental education are all important predictors of high school completion. In addition, Pan’s (2007) study revealed gender, mother’s language preference, and family structure to be related to educational achievement (p. 20).

Family Income

One measure of financial capital included in this study is income. Driscoll’s study found that having a higher family income is a protector against early dropout (1999, p. 871). Rumberger (1983) found that having more reading material in the house, as well as a higher socioeconomic status was associated with less high school dropouts (pp. 206, 210). The probability of having reading material in the house, however, may be partly attributable to socioeconomic status. Finally, Warren (1996) found that a majority of the gap between Mexicans and Whites, and almost all between Chicanos and Whites, can be attributed to family background characteristics, irrespective of ethnicity (p. 150).

With such background in mind, this paper uses a different data set than those used for similar studies in the past, aiming to confirm whether household language, mother’s educational attainment, father’s nativity, gender, and income are significant predictors of high school graduation.

Methodology**Data Set**

This research used the Los Angeles Family and Neighborhood Study (L.A.FANS), which surveyed 3,085 households from a stratified random sample of 65 neighborhoods in Los Angeles County, California, with an oversample of poor neighborhoods and households with children under 18. Analysis was conducted on respondents who identified as being Latino(a) ($n=3,832$).

Analysis

The variables for the final model were identified based on several factors. First, there is literature to support all five variables as having an impact on educational outcomes. Second, many of the variables were previously shown through the researcher's bivariate and stepwise regression analysis to significantly predict high school graduation. Furthermore, the final variables seemed less likely to be collinear, making the model less biased and bringing each variable to a more accurate portrayal of its significance. The final model analyzed with logistic regression included Father's nativity (Born in U.S. or not), Gender, Household language, Income, and Mother's educational attainment (recoded into Primary, Middle School, Some Secondary, Graduated from High School, and Attended or Completed Higher Education). The dependent variable "Have you graduated from high school?" was recoded into a binary variable indicating whether or not the respondent graduated or earned a GED (yes=1; no=0). In order to further ensure the study's model was unbiased and had the best fit, diagnostics were then conducted to test for linearity, independence of errors, multicollinearity, outliers, and influential cases. In the final model, the sample size was 661.

Results and Discussion

Bivariate Analysis

The following Table 1 shows descriptive statistics resulting from a bivariate analysis, including the mean income of the sample, as well as the t and χ^2 statistics:

Table 1: Bivariate Statistics for 1,531 Latino Respondents

	Graduated (n=1223)		Did Not Graduate (n=308)		t (df); p
	Mean	SD	Mean	SD	
Income (n=731)	\$30,464.53	\$32,398.72	\$27,700.00	\$27,048.76	t (278.78) = -1.08; p>0.05
	n	%	n	%	χ^2 (df); p
Gender (n=1530)					χ^2 (1) = 2.39; p>0.05
Male	593	48.5	164	53.4	
Household Language (n=1140)					χ^2 (1) = 0.001; p>0.05
English	548	60.6	142	60.4	
Spanish	357	39.4	93	39.6	
Father's Nativity (n=1508)					χ^2 (1) = 65.57; p<0.001
Born in U.S.	684	56.6	91	30.4	
Born in Another Country	525	43.4	208	69.6	
Mother's Educational Attainment (n=1407)					χ^2 (4) = 152.66; p<0.001
Primary	123	10.7	91	35.4	
Middle School	162	14.1	58	22.6	
Some Secondary	96	8.3	34	13.2	
Graduated High School	379	33.0	46	17.9	
Higher Education	390	33.9	28	10.9	

Based on a comparison of graduation status and demographics, only Father's Nativity (χ^2 (1) = 65.57; p<0.001) and Mother's Educational Attainment (χ^2 (4) = 152.66; p<0.001) are significantly related to whether or not the respondent graduated. While 56.6 percent of those whose fathers were born in the United States graduated from high school, 69.6 percent of those whose fathers were born abroad did not graduate. Similarly, only 8.3 percent of those whose mothers had some secondary education ended up graduating from high school themselves, and 35.4 percent of those whose mothers had only a primary level of education did not graduate from high school. However, 33 percent of those whose mothers graduated from high school also graduated from high school.

Logistic Analysis

A logistic regression analysis to examine the relationship between high school graduation and household language, mother's educational attainment, immigrant generation, gender, and

income resulted in the following model:

$$\begin{aligned}
 \text{logit (high school graduation)} \\
 = & 17.26 - 2.03 (\text{mother's primary education}) \\
 & - 1.28 (\text{mother's middle school education}) \\
 & - 1.42 (\text{mother's secondary education}) \\
 & - 0.30 (\text{mother's high school graduation}) + 0.45 (\text{gender}) \\
 & - 0.65 (\text{father's nativity}) - 0.10 (\text{household language}) + 0.00 (\text{income})
 \end{aligned}$$

This overall model is significant, at $\chi^2 (8)=90.27$, $p<0.001$, and Nagelkerke's $R^2 = 0.21$.

The significant predictors are gender, father's nativity, and mother's educational attainment – through primary, middle school, and some secondary, as compared to higher education. The odds ratio for gender is 1.56 (CI=1.02-2.39), for father's nativity 0.52 (CI=0.31-0.86), and within mother's educational attainment, 0.13 (CI=0.06-0.27) for primary, 0.27 (CI=0.13-0.58) for middle school, and 0.24 (CI=0.11-0.54) for some secondary. In other words, people whose mothers did not graduate from high school are less likely to graduate themselves, compared to people whose mothers received a higher education. Those whose mothers attained a primary education are 86.9 percent less likely to graduate from high school, those whose mothers had a middle school education are 72.3 percent less likely to graduate, and those whose mothers achieved some secondary education are 75.9 percent less likely to graduate from high school. Similarly, those whose fathers were born outside the United States are 48 percent less likely to graduate from high school. Finally, the odds of graduating from high school are 1.56 higher for females than for males.

Diagnostics

This study assessed the assumptions and conducted diagnostics to determine if the model was unbiased and a good fit for the data. Multicollinearity and the tolerance level were tested, and neither violated their assumptions. Thus, this model is unbiased and can be generalized to a larger population. For diagnostics, studentized residuals were plotted against the case numbers,

and independence of errors was assumed because the graph was homoscedastic. The Cook's D and leverage values shared 12 influential cases, so the model was run again without them. The results were in fact not much different than the model that included the 12 influential cases and outliers; the same predictors were significant in both models, and the model without the 12 cases made those predictors only slightly more significant. Because of the ethical issues involved in taking away data for the sake of good diagnostics, these 12 cases were returned to the data set, and this is the model analyzed in this results and discussion section. In the end, it was concluded that this model is a good fit for the data.

Predicted Probabilities

The following Table 2 provides predicted probabilities for six cases within the full data set:

Table 2: Selected Predicted Probabilities for Six Latinos, Half of Which Graduated

Mother's Educational Attainment	Gender	Income	Household Language	Father's Nativity	Predicted Probability of Graduating from High School	Actual High School Graduation Outcome
Higher Education	Female	30,000	English	U.S.	0.96	Graduated
Graduated High School	Male	15,000	Spanish	U.S.	0.90	Graduated
Higher Education	Female	13,000	Spanish	U.S.	0.95	Graduated
Primary	Male	13,000	Spanish	Another Country	0.44	Not graduated
Primary	Female	20,000	English	Another Country	0.58	Not graduated
Primary	Male	34,000	Spanish	Another Country	0.46	Not graduated

As can be seen from the table, the first three cases graduated, while the last three did not. These selected cases show the trends suggested by the results mentioned above. First, the respondents' income varies within the outcome variable; although one individual who graduated has an income of \$30,000, the other two who graduated make \$13,000 and \$15,000, and one individual who did not graduate has an income of \$34,000. This is consistent with the

insignificant chi-squared test and beta value for income. Similarly, household language also varies across the outcome variable. However, those who did not graduate had fathers who were born outside the country, and mothers with only a primary level education. Finally, although there is variation in gender, two of the three who graduated were females, whereas two of the three who did not graduate were males.

One interesting thing to note here is that one respondent's predicted probability was 0.58, and was predicted to graduate but did not. This could have to do with how well the model predicted the actual outcomes; Table 3 shows that the model was more sensitive than specific, meaning it was more accurate at predicting those who graduated than those who did not graduate. Ninety-seven percent of those who graduated were predicted correctly, whereas only 19.2 percent of those who did not graduate were predicted correctly.

Table 3: Observed vs. Predicted Graduation, Using Above Regression Model

Observed	Predicted		% Correct
	Yes	No	
Yes	522	14	97.4
No	101	24	19.2
Total			82.6

Limitations

This study has three major limitations. First, although it focused on the fastest-growing immigrant group with one of the highest education and income gaps, future research should include comparison groups. This will allow researchers and policy-makers to better understand how Latinos differ from others. Second, this model's immigrant generation variable provided a different result than one of the articles read in preparation for the study. Schmid (2001) found that "those whose two parents were foreign born had significantly higher chances of completing high school than did those whose parents were U.S. Born" (p. 74). However, this study's model

predicts the opposite. Further research should therefore be conducted to explore the phenomenon of immigrant generation and its effects on educational outcomes.

The third limitation of this study is that the data is cross-sectional. As such, the family income and household language variables are indicative of the children's educational environment, rather than the adult respondents' environment at the time when they were going through school. Time order therefore likely prevented significance and causation in this study. When the second wave of the L.A.FANS becomes public data, it may be useful to conduct a similar study using the children as respondents, to find whether the variables that were insignificant in this study become significant. Furthermore, future research should use a measurement of the children respondents' educational outcomes, such as achievement, in order to include additional community and family variables that present a broader picture of a Latino's environment. Such additional characteristics could be broken down ecologically, through community, family, and individual factors, or by additional types of capital – social, human, and financial.

Conclusions

Consistent with prior research, this study provides additional evidence that mother's educational attainment, gender, and father's nativity are all significantly associated with the likelihood of high school graduation for Latinos. According to the above model, the following Latino populations are less likely to graduate from high school: males, those whose mother did not graduate from high school, and those whose father was born in another country.

Ultimately, with the Latino population growth, the increasing gap in education, and the intensifying immigration debate, it is crucial to provide accurate information about what most helps these people succeed once arriving in the country. Decision-makers will be better informed

and more able to spend local, state, and federal dollars more efficiently, if they know where to direct interventions, and which interventions are more effective.

With this in mind, future social workers, educators, and community members can create interventions geared specifically towards second-generation males whose mothers did not receive a high school degree or higher education, knowing this is a population that likely needs extra support in navigating the educational system. Even within programs that already exist, such as those for English-as-a-second-language students, this study suggests one best practice for leaders of these programs: a specific targeting of second-generation males whose mothers did not graduate from high school. These interventions will then lead to higher educational attainment and socioeconomic status for Latinos, and a greater contribution to the economy and society of this country.

References

- Bleakley, H. & Chin, A. (2008). What Holds Back the Second Generation? The Intergenerational Transmission of Language Human Capital Among Immigrants. *Journal of Human Resources*, 43(2). Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2440707/>
- Cauce, A.M. & Domenech-Rodriguez, M. (2002). Latino Families: Myths and Realities. In J.M. Contreras, K.A. Kerns, & A.M. Neal-Barnett (Eds.), *Latino Children and Families in the United States: Current Research and Future Directions* (pp. 3-26). Westport, CT: Praeger Publishers.
- Cheeseman Day, J. & Jamieson, A. (2003). School Enrollment: 2000. *Census 2000 Brief*. Retrieved from <http://www.census.gov/prod/2003pubs/c2kbr-26.pdf>
- Driscoll, A.K. (1999). Risk of High School Dropout among Immigrant and Native Hispanic Youth. *International Migration Review*, 33(4), 857-875. Retrieved from <http://www.jstor.org/stable/2547355>
- Fry, R. (2008). Latinos Account for Half of U.S. Population Growth Since 2000. *Pew Research Center Publications*. Retrieved from <http://pewresearch.org/pubs/1002/latino-population-growth>
- Iyigun, M.F. (1999). Public Education and Intergenerational Economic Mobility. *International Economic Review*, 40(3), 697-710. Retrieved from <http://ideas.repec.org/a/ier/iecrev/v40y1999i3p697-710.html>
- Mistry, R.S., Biesanz, J.C., Chien, N., Howes, C., & Benner, A.D. (2008). Socioeconomic Status, Parental Investments, and the Cognitive and Behavioral Outcomes of Low-

- Income Children from Immigrant and Native Households. *Early Childhood Research Quarterly*, 23, 193-212. Retrieved from <http://www.sciencedirect.com/science>
- Pan, Y. (2007). *Gains from Green Cards: Immigrant Parents' Legal Status and Children's Scholastic Achievement*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1071202
- Pebbley, Anne R., and Narayan Sastry. Los Angeles Family and Neighborhood Study (L.A.FANS) [Computer file]. ICPSR22940-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2009-11-09. doi:10.3886/ICPSR22940
- Ream, R.K. (2005). Toward Understanding How Social Capital Mediates the Impact of Mobility on Mexican American Achievement. *Social Forces*, 84(1), 201-224. Retrieved from <http://www.jstor.org/stable/3598301>
- Reese, L. & Goldenberg, C. (2008). Parental Involvement and the Academic Achievement of Hispanic Students. *Marriage and Family Review*, 43(1/2), 109-139. doi: 10.1080/01494920802010272
- Rumberger, R.W. (1983). Dropping out of High School: The Influence of Race, Sex, and Family Background. *American Educational Research Journal*, 20(2), 199-220. Retrieved from <http://www.jstor.org/stable/1162594>
- Schmid, C.L. (2001). Educational Achievement, Language-Minority Students, and the New Second Generation. *Sociology of Education*, 74, 71-87. Retrieved from <http://www.jstor.org/stable/2673254>
- Schriver, J.M. (2004). *Human Behavior and the Social Environment*. Boston: Pearson Education, Inc.

- Stamps, K. & Bohon, S.A. (2006). Educational Attainment in New and Established Latino Metropolitan Destinations. *Social Science Quarterly*, 87(5), 1225-1240. Retrieved from <http://paa2006.princeton.edu/download.aspx?submissionId=60028>
- Stanton-Salzar, R.D. & Dornbusch, S.M. (1995). Social Capital and the Reproduction of Inequality: Information Networks among Mexican-Origin High School Students. *Sociology of Education*, 68(2), 116-135. Retrieved from <http://www.jstor.org/stable/2112778>
- Villenas, S. & Deyhle, D. (1999). Critical Race Theory and Ethnographies Challenging the Stereotypes: Latino Families, Schooling, Resilience, and Resistance. *Curriculum Inquiry*, 29(4), 413-445. Retrieved from <http://www.jstor.org/pss/3185923>
- Warren, J.R. (1996). Educational Inequality among White and Mexican-Origin Adolescents in the American Southwest: 1990. *Sociology of Education*, 69(2), 142-158. Retrieved from <http://www.jstor.org/stable/2112803>
- Wojtkiewicz, R.A. & Donato, K.M. (1995). Hispanic Educational Attainment: The Effects of Family Background and Nativity. *Social Forces*, 74(2), 559-574. Retrieved from <http://jstor.org/stable/2580492>