

ESSAYS ON RESIDENTIAL SEGREGATION, THE ECONOMICS OF PRESTIGE AT
THE ELITE COLLEGES, AND PROPERTY RIGHTS

Peter Q. Blair

A DISSERTATION

in

Applied Economics

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2015

Supervisor of Dissertation

Todd Sinai, Professor, Real Estate and Business Economics and Public Policy

Graduate Group Chairperson

Eric Bradlow, Professor of Marketing, Statistics, and Education

Dissertation Committee

Fernando Ferreira, Associate Professor of Real Estate and Business Economics and Public Policy; Gilles Duranton, Dean's Chair in Real Estate Professor; Kent Smetters, Boettner Professor of Business Economics and Public Policy

ESSAYS ON RESIDENTIAL SEGREGATION, THE ECONOMICS OF PRESTIGE AT
THE ELITE COLLEGES, AND PROPERTY RIGHTS

© COPYRIGHT

2015

Peter Q. Blair

This work is licensed under the
Creative Commons Attribution
NonCommercial-ShareAlike 3.0
License

To view a copy of this license, visit

<http://creativecommons.org/licenses/by-nc-sa/3.0/>

Dedicated to

I dedicate this thesis to my mom Judith Carol Blair, who sacrificed finishing her PhD in order to raise my six brothers and me. Mom your love for us and your dedication to seeing us be all that we can be inspires me. Many times when I was feeling uncertain or discouraged along this process you were always there to pray with me, encourage me and challenge me to do something that took me out of my comfort zone. I love you more than the words and equations in this thesis can express – they are all for you and because of you Mom.

ACKNOWLEDGEMENT

So many people have journeyed with me over the past 14 years that I have been in school – nine of them in graduate school. As I turn to thank these journey mates, I am reminded of the words of the Jewish King Solomon, who said “Except the Lord build the house, they labor in vain that built it.” I thank the Lord for being with me through the highs and lows of graduate school and for gifting me with incredible advisers and dedicated friends.

To my committee, Todd, Fernando, Kent and Gilles – thank-you for sharpening me through this process. Todd, I value your deep economic intuition and the way that you understand my strengths and weaknesses. Fernando, thank-you for always making the time to talk with me about research and for pushing me to cultivate my empirical skills. Kent, thank-you for being open-minded, admitting this particle physicist-turned economist to this program. I have enjoyed working with you on our paper, and learning from you how to build simple, yet powerful models! Gilles, in addition to being awesome (just by virtue of being French), I value the way that you challenged me to think and figure out what was deeply wrong, and in some cases, correct, about the work that I did :). I also owe a special debt of gratitude to Maisy Wong, who met with me consistently during my first few years of graduate school and encouraged me to begin writing the first chapter of this thesis during my first year at Wharton. Thank-you, Maisy. I also thank Prof. Susan Wachter, Stuart Hean, Lynn Selhat as well as my fellow graduate students at Wharton for generously giving their time to help me develop as a scholar.

To my friends Daniel DeVougas, Mickey Whitzer, Dion Cunningham, Kinde Wubneh, Tout Wang, Sorell Massenberg, Stephen Lafata, Mike and Hannah Rizk, Eliud Ramocan, Luke Stewart, Thummim Cho, Stephen Odaibo, Deriba Olana, Konrad Osa, Chris Trimble, Philip Winter, Franco Nilo, Joseph Bataille, Daniel Summerhayes, Neil Abraham, Charles Robinson, Nicholas Litombe, Greg Akingbiyi, Fred Roland, Reggie Roberson, Marthony Robins, Kory Efunnuga, David Stremic, Gabe Bouch, Nadia Souri, Nico Cernek, Onoso Imoagene

the ManSlon and Antioch Calvary Chapel Communities and Rafael Luna – thank you for being such an integral part of this journey. Your prayers and encouragements have sustained me.

To my family members, my dad Leslie Blair, brothers Steve, Julian, Marvin, Andrew (deceased), Tim, and Paul; and my second mom Janet Jibberson and my nephew Moses and my nieces Chassity-Twyla and Sierra-Kennedy, thank you for always challenging me to be my best.

I also could not have done this PhD without the financial support of the Wharton School, the Mustard Seed Foundation, the National Science Foundation, the Amy Morse Foundation, the Government and people of the Bahamas, the Lyford Cay Foundation, the Mellon Mays Foundation, and the Bell Labs Foundation. Importantly, I owe deep debt of gratitude to Mrs. Sheila Ferguson, Prof. Melissa Franklin and Prof. Chris Stubbs of the Harvard Physics Department, who encouraged me to cultivate my interest in economics during my time at Harvard, even as I pursued a PhD in Physics. I am also grateful to the following economics professors, who supported me during my first forays into economics, while I was contemplating the transition from physics to economics: Edward Glaeser, David Laibson, Devin Pope, Ernst Fehr, Chris Avery and Roland Fryer.

As an undergraduate student, my initial passion for research was sparked by Duke Professor Arlie Petters, who is a deeply influential and inspirational mentor. Arlie, thank you believing in me, expecting nothing but the highest level of excellence from me and for being an example to me of how to do it and succeed despite the odds. Lastly, I wish to express my gratitude to some of my earliest teachers – Julie Arthur, Sheri Anderson and Janet Jibberson who are responsible for seeing talent in me and cultivating it. I hope to pay it forward by mentoring others as you have mentored me. There are innumerable others that are not mentioned here but who have helped me throughout the process. Thank-you all!

ABSTRACT

ESSAYS ON RESIDENTIAL SEGREGATION, THE ECONOMICS OF PRESTIGE AT THE ELITE COLLEGES, AND PROPERTY RIGHTS

Peter Q. Blair

Todd Sinai

This dissertation consists of three chapters. In the first chapter of the dissertation, I propose a new method for estimating the racial tipping points of US neighborhoods. Applying this method to US census data, I find that the average neighborhood tipping point has increased from 15% in 1970 to 42% in 2010. Moreover, I find that from 1970-2010 that differences in racial preferences play a diminishing role in explaining cross city differences in tipping points. Differences in the outside options of white household's, on the other hand, play an increasingly important role in explaining cross city differences in tipping points from 1970-2010.

In the second chapter of the dissertation, which is joint work with Kent Smetters, we document an important fact about Ivy League Schools. In the past 13 years the demand for undergraduate admission to Ivy League schools has doubled while the number of available spaces has increased by less than 3%. We show that this pattern can be explained by a model in which elite schools restrict supply in order to maintain their relative selectivity. Focusing on Harvard, Yale and Princeton, we find that the resulting loss in social surplus is an average of \$73M for each school (annually), with producers (schools) bearing the larger absolute loss in social surplus and consumers (students) bearing the larger relative loss in social surplus.

In the third chapter of this dissertation, I use a two country treatment-control design to study an economic policy in the Bahamas that restricted the property rights of non-natives. Contrary to the prevailing view in the development literature that restricted property rights

reduces GDP growth, I find that GDP growth in the Bahamas increased while this policy was in place (relative to a peer country Barbados, which did not have a similar policy). When the policy is repealed, I find that there are large inflows of foreign capital to the Bahamas with no corresponding increase in GDP growth. This study highlights the importance of differentiating between the property rights of natives and non-natives as separate channels for promoting economic growth in small developing countries.

ACKNOWLEDGEMENT	iv
ABSTRACT	vi
LIST OF TABLES	x
LIST OF ILLUSTRATIONS	xii
PREFACE	xiii
CHAPTER 1 : The Effect of Outside Options on Neighborhood Tipping Points . .	1
1.1 Introduction	2
1.2 Model	5
1.3 Tipping Point	9
1.4 Data	14
1.5 Results	15
1.6 Conclusion	26
CHAPTER 2 : Why Don't Elite Colleges Expand Supply?	28
2.1 Introduction	28
2.2 Model	31
2.3 Welfare Analysis	37
2.4 Equilibrium Selection	42
2.5 Parameter Estimation	43
2.6 Data	48
2.7 Empirical Framework	48
2.8 Results	52
2.9 Conclusions	53

CHAPTER 3 : Property Rights and Economic Growth: A Caribbean Natural Ex-	
periment	54
3.1 Introduction	54
3.2 Literature Review	56
3.3 Addressing Endogeneity	59
3.4 Data	61
3.5 Identification	62
3.6 Results	64
3.7 Robustness Checks	68
3.8 Conclusion	72
APPENDIX	73
BIBLIOGRAPHY	108

List of Tables

TABLE 1 : Relative Marginal Utility of Minority Neighbors (Diff-in-Diff)	73
TABLE 2 : Census Tract Tipping Points (Diff-in-Diff)	73
TABLE 3 : Census Tract Tipping Point (Semi-Parametric)	76
TABLE 4 : Correlation Tract Tipping Points (Semi-Parametric)	76
TABLE 5 : Average MSA Tipping Points	79
TABLE 6 : Median MSA Tipping Points	80
TABLE 7 : Correlation in Mean MSA Tipping Points	81
TABLE 8 : Comparison of CMR and Blair Tipping Points	81
TABLE 9 : Effect of Preferences and Options on Tipping Points	81
TABLE 10 : Ivy League Admissions Numbers (2003-2015)	92
TABLE 11 : Characteristics of HYP Ivies: Harvard, Yale, & Princeton (2003-2015)	92
TABLE 12 : Characteristics of non-HYP Ivies (2003-2015)	93
TABLE 13 : Estimating the Demand Elasticity of Supply at HYP Ivies	93
TABLE 14 : Estimated Parameters of the Model (2003-2015)	93
TABLE 15 : Welfare Results for HYP Ivies (2003-2015)	94
TABLE 16 : Effect of Law Change on GDP Growth	109
TABLE 17 : Using Legal Reform as Instrument for Net FDI	110
TABLE 18 : Robustness Checks, Counter-factuals, and Placebo Tests	111
TABLE 19 : Robustness Checks: Auto-correlation	111
TABLE 20 : Contemporaneous Difference Estimates	112

List of Figures

FIGURE 1 :	Kernel density plot of estimated census tract tipping points of Jersey City, NJ and Mobile, AL in 1970.	74
FIGURE 2 :	Kernel density plot of the racial preference parameter for each MSA from 1970 to 2010, using the diff-in-diff estimates.	75
FIGURE 3 :	Distribution of census tract tipping points for each census year using the tipping points from the semi-parametric (inverse) method.	77
FIGURE 4 :	Distribution of mean MSA tipping points for each census year using the tipping points from the semi-parametric (inverse) method.	78
FIGURE 5 :	Kernel density plot of the difference between the CMR tipping points and mean MSA tipping points using all tracts, and using the marginal census tracts that are within 5 percentage points of their tipping point.	82
FIGURE 6 :	Kernel density plot of the difference between the CMR tipping points and median MSA tipping points using all tracts, and using just the marginal census tracts that are within 5 percentage points of their tipping point.	83
FIGURE 7 :	A plot of coefficient estimates from a regression of MSA tipping points on standardized measures of racial preferences and Herfindahl-Hirschman Indices (HHI) of the minority and non-minority concentration of the MSA in census years C1970-C2010.	84
FIGURE 8 :	Number of College Applicants and Admits at Harvard, Yale, Princeton and Penn (2003-2015)	94
FIGURE 9 :	Number of College Applicants and Admits at Columbia, Brown, Dartmouth, Cornell (2003-2015)	95
FIGURE 10 :	Estimated Parameters from College Prestige Game (2003-2015)	95

FIGURE 11 : Country Comparison.	108
FIGURE 12 : Effect of Reduced Property Rights for Non-Natives on per capita Net FDI.	112
FIGURE 13 : Effect of Reduced Property Rights for Non-Natives on per capita GDP Gap.	113
FIGURE 14 : Effect of Property Rights for Non-Natives on per capita Net FDI.	114
FIGURE 15 : Effect of FDI Inflow on per capita GDP Gap.	115

PREFACE

“There are those who seek knowledge for the sake of knowledge; that is Curiosity.
There are those who seek knowledge to be known by others; that is Vanity.
There are those who seek knowledge in order to serve; that is Love.” (Bernard
of Clairvaux)

These essays highlight my deep interest as an economist into researching issues of access and equity in the domains of housing, education and economic development in the Third World.

CHAPTER 1 : The Effect of Outside Options on Neighborhood Tipping Points

Abstract: Recent estimates of tipping points in the literature suggest that racial progress in the United States has been slow. According to these estimates, the mean tipping point of US cities increased by 1 percentage point per decade – from 12% in 1970 to 14% in 1990 (Card, Mas, and Rothstein 2008). I develop a new method for estimating the tipping points of census tracts within cities. These new estimates paint a more optimistic picture of racial progress in the United States: the mean tract-level tipping point has increased from 15% in 1970 to 42% in 2010; and the mean city tipping point has increased from 13% in 1970 to 35% in 2010. I show that the previous estimates of city or metropolitan statistical area (MSA), tipping points are smaller than my estimates because they are local averages of the tipping point of the marginal census tracts – the tracts that are close to tipping. Because I estimate tipping points at the census tract level, my city tipping points are averages of the tipping points of both marginal and infra-marginal census tracts.

1.1. Introduction

A neighborhood tips when a marginal increase in its minority population leads to white flight from the neighborhood. Given the link between racial segregation and adverse outcomes for minorities, it is important to understand the mechanisms that drive neighborhood tipping.¹ Furthermore, neighborhood tipping points can be important parameters for place-based policies like the Moving to Opportunity Experiment (MTO). In place-based programs where the treatment is the destination neighborhood, it may be important to discern whether the act of assigning minority households to a destination might itself result in the neighborhood tipping, thereby undermining the intended treatment (Kling, Liebman, and Katz 2007).

I define a household's *neighborhood* as the census tract where it resides, and its *outside options* as the set of all of the other census tracts in its MSA of residence. I impose an incentive compatibility constraint on the exit decision of non-minority households. Non-minority households exit their current neighborhood only in the face of racial integration if exercising their outside option delivers more utility than remaining in their current neighborhood. By imposing this constraint, I am able to exploit sorting patterns in the data to estimate static tipping points at the census tract level which are defined within the context of a discrete choice model of housing (McFadden 1978; Berry, Levinsohn, and Pakes 1995; Bayer et al. 2007). The two main contributions of this papers are: (i) I provide a method for computing census tract tipping points (ii) I produce estimates of census tract tipping points for the census tracts of 123 US cities that cover the five most recent censuses (1970-2010).

Computing tipping points for individual census tracts is a methodological contribution to the empirical literature on tipping, which has progressed from the national level (Easterly 2009) to estimating tipping points at the MSA level (Card et. al. 2008). In some cases, the distribution of census tract tipping points is disperse and an MSA average may mask this

¹Segregation impacts the provision of public goods to minorities, particularly if preferences for redistribution are local (Zeckhauser 1993; Bayer, McMillan, and Rueben 2005). Additionally, residential segregation creates spatial mismatches: minorities are disconnected from jobs, role models, and opportunities to interact with non-minorities, which may result in the persistence of racial stereotypes (Kain 1968).

heterogeneity. Mobile, AL and New Jersey provide an illustrative case (Figure 1). In 1970, both cities have a mean MSA tipping point of 22%. In Mobile, the dispersion about this mean is large, whereas for New Jersey there is relatively less dispersion about this mean. As such, an MSA-wide place-based policy is more likely to work consistently in New Jersey, whereas Mobile would require more locally targeted policy to account for the heterogeneity in tipping points by census tract.

In computing the tipping points for the 38,466 tracts in my data, I find that the mean neighborhood (tract) tipping point in the United States has increased at a rate of 6 percentage points per decade – from 15% in 1970 to 42% in 2010. To compare these results with the literature, I aggregate my tract tipping points to the city level and find that the mean metropolitan statistical area (MSA) tipping points also increased at an average rate of 5 percentage points per decade from 13% (1970) to 35% (2010). According to other estimates in the literature, the mean tipping point of US cities increased from 12% in 1970 to 14% in 1990 – an average of 1 percentage point per decade (Card, Mas, and Rothstein 2008). I show that prior estimates understated city tipping points because they reflected the average tipping points of only marginal census tracts in the city (i.e., those that were close to tipping), whereas my estimates are an average of the tipping points of both the marginal and infra-marginal census tracts in a city.

In the data, I also find evidence that cross-sectional differences in city tipping points depend crucially on the outside options of non-minority households. Moreover, the relative importance of outside options in explaining cross-sectional differences in city tipping points vis-à-vis racial preferences is increasing over time. In 1970, a decrease of one standard deviation in the clustering of minorities was correlated with a 1.1 percentage point increase in the city tipping point.² In 2010, a similar decrease in the clustering of minorities was correlated with an increase of 14 percentage points in the tipping point. By contrast, in 1970, a one-standard-deviation increase in the relative preference of non-minorities for mi-

²A reduced level of minority clustering means that the outside option consists of fewer tracts with only non-minority households.

nority neighbors was correlated with a 6.6 percentage point increase in the mean MSA tipping point; in 2010, however, the effect was statistically indistinguishable from zero. Since outside options become increasingly relevant over time, and racial preferences become less predictive of cross-sectional differences in tipping points, it is important to model tipping phenomena with outside options playing a key role. Moreover, in many branches of economics, outside options matter for modeling the decision of agents.³

Historically, the debate between sociologists and economists on tipping centered on whether neighborhood tipping reflects racial animus of non-minorities (whites) toward minorities. In an early essay, “The Metropolitan Areas as Racial Problem,” University of Chicago political scientist Morton Grodzins (1958) asserted that neighborhood tipping reflects “the unwillingness of white groups to live in proximity to large numbers of [African Americans].” Later work by economists, notably Thomas Schelling (1969, 1971), challenged this view, demonstrating with a set of intuitive and simple models that segregation, at the neighborhood and city levels, could occur even if white households, at the individual level, did not possess a strong aversion to living in communities with minorities. One of the limitations of this model, as Schelling himself noted, was that it did not include a role for outside options: “This is but a small sample of possible results, using straight-line schedules and simple dynamics. There are no expectations in the model, no speculation, no concerted action, no restriction on the alternative localities available” (Schelling 1969).

The key insight of this paper is that the tipping threshold for a non-minority household depends on its preferences for minority neighbors and on the household’s outside options. In some cases, the dearth of preferable outside options will result in a tipping threshold that is high (more racial integration is tolerated) even though non-minority households have a strong relative preference for living with other non-minorities. As an example, consider

³The canonical principal agent model of contract theory (Jullien 2000) is one example in economics where outside options matter for decision making by individual agents and the firm. In job search models in macroeconomics (McCall 1970; Mortensen and Pissarides 1994), outside options matter for aggregate market outcomes such as mean unemployment duration. Outside options have also been used to illustrate scenarios in which the Coase Conjecture fails (Board and Pycia 2014).

a city with a choice-set of two census tracts: tract W, which is 100% non-minority, and tract M, which is 100% minority. Suppose further that non-minority households in the city have a relative preference for non-minority neighbors. What happens if minorities integrate tract W? Would non-minority households exit tract W in preference for tract M? No. Non-minority households are better off if they remain in tract W, which has marginally fewer non-minorities (even after integration) than if they exit to tract M, which is maximally minority occupied. In fact, contrary to the intuition that a stronger relative preference for same-race neighbors leads to more white flight, the stronger the white household's preferences for white neighbors, the more likely it will be to remain in tract W even as the neighborhood integrates. In the same way that competition for employees among firms sets the cost of discriminating against minority employees (Becker 1993), the availability of preferable outside options sets the cost of acting on racial preferences in the housing market.

The rest of the paper is organized as follows: In section 2, I model the neighborhood choice of households. In section 3, I define the tipping point. I discuss the data in section 4, followed by a description of the empirical strategy in section 5 and a discussion of the results in section 6. I end with a summary of the key findings.

1.2. Model

The key goal of my model is to generate a relationship between the neighborhood tipping point and two quantities: (a) the utility wedge between an agent's current neighborhood and his/her outside options; and (b) the marginal utility for minority neighbors. The model consists of a demand side, in which households have neighborhood preferences that depend on the endogenous racial mix of the neighborhood and the price of housing services in the neighborhood, as well as other amenities therein. I follow the literature in focusing on the demand side and abstracting from the influence of changes in housing supply on tipping points (Card et. al. 2008, Caetano and Maheshri 2013).

A household's choice-set consists of all of the census tracts in its MSA. Accordingly, its

outside option consists of all of the tracts in its MSA excluding its current tract of residence. In cities where there are many minority tracts, the outside option will impact the ability of non-minorities to exit their neighborhoods of residence. With this definition of households' choice-set, I use a discrete choice model to exploit within-MSA sorting patterns in the data to obtain tipping points for each census tract-year observation (McFadden 1978; Berry, Levinsohn, and Pakes 1995; Bayer et al. 2007). I use data from N census tracts in an MSA from two consecutive census periods to estimate $2N$ tipping points – one for each census-tract-year observation. Using the approach in Card et. al. (2008) and similar data generates a single tipping point – the MSA tipping point.

In the model, I construct the tipping point in two steps. First, I use the estimates of the sorting model to compute an exit function of white households from the neighborhood. For a given exogenous change in the mean utility of whites in neighborhood, τ , the exit function measures the probability that a white household exits its current neighborhood for its best alternative in its choice-set. I refer to τ as the utility tolerance of white households since it parametrizes the exit probability as a function of changes in the mean utility of whites. At the tipping point, the first derivative of the exit function, the exit rate, equals zero, and the tolerance equals τ^* . This definition of tipping is similar to the approach in Card et al. (2008), which associates the tipping point with the share of minorities for which the rate of decline in white population is maximal. I get the tipping point by converting the utility tolerance into a percent minority by using an empirical relationship between the percent minority and the mean utility.

1.2.1. Demand Side

In the model, there are C cities indexed $c \in \{1, 2, \dots, C\}$, and two types of households that are differentiated by a type index, $r \in \{w, m\}$. The type index $r = w$ references white households, while the type index $r = m$ references minority households. Each city is exogenously assigned a total of Q_{tot}^w white households and total of Q_{tot}^m minority households. Each household, in turn, endogenously sorts into one of the N neighborhoods in that city,

indexed by $n \in \{1, 2, \dots, N\}$. The sorting of households to neighborhoods depends on the household income, the price of housing in equilibrium, and the equilibrium level of amenities in each of the N neighborhoods.

A household's problem is to choose the neighborhood that delivers the maximum utility. Solving the household's problem requires first solving for the indirect utility for each of the N possible neighborhoods, and then choosing the neighborhood that delivers the maximum indirect utility. Households h of type r have utility over neighborhood amenities, consumption, and housing services in each neighborhood n . The utility function takes the form:

$$U_{hnr} = \underbrace{\log(A_{hnr})}_{\text{Amenities}} + \underbrace{\alpha \log(C_{hnr})}_{\text{Consumption}} + \underbrace{\beta \log(H_{hnr})}_{\text{Housing}}, \text{ for } r \in \{w, m\}. \quad (1.1)$$

The parameters α and β are the consumption and housing shares. The neighborhood amenity, A_{hnr} , consists of an endogenous component and an exogenous component in addition to an idiosyncratic taste shock:

$$\log(A_{hnr}) = \underbrace{\gamma_r f_n}_{\text{Endog. Amenity}} + \underbrace{\theta X}_{\text{Exog. Amenities}} + \underbrace{\xi_n}_{\text{Unobs. Quality}} + \underbrace{\epsilon_{hnr}}_{\text{Taste Shocks}}. \quad (1.2)$$

The endogenous amenity is the racial composition of neighborhood, $f_n^m = \frac{Q_n^m}{Q_n^m + Q_n^w}$, which is the percent minority in the neighborhood. The value of the endogenous amenity varies by agent type, with whites valuing a 1% increase in the minority share by an amount γ_w , and minorities valuing a 1% increase in the minority share by an amount γ_m . The X 's represent observable characteristics of the neighborhood, which also capture the overall quality of the neighborhood and the ξ_{nr} unobservable measures of neighborhood quality, which may vary by race. I assume that the taste shocks are i.i.d. and follow a type 1 extreme value distribution. This assumption makes it convenient to obtain closed-form solutions without compromising the key insight of the model – which is that the choice-set of white agents impacts the neighborhood tipping points – and allows for the estimation of sub-MSA tipping points.

Solving for the Indirect Utility of a Neighborhood

For each neighborhood n households choose a bundle of consumption C_{hnr} and housing H_{hnr} to maximize utility, subject to the household's budget constraint:

$$C_{hnr} + p_n H_{hnr} \leq I_h. \quad (1.3)$$

Consumption is the numeraire good, and housing price p_n is in terms of units of consumption. The household's income I_h is exogenously determined and independent of the household's choice of a neighborhood n . For neighborhood n , the optimal bundle (C_{hnr}^*, H_{hnr}^*) is:

$$C_{hnr}^* = \left(\frac{\alpha_r}{\alpha_r + \beta_r} \right) I_h \quad (1.4)$$

$$H_{hnr}^* = \left(\frac{\beta_r}{\alpha_r + \beta_r} \right) \left(\frac{I_h}{p_n} \right), \quad (1.5)$$

and the associated indirect utility is:

$$\tilde{V}_{hnr} = \gamma_r \left(\frac{Q_n^m}{Q_n^m + Q_n^w} \right) + (\alpha_r + \beta_r) \log(I_h) - \beta_r \log(p_n) + X\theta + \xi_n + \epsilon_{hnr}. \quad (1.6)$$

To simplify notation, I define V_{hnr} , the deterministic part of the indirect utility, using the following relation:

$$\tilde{V}_{hnr} = V_{hnr} + \epsilon_{hnr}. \quad (1.7)$$

Solving for Neighborhood Demand

After having solved for the indirect utility for each neighborhood, each household of income I_h and type r chooses the neighborhood, n_{hr}^* , that delivers the highest indirect utility:

$$n_{hr}^* = \arg \max \{ V_{hnr} \} \quad (1.8)$$

The household's utility-maximizing behavior across the N neighborhoods in the city generates a conditional demand function, $Q_n^r(\vec{p}|I_h)$, for each neighborhood by both household type and household income category I_h . The conditional demand functions take the form:

$$Q_n^r(\vec{p}|I_h) = Q_{tot}^r \left(\frac{\exp(V_{hnr})}{\sum_{n'=1}^N \exp(V_{hn'r})} \right), \quad \text{for } r \in \{w, m\}, \quad (1.9)$$

where $\vec{p}_n = \{p_1, p_2, \dots, p_N\}$ is the vector of house prices for all neighborhoods in the city. The unconditional demand for neighborhood n by households of type r equals the sum of the conditional demand functions over the income categories:

$$Q_n^r(\vec{p}_n) = \sum_h Q_n^r(\vec{p}|I_h), \quad \text{for } r \in \{w, m\} \quad (1.10)$$

1.3. Tipping Point

In the empirical literature on tipping, the tipping point of a neighborhood n is defined by a threshold minority fraction, f_n^* . When the minority fraction of the neighborhood exceeds this threshold, whites exit the neighborhood at a rapid rate. Below this threshold, changes in the white population of the neighborhood are less stark. In the context of this model, I define the tipping point of a neighborhood as corresponding to the minority fraction for which the exit rate of whites from the neighborhood is maximal.

In order to compute the tipping point, I first construct the exit function for each neighborhood. This exit function traces out the probability of white flight from the neighborhood n as a function of the decrease in utility experienced by whites in the neighborhood due to the arrival of minorities. I adopt a similar approach to Caetano and Maheshri (2013) by using counter-factual decreases in the utility of whites to construct the exit function. The exit function also depends on the utility wedge between the household's inside option, V_{hnw} , and the household's next best alternative, $V_{ha(n)w}$, where the notation $a(n)$ is the neighborhood

that is the household's best alternative should it choose to relocate to another census tract in the same MSA.

After constructing the exit probability as a function of the counter-factual decrease in utility, I will solve for the utility of whites in the neighborhood at the tipping point by solving for the inflection point of the exit function: the level of utility for which the second derivative of the exit function is zero. The first derivative of the exit function is the exit rate. The second derivative, which is required to solve for the inflection point, captures the marginal exit rate. When the marginal exit rate equals zero, the exit rate is maximal.

1.3.1. Conditional Exit Functions

Following the arrival of new minority households to a neighborhood n , some white households may find it preferable to exit the neighborhood and relocate to the best alternative among the other $N-1$ neighborhoods in its city, neighborhood $a(n)$, instead of remaining in neighborhood n . I use τ_{nw} to represent the loss in indirect utility that white households of income category h experience due to the arrival of new minority households to their host neighborhood n . The conditional exit function of whites, which represents the exit probability of whites of a given income category, is given by:

$$E(\tau_{nw}; \vec{V} | I_h) = \sum_{a(n)} \left[\underbrace{\text{Prob}(V_{hnw} - \tau_{nw} + \epsilon_{hnw}^n < V_{ha(n)w} + \epsilon_{ha(n)w})}_{\text{Prob. exit n for a(n)}} \times \underbrace{\omega_{a(n)}}_{\text{Prob. a(n) is best opt.}} \right] \quad (1.11)$$

$$= \sum_{a(n)} \left(\left[\int_{-\infty}^{\infty} F(V_{ha(n)w} - V_{hnw} + \tau_{nw} + \epsilon_{ha(n)w}) f(\epsilon_{ha(n)w}) d\epsilon_{ha(n)w} \right] \omega_{ha(n)w} \right), \quad (1.12)$$

where $\omega_{ha(n)w}$ is the probability that neighborhood $a(n) \in \{1, 2, ..n-1, n+1, ...N\}$ is the best alternative among the $N-1$ options in the household's choice-set, and $F(\cdot)$ is the cumulative distribution function for the taste shocks, which I assume follow a type 1 extreme value

distribution. The probability weight $\omega_{ha(n)w}$ is assumed to be the share of whites in the alternative neighborhood $a(n)$ relative to the total number of whites in the MSA excluding the current tract n :

$$\omega_{ha(n)w} = \frac{\exp(V_{ha(n)w})}{\sum_{a \in \{a(n)\}} \exp(V_{haw})}. \quad (1.13)$$

Each non-minority household living in a given tract will have a single best alternative. However, since I do not observe all of the covariates of an individual non-minority household, I average over all the non-minority households in a neighborhood to obtain a probability than a given tract in the MSA is best alternative for non-minority households in this tract. The probability weights in equation (1.13) are type of counter-factual market shares for utility maximizing non-minority households who face a choice-set of the N-1 census tracts in the MSA, where census tract n has been excluded from consideration. As such these weights present the probability that a tract $a(n)$ is the best option of the N-1 tracts for non-minority households. Since I have assumed that preferences are homogeneous within racial group but heterogeneous across racial group, these probability weights are natural measures of the probability that tract $a(n)$ is the best alternative for a moving agent.⁴

1.3.2. Unconditional Exit Function

The unconditional exit probability of whites from neighborhood n is the sum of the conditional exit functions weighted by the number of households in the income category that corresponds to the individual conditional demand functions, Q_{hn}^w , as a fraction of the total number of whites in neighborhood n , Q_n^w :

$$E(\tau_{nw}; \vec{V}) = \sum_{h=1}^{15} \left(E(\tau_{nw}; \vec{V} | I_h) \cdot \frac{Q_{hn}^w}{Q_n^w} \right) \quad (1.14)$$

The unconditional exit function will be dominated by the behavior of whites in the most highly represented income categories in the neighborhood. This is captured in the weight-

⁴This assumption is particularly reliable for cases where there are a large number of census tracts in the MSA. In these cases, the removal a single census tract has a diminishing effect on the overall sorting in the MSA as the number of tracts in the MSA increases. In the paper, we follow the literature and restrict our analysis to MSAs with at least 100 census tracts (Card et. al. 2008).

ing. As the utility drop becomes large and positive, due to the arrival of minorities, the exit probability goes to one, and all whites exit the neighborhood. In the opposite limit, as τ_{nw} gets arbitrarily large and negative, which corresponds to whites moving into the neighborhood, the probability of white residents exiting the neighborhood converges to zero. In general, the exit function will resemble an S-curve with τ_{nw} on the horizontal axis and the associated conditional exit probability on the vertical axis.

1.3.3. Tipping Point

The tipping point of the neighborhood is the percent minority at the inflection point of the exit function. This is the point at which the exit function changes concavity and the exit rate (the derivative of the exit function with respect to τ_{nw}) is maximal:

$$\frac{d^2 E(\tau_{nw}^*; \vec{V})}{d\tau_{nw}^2} = 0. \quad (1.15)$$

At the tipping point, the mean utility of white households in neighborhood n has decreased by an amount $-\tau_n^*$. I use the the relative marginal utility for minority neighbors to convert this decrement in mean utility into a change in the percent minority. Accordingly, the percent minority at the tipping point is given by:

$$f_n^* = f_n - \frac{\tau_n^*}{\gamma_w - \gamma_m} \quad (1.16)$$

The first piece of the tipping point is the initial percent minority in the census tract, $f_n = \frac{Q_n^m}{Q_n^m + Q_n^w}$. The second part of the tipping point is the change in the percent minority that takes the neighborhood to the critical point of the exit function. The key take-away from equation (1.16) is that the tipping point is directly proportional to the utility tolerance of non-minorities for minorities, τ , and inversely proportional to the relative preference of non-minority households for minority neighbors, $\gamma_w - \gamma_m$, which in the data is negative. If $\tau_n^* > 0$, then the neighborhood n is a more desirable neighborhood than the alternatives in

the choice-set. In order for this neighborhood to tip, minorities must move in to lower the utility of non-minorities to the point where the neighborhood tips. If $\tau < 0$, the opposite is true, and the tipping point is lower than the current fraction of minorities. One merit of estimating tipping points in this manner is that it allows researchers to estimate the tipping point of census tracts that have tipped, that have yet to tip ($\tau > 0$), and that are beyond their tipping points ($\tau < 0$).

Estimating Preferences

This preference parameter, $\gamma_w - \gamma_m$, is identified from the differential sorting of non-minorities into neighborhoods as a function of the fraction of minorities in the neighborhood. From equation (1.6), we relate the ratio of the non-minority to minority market share of a neighborhood to the percent minority in the neighborhood and the relative preference parameter $\gamma_w - \gamma_m$:

$$\log\left(\frac{Q_n^w}{Q_{tot,c}^w}\right) - \log\left(\frac{Q_n^m}{Q_{tot,c}^m}\right) = (\gamma_w - \gamma_m) \left(\frac{Q_n^m}{Q_n^m + Q_n^w}\right) + e_{n,m} \quad (1.17)$$

The term on the left-hand side is the relative market share of whites to minorities in neighborhood n . The market share of a neighborhood is the fraction of households in the MSA of a given type that reside in the neighborhood. Moreover, the log of the market share is mean utility of an household of the given racial type. The regressor on the right-hand side is the percent minority in the census tract. By taking the relative market share, I can difference out characteristics of the neighborhood that are valued equally by minorities and non-minorities. Here I make the assumption that white and minority households value everything similarly except the percent minority in the tract.

1.3.4. Semi-Parametric Estimate of Tipping Point

I also use the relative market shares to develop a semi-parametric estimator of the tipping point, which is non-linear parallel to equation (1.16). In equation (1.17), the relative market

shares are a linear function of the fraction of minorities, f_n . One limitation of this specification is that it can produce tipping points that lie outside of the interval $[0, 1]$. I relax this assumption by allowing the percent minority in a neighborhood to depend on the relative market shares in a non-linear fashion. I obtain this relationship, empirically, by regressing the percent minority in the census tract on powers of the log of the relative market shares:

$$\frac{Q_n^m}{Q_n^m + Q_n^w} = \alpha_0 + \sum_{j=1}^5 \alpha_j \left[\log \left(\frac{Q_n^w}{Q_{tot,c}^w} \right) - \log \left(\frac{Q_n^m}{Q_{tot,c}^m} \right) \right]^j \quad (1.18)$$

The coefficients of this regression define the inverse mapping from the ratio of the non-minority to minority market shares to the percent minority in the tract. This inverse mapping is important because we know the mean utility of white households at the tipping point, but the ultimate quantity of interest is the percent minority, which defines the tipping point; therefore we need the inverse mapping of the relative utility to percent minority. I use the estimated parameters to obtain the percent minority at the tipping point by inserting a value of $V_{nw} - \tau_{nw}^* - V_{nm} = \log \left(\frac{Q_n^w}{Q_{tot,c}^w} \right) - \log \left(\frac{Q_n^m}{Q_{tot,c}^m} \right) - \tau_{nw}^*$ for the relative market share at the tipping point.

1.4. Data

To estimate the model, I use data from the U.S. census covering five decades: 1970, 1980, 1990, 2000, and 2010. This data consists of the demographic characteristics of the households living in each of the census tracts, as well tract-level measures of the local housing stock and local economic conditions. The key variables of interest for this study are the population shares of each census tract broken down by race. Prior work has used similar data from the 1970–2000 extracts of the census data to compute MSA tipping points (Card et al. 2008). I build on this work by updating the previous results to include estimates of tipping points from the 2000 and 2010 censuses. With these five decades of tipping points, I trace the time evolution of tipping points to show that, over time, tipping points in the United States have increased. In addition to this empirical contribution, my paper also

makes a methodological contribution. Whereas these data have been used in Card et al. (2008) to compute MSA tipping points, I use these data to compute census tract tipping points. These tract-level estimates capture the distribution of neighborhood tipping points within an MSA.

I follow Card et al. (2008) in making the following cuts in the data. First, I eliminate any tracts whose population growth between consecutive census years surpasses average population growth in the MSA by more than five standard deviations. Second, I drop all tracts that experience an increase of more than 500% in their white population between consecutive census years. These first two cuts reduce the effect of outliers on the results of this study. For the final cut, I focus my analysis on MSAs that have 100 census tracts or more, also following Card et al. (2008). There are 123 MSAs that satisfy these criteria, and these MSAs cover the 38,489 census tracts that comprise my final data set.

1.5. Results

1.5.1. Descriptive Statistics: Racial Preferences

In Table 1, I report the mean, median, and standard deviation of these relative preference estimates from the “diff-in-diff” procedure of equation (1.17). To compute standard errors on the point estimate and preference parameters, I use an $N=1000$ bootstrap. The estimates for $\gamma_w - \gamma_m$ range from -8.84 in 1970 to -6.11 in 2010. For 1970, the diff-in-diff point estimate of -8.84 means that a 7.8 percentage point increase in the fraction of minorities was associated with a 50% reduction in the non-minority population of the average neighborhood. The diff-in-diff point estimate of -6.11 for 2010 indicates that an 11.3 percentage point increase in the fraction of minorities was needed for the non-minority population in a neighborhood to halve.

In Figure 2, I graph decadal changes in the distribution of the relative racial preferences. Each of the kernel density plots uses data from the 5th to 95th percentile to limit the effect of outliers on the shape of the graphs. In each ten-year period, the distribution

of preferences shifts to the right, indicating that the mean is decreasing over time. A decreasing mean over time is consistent with white households becoming more tolerant of living with minorities. Over time, the distribution of preferences also narrows. This suggests that, on average, white households' increase in tolerance is occurring across all levels of the preference distribution. The compression in the distribution of racial preferences across cities, over time, is responsible for the declining importance of racial preferences as an explanatory factor in cross-sectional differences in tipping points across cities.

1.5.2. Descriptive Statistics: Tipping Points

By applying the model to the data, I obtain two sets of census tract tipping points. The first set comes from the the linear model in equation (1.16), and the second set comes from the semi-parametric estimator of equation (1.18). Since the predicted outcome of both approaches is a tipping point that lies in the interval $[0,1]$, an apt analogy for describing the two methods is that the linear (diff-in-diff method) is analogous to a linear probability model, while the semi-parametric method is analogous to a non-linear, e.g. probit model, which produces estimates that lie in the interval $[0,1]$.

Tract Tipping Points

In Table 2, I report summary statistics for the census tract tipping points from the diff-in-diff method of equation (1.16). The table is divided into three panels. In the first panel I report the mean, median, and standard deviation of the census tract tipping points for the full sample in each of the census years. In the second panel I report the identical statistics for the census tract tipping points that are in the allowable $[0,1]$ range for the given census year. In the third panel, I report the identical statistics for each census tract that is in range in 2010. This restriction gives a consistent set of tracts across all census years. Because some of the tipping points are not "in range," I use the results from these three panels as checks that the time trends in the tipping points that I observe are consistent under the three sample restrictions: the full sample, the sample of tracts that are in range in the

census year, and a consistent set of tracts that are required to be in range in the 2010 census year. In Table 3, I present results from the semi-parametric (inverse) method of equation (1.18).

The mean and median of the census tract tipping points increase monotonically over time for both the diff-in-diff and the semi-parametric (inverse) mapping methods. Moreover, the mean tipping point is greater than the median tipping point for all years. The magnitudes of the estimates are also comparable in both methods. I focus my discussion on the estimates from the inverse mapping method, because between 98% and 99.6% of the tipping points for this method are in the allowable $[0,1]$ range.⁵ For this method the tipping points have a mean of 15% in 1970, 22% in 1980, 28% in 1990, 36% in 2000, and 41% in 2010. The median-tract tipping point also monotonically increased from 1970 to 2010. In 1970, the median-tract tipping point was 13% and by 2010 it was 34%. The inter-censal correlation between the tract tipping points is between 0.71 and 0.78, as reported in Table 4. This demonstrates that while the mean tipping points of the tracts has increased over time, there has been strong persistence in the ranking of tracts across time.

In Figure 3, I report kernel density plots of the tract tipping points in each census year. The distribution for each year is a single peaked distribution that is left skewed. Over time the peak pushes out the right, and the distribution flattens and gains more mass in the right tail. With each succeeding census year the curve shifts out by less, indicating that the tipping point is increasing at a decreasing rate.

MSA Tipping Points

Using the census tract tipping points, I construct two measures of MSA tipping points. The first is a mean MSA tipping point and the second is a median MSA tipping point. The mean MSA tipping point is the average tipping point of the census tracts in the MSA. The median MSA tipping point is the median census tract tipping point in the MSA. In Table 5,

⁵By comparison, only 57%–84% of the diff-in-diff linear estimates are in range. Nevertheless, the results in the second panel of Table 2 agree with the results in the inverse mapping method for all years.

I report the mean, median, 25th percentile, 50th percentile, and 75th percentile of the mean MSA tipping points for all MSAs and also broken down by geographic region – Northeast, Midwest, South, and West. The results in Table 6 are for the median MSA tipping points.

From 1970 to 2010, both MSA tipping points increased monotonically over time. In 1970 the mean MSA tipping point was 11%; by 2010 it rose to 33%. This increase in the MSA tipping points over time was undergirded by an increasing time trend in tipping points in all regions of the US. MSA tipping points in the West increased fastest, at a rate of 7.5 percentage points per decade, while MSA tipping points increased slowest in the Midwest - 3.25 percentage points per decade. The distribution of tipping points in the Northeast parallels the distribution of MSA tipping points in the Midwest. The mean tipping points in both regions over time were (9%/9%) (12%/12%), (15%/14%), (22%/19%), (26%/22%) in 1970, 1980, 1990, 2000, and 2010, respectively. Likewise, the MSA tipping points in the South mirrored those in the West.

The MSA tipping points exhibited a high degree of correlation across consecutive census periods (0.89–0.98), notwithstanding the fact that they increased substantially across time (Table 7). One striking fact about the MSA tipping points is that the mean MSA tipping points and the median MSA tipping points had very similar distributions. For example, in the full sample the mean of the mean MSA tipping points and the mean of the median MSA tipping points were (13%/11%), (18%/16%), (22%/21%), (30%/28%), and (35%/33%) in 1970, 1980, 1990, 2000, and 2010, respectively. For each year the difference between the mean and the median MSA tipping points was between 1 and 2 percentage points.⁶ Moreover, the mean of the mean MSA tipping points is similar to the median of the census tract tipping points. I exploit this fact in the next section, where I show that the tipping points estimated by Card et al. (2008) are similar to the local mean and the local median of the tipping points of the marginal census tracts in the MSA.

⁶By comparison, the mean and median of the census tract tipping points were more dissimilar (15%/13%), (22%/16%), (28%/21%), (36%/28%), and (41%/34%) in the respective census years. For each year the difference between the mean and the median tract tipping points is between 2 and 8 percentage points.

1.5.3. Comparison with Prior Estimates of Tipping Points

The Card et al. 2008 (CMR) tipping points cover three decades – 1970, 1980, and 1990. On average, the tipping points that I get are 3 percentage points higher in 1970, 9 percentage points higher in 1980, and 12 percentage points higher in 1990 than CMR. This difference occurs because the two tipping points capture different aspects of the underlying distribution of census tract tipping points. The MSA tipping points that I report are an average of the underlying census tract tipping points, which I am able to estimate because I model the location decision of households within the MSA and use the counter-factual exercise to obtain census tract tipping points. The CMR approach accurately measures a local average of the marginal census tracts (i.e., tracts that are close to the tipping threshold). This distinction between the two MSA tipping points is evident from the theories guiding their construction.

The CMR tipping points are the result of a fixed-point procedure. To obtain the MSA tipping point, the CMR fits a polynomial of the change in the percent of whites between census years (above the MSA average) as a function of the fraction of minorities in the base census year. Each observation used to fit this function is a census tract in the MSA which appears into consecutive periods. The first method for determining the tipping point is to solve for the zeros of this polynomial. The key point is that tracts below the tipping point experience above-average growth in their white populations, whereas tracts beyond the tipping point experience below average growth in their non-minority populations. The minority fraction at the zero of this polynomial is taken to be the MSA tipping point.

In cases where there are multiple zeros, the authors took the zero that delivered the most negative first derivative. This equilibrium selection procedure parallels the approach that I take in this paper, where for each tract I stipulate that the tipping point occurs at the level of utility for which the exit rate of non-minorities is maximal (and the marginal exit rate equals zero). Since the CMR tipping point is the zero of a fitted polynomial, it depends crucially on the behavior of the census tracts in the vicinity of this zero. I call these

census tracts the “marginal census tracts.” These are tracts that are close to their tipping points. I call tracts that are farther away from the zero of the polynomial “infra-marginal census tracts.” Changes in the behavior the infra-marginal tracts have less bearing on the estimated value of the CMR MSA tipping points.

The setup of the CMR procedure suggests that the CMR tipping points are local averages of the tipping points of the marginal tracts, or perhaps the median of the tipping points of the marginal census tracts in the MSA. Since I have estimates of tipping points for each census tract in an MSA, I can test the hypothesis that the CMR tipping points are local averages of the tipping points of marginal census tracts or the median of the tipping points of the marginal tracts. In Table 8, I present results from a regression of the difference between the CMR and Blair MSA tipping points, which I call the *tipping difference*,⁷ and the fraction of marginal census tracts in the MSA. Here, a marginal census tract is a tract whose percent minority is within 5 percentage points of its estimated tract tipping point. For example, if a tract has a tipping point of 35% and a current percent minority of 32% it is considered a marginal tract. Likewise, if another tract in the same MSA has a tipping point of 7% and a current percent minority of 11%, I also consider it a marginal tract. To allow for asymmetry in the impact of marginal tracts that lie to the left and to the right of their respective tipping points, I include separate explanatory variables for (a) the fraction of tracts in the MSA that are marginal and have minority fractions below their tipping points; and (b) the fraction of tracts in the MSA that are marginal and have minority fractions greater than their tipping points.

The constant terms from the regressions in Table 8 capture the mean *tipping difference*. In 1990, the tipping difference was -15%. It was -10% in 1980 and -8% in 1970. These regression results accord with the -13%, -9%, and -3% tipping differences from the raw data. Based on the regression results, the marginal tracts that were below the tipping point

⁷I do not call this quantity a bias because my hypothesis is that the CMR tipping points measure the distribution of the marginal census tracts, which in and of itself is an important quantity. We care about which tracts are marginal and how the distribution of marginal tracts varies across time and across space.

drove the tipping difference in 1990 and 1980. In 1970, the marginal tracts that were above the tipping point (i.e., those that had already tipped), drove the tipping difference. One reason there is a tipping difference at all is that there were few marginal tracts, and so an average of the marginal tracts was different from an average of all tracts. To illustrate this point, I combine the constant term from the regression and the significant coefficient on the marginal tracts to compute the threshold fraction of marginal tracts, \tilde{f} required for there to be no tipping difference in each year:

$$\tilde{f} = \frac{\# \text{ marginal tracts in MSA}}{\text{Total } \# \text{ tracts in MSA}}. \quad (1.19)$$

To solve for \tilde{f} , I set the sum the constant term from the regression and the product of the (significant) coefficient on the marginal tracts times \tilde{f} equal to zero. At this value of \tilde{f} , the tipping difference is zero, and the CMR tipping points and the tipping points that I obtain are equal. In 1990, 72% of the tracts would have to be marginal for there to be no tipping difference. In 1980, 53% of the tracts would have to be marginal; and in 1970, 13% of tracts would have to be marginal. In the reality, the average MSA consisted of 14% of marginal tracts (below) in 1990, 13% of marginal tracts (below) in 1980, and 11% of marginal tracts (above) in 1970. Since the mean number of marginal tracts was closest to the required level in 1970, it is not surprising that the tipping difference was smallest in 1970. The opposite is true for 1990, the year when the difference between the required threshold and the fraction of tracts that were marginal was largest.

The results of Figure 5 and Figure 6 further illustrate the point that the CMR tipping points are a local mean or median of the tipping points of the marginal census tracts. In both Figure 5 and Figure 6, there are three panels, one each for 1970, 1980, and 1990. The dashed line in each panel of Figure 5 is a kernel density plot of the tipping difference using all tracts. The solid line in each plot is a kernel density plot of the *tipping difference* using only the marginal tracts that were predicted to explain the tipping difference. For example,

in 1980, the tipping difference using only marginal tracts is the difference between the CMR tipping point (unmodified) and the mean of the census tracts, which are to the left of the tipping points, with the means of marginal tipping points computed by MSA. For each year, the dashed line peaks to the left of zero. This reflects the fact that the CMR tipping points are, on average, smaller than the tipping points that I get when I compute means using all census tracts in an MSA. On the other hand, in every year, the solid line peaks near zero. This reflects the fact that the tipping difference disappears when I compute MSA tipping points using only the marginal tracts. With this sample restriction, the tipping difference of the median MSA is reduced from -11.2% to -2.4% in 1990, from -7.8% to -1.1% in 1980, and from -3.7% to 1.9% in 1970. This suggests that the CMR tipping points are a local average of the tipping points of the marginal census tracts, as hypothesized.

To verify this, I perform a similar exercise, this time changing the definition of the tipping difference to be the difference between the CMR tipping point and the tipping point of the median census tract in each MSA. When I restrict the sample to only the marginal tracts, the tipping difference equals the difference between the CMR tipping point and the median tipping point of the marginal census tracts. Apart from this change in the definition of the tipping difference, Figure 6 is laid out identically to Figure 5. The dashed lines peak to the left of zero, reflecting the fact that the CMR tipping points are smaller, on average, than the MSA tipping points that I compute from the median. The solid lines, however, peak even more sharply around zero than the solid lines using the mean. This reflects the fact that the tipping difference also disappears when I compute MSA tipping points using only the marginal tracts. With this sample restriction, the tipping difference of the median MSA is reduced from -11.2% to -1.7% in 1990, from -7.8% to 0.1% in 1980, and does not substantively change in magnitude in 1970 (from -3.7% to 4%). Taking the best of the local mean and median results, the tipping difference of the median MSA is bounded above by 1.9% and bounded below by -0.1%. These results provide further support for the hypothesis that the CMR tipping points capture the shape of the distribution of marginal census tracts in an MSA.

From this comparison of marginal and infra-marginal tracts, we learn that the tipping points of the marginal census tracts evolve more slowly over time than those of the infra-marginal tracts. The CMR tipping points, which were shown to be an average of the marginal tipping points, increased an average of 1 percentage point per decade (1970–1990). We also learn that the infra-marginal tracts play an important role in the secular time trend of tipping points. The MSA tipping points using all tracts (both marginal and infra-marginal) increased at an average rate of 5.5 percentage points per decade (1970–2010), which is substantially higher than the growth rate of the tipping points of infra-marginal tracts. An important contribution of the method in this paper is that it enables researchers to compute the tipping points of all census tracts and derived MSA tipping points, which are aggregates of the underlying census tract tipping points. The dynamics of these MSA tipping points better reflect the dynamics of the underlying census tracts.⁸

1.5.4. Results: Preferences versus Outside Options

The motivating insight of this paper is that outside options affect the ability of households to act on their preferences for neighborhood racial composition. A household may remain in a neighborhood despite its racial composition if the outside options do not offer higher utility. The reverse is also true – a household may exit its current neighborhood because of the availability of desirable alternative neighborhoods in its city. For each census year, I decompose the mean MSA tipping point into a component due to the mean preferences of the households and a measure of their outside options, which depends on the extent of clustering in the city by race.

To measure minority clustering, I use a Herfindahl-Hirschman Index (HHI) that is standardized to have a mean of zero and standard deviation of one. To compute the index, I construct the minority market share of each census tract n in a given census year y by dividing the number of minorities in the tract by the total number of minorities in its MSA

⁸The mean tipping point of census tracts in the data increased by 6.75 percentage points per decade (1970–2010)

c : $s_{n,m,y}^c = \frac{Q_{n,m,y}^c}{Q_{tot,m,y}^c}$. I then square the minority market shares, $s_{n,m,y}$ and sum over them for each MSA, c , to get the MSA HHI^c : $HHI_c = \sum_n (s_{n,m,y}^c)^2$. (1.20) To finish, I de-mean the Minority HHI and normalize it to have variance 1 in each year. This yields a minority HHI z-score, $(HHI_{m,z,y}^c)$ for each MSA for each census year from 1970 to 2010. I construct a standardized non-minority HHI using an identical procedure $(HHI_{w,z,y}^c)$. To construct the standardized measure of racial preferences, $(\Delta\gamma_{z,y}^c)$, I de-mean the relative marginal utility for minority neighbors from the diff-in-diff procedure of equation (1.17) and normalize it to have standard deviation 1.

A one standard deviation *decrease* in the minority HHI corresponds to less clustering of minorities in the MSA, or more census tracts in the MSA having some minority households. Less clustering of minorities creates a choice-set in which many census tracts are sprinkled with at least some minorities, making it difficult for non-minority households to sort into all-white neighborhoods. A one standard deviation *increase* in the non-minority HHI corresponds to more clustering of non-minorities into fewer census tracts within an MSA. Greater clustering of non-minorities creates a choice-set that is bimodal, with some tracts having many non-minorities and others having few non-minorities. A one standard deviation *increase* in the race preferences is associated with non-minorities having preferences for minority neighbors that is more similar to the preferences that minorities have for minority neighbors.

In Table 9, I report the results of a regression of mean MSA tipping points (T_y^c) on the minority HHI ($HHI_{m,z,y}^c$), the non-minority HHI ($HHI_{w,z,y}^c$), and the standardized race preferences ($\Delta\gamma_{z,y}^c$):

$$T_y^c = \eta_1^y HHI_{m,z,y}^c + \eta_2^y HHI_{w,z,y}^c + \phi^y \Delta\gamma_{z,y}^c. \quad (1.21)$$

Since the tipping points were constructed using the choice-set faced by households and the estimated preferences, I read this regression as providing a decomposition of the MSA

tipping points into a component due to the configuration of the outside option and the mean preferences in the MSA. The results of this regression are summarized in Figure 7. The first result is that less minority clustering, or having the presence of a more diverse choice-set, is associated with a higher tipping point. In 1970, for example, a one standard deviation decrease in the minority HHI results in a $\eta_1^{1970} = 1.1$ percentage point increase in the tipping point. From 1970 to 2010, this effect of a reduction in minority clustering on mean MSA tipping points strengthens monotonically. By 2010, a one standard deviation decrease in minority clustering increases the mean tipping point by 14 percentage points, which is roughly 50% of the base tipping point of 30% in 2010. Since reductions in the clustering of minorities result in many tracts having at least some minorities, this reduces the number of all-white tracts, which in turn creates a barrier to neighborhood exit by non-minorities.

Increasing the clustering of non-minorities has no statistically significant effect on MSA tipping points in 1970. In 1980, however, a one standard deviation increase in the non-minority HHI increases the tipping point by 2.6 percentage points or 15%. By 2010, a one standard deviation increase in the non-minority HHI is associated with a 11.2 percentage point increase in the tipping point. When the non-minority HHI increases, non-minorities are clustered in fewer census tracts. This creates a bimodal distribution of tracts by racial composition. There are fewer tracts with some non-minorities because of the increased clustering of minorities; and there are more tracts with higher minority fractions also because of the clustering. Both of these factors act as barriers to the exit of non-minorities from neighborhoods, resulting in an increasing tipping point. Hence both the diffusion of minorities through the MSA by a decrease in the minority HHI and an increase in the clustering of non-minorities by an increase in the non-minority HHI are associated with greater tipping points in the cross-section. This effect also strengthens over time, as illustrated by the non-zero slope of the HHI coefficients over time in Figure 7.

While the role of outside options becomes increasingly important over time, the role of racial

preferences diminishes. This is illustrated by the negative downward slope of the racial preferences coefficients. In 1970, a one standard deviation increase in racial preferences was associated with a $\phi^{1970} = 6.6$ percentage point increase in the tipping point, which equals a more than 50% increase in the tipping point. In 2010, a one standard deviation increase in racial preferences had no effect on the tipping point. The decline in the effect of racial preferences on the tipping point is nearly monotonic over time. Interestingly, 1970 and 2010 are opposite sides of the coin when it comes to the respective roles of racial preferences and outside options in MSA tipping points. In 1970, the clustering of minorities and non-minorities had a small effect on tipping points, whereas racial preferences were at the zenith of their importance. In 2010, the opposite was true – the configuration of the outside options was paramount, and racial preferences appear to have played no role in explaining cross-sectional differences in MSA tipping points. With the configuration of a household’s choice-set mattering more now than in the past, modeling the tipping with outside options, as I do here, is of principal importance in understanding the future evolution of city tipping points.

1.6. Conclusion

A neighborhood tips when non-minorities exit in response to integration. Prior literature has focused on racial preferences as a key driver for tipping. I show that in addition, the outside options of households also matter. To incorporate outside options into a model of tipping, I start with the assumption that a household’s outside options are the other neighborhoods in its city of residence. I further require that a household’s response to integration is incentive compatible – the household only exits its current neighborhood if relocating delivers higher utility than staying. I pair this assumption about the choice-set and the incentive compatibility constraint with a discrete choice model in order to exploit the sub-MSA sorting patterns in the data to estimate census tract tipping points.

The census tract tipping points that I estimate reveal two key findings. First, tipping points we learn that tipping points have increased over time by more than previously thought. This

result also holds when I aggregate the tract tipping points at the MSA level. Prior estimates in the literature pegged the growth in city tipping points at an average of 1 percentage point per decade, whereas I find that city tipping points have grown by an average rate of 5 percentage points per decade. I show that my estimates of city tipping points are different from the prior literature because the city tipping points that I estimate are an average of the underlying distribution of tract tipping points. The CMR tipping points appear to be local averages of the tipping points of census tracts that are close to tipping.

The second key finding of this paper is that outside options are increasingly important for explaining cross-sectional variation in tipping points. In 1970, a one standard deviation increase in the diffusion of minorities across a city was associated with a 1 percent increase in the tipping point. By 2010, a similar change in the outside option was associated with a 14 percentage point increase in the tipping point. By contrast, differences in racial preferences have become less important for explaining cross-sectional differences in city tipping points.

From a policy standpoint, the distinction between preferences and outside options is important. If heterogeneity in tipping points were driven primarily by racial preferences, then the government would require policy levers that change preferences in order to mitigate neighborhood tipping. If, on the other hand, heterogeneity in tipping points were due to differences in the outside options of households, then this would provide more scope for place-based policies to promote integration. The results of this study suggest that the latter is the case, since the effect of outside options has become larger over time, while the effect of racial preferences on tipping points has diminished in significance.

CHAPTER 2 : Why Don't Elite Colleges Expand Supply?

Abstract: While both the skills premium and the demand for elite colleges have increased over time, elite colleges have not increased supply to capture these profits, instead choosing to reduce their acceptance rates. We propose a model in which colleges compete on relative selectivity as an explanation for this puzzle. There are two equilibria of this game: one in which elite colleges coordinate on a low admissions rate and another in which elite colleges coordinate on a high, Pareto dominant, admissions rate. In the low equilibrium, increased demand reduces the admissions rate. The opposite is true in the high equilibrium. We take the model to data using recent admissions statistics for Harvard, Yale and Princeton (2003-2015). We find that the colleges play the low equilibrium, which results in an average annual loss of \$73M in social surplus for each school. For comparison, the loss in social surplus resulting from the low equilibrium play is 50% larger than the size of the dead weight loss (\$47M) in the standard monopoly case where Harvard, Yale and Princeton (HYP) have no preferences over relative prestige. From a policy standpoint, allowing schools to collude on their admissions rates could be Pareto improving if this collusion results in the adoption of the high equilibrium admissions rate.

2.1. Introduction

In 1979, the incoming class of Yale College freshmen stood at 1,346 students (Walters, 2001). In 2014, the size of the incoming class of Yale College freshmen stood at 1,360 students.¹ Over the same period, the number of applications to Yale College more than tripled from 9,331 students in 1979 to 30,932 in 2014. Across the Ivy League and its peer institutions, the story is the same story— an increasing demand for spaces by increasingly more qualified students ² with little change in the supply of spots. In Figures 8 and 9, we provide plots of the number of applicants and the number of admitted students at each

¹The peak enrollment at Yale College from 1979 to 2014 was 1409 students in 2000 and the minimum enrollment was 1255 students in 1986.

²At Yale College, the inter-quartile range of the SAT Math of admitted students increased from 620-730 (1975) to 710-790 (2014) and likewise for Sat Verbal from 670-780 (1975) to 710-800 (2014).

undergraduate college of the eight Ivy League universities during the past 13 years. They all demonstrate a pattern of rapidly increasing number of applications and stable number of students admitted. Given the increasing returns to education for students, and particularly to education at elite colleges, and the increasing endowments due to alumni contributions, it is a puzzle that elite colleges have not expanded supply in the past four decades (Blackburn and Neumark 1993, Hoxby 2000, Katz and Murphy 1992, Baum et. al 2010, Long 2007).

One potential explanation for this puzzle is that elite universities face a hard supply constraint. The existence of a hard supply constraint may explain why elite universities in city centers, for example Columbia University and the University of Pennsylvania, have not appreciably increased the size of their undergraduate population. This explanation, however, is less convincing for schools like Stanford University³ and Duke University, where available land is plentiful.

Another alternative explanation to this puzzle is that elite schools have not expanded supply in order to maintain the quality of the student body. This explanation is also unlikely given the increased quality of the applicant pool at elite schools documented in Hoxby (2009). For example, from 1970-2000, the number of applicants to Stanford with Math scores > 700 increased by a factor of 2.3 (Hennessey 2007).

The solution to this puzzle, which we offer in this paper, is that elite schools have not expanded supply because of competitive pressure. In our model, when deciding whether to expand the supply of spaces in its entering class, colleges face the following important trade-off: by adding more students, there is the possibility that this increases profits. On balance, adding more students could also lead to the college becoming less selective by increasing its admissions rate relative to that of its peer institutions. We show that when elite colleges compete over relative prestige, which we define as the ratio of the college's admissions rate to the average admissions rate of schools in its peer group, this leads to

³In 1970, Stanford received 9,800 applications and had a class size of 1,555 students. In 2014, Stanford received 42,167 applications and had a class size of 1,691.

a “prestige externality” that produces multiple equilibria. In our simple case, we take the $N \geq 2$ competing colleges to be identical in their preferences over relative prestige.⁴ We find that this prestige game yields two competitive equilibria – one in which the admissions rate of colleges is low (\bar{a}_L) and the other in which admissions rates are high (\bar{a}_H).⁵

Since all colleges are identical in the model, in the low admissions rate equilibrium, all colleges coordinate on an identical admissions rate, resulting in zero net utility from relative prestige at the equilibrium admissions rate. The same is true of each college’s actions in the high admissions rate equilibrium. The Pareto ranking of the two equilibria is therefore driven by the differences in net profits between the high and low admissions rate equilibrium. The high admissions rate equilibrium yields a larger producer surplus for colleges since the high admissions rate equilibrium takes the colleges closer to the monopoly admissions rate in which the college places zero weight on relative prestige. Furthermore, students are also better off under the high admissions rate equilibrium since more students are admitted; moreover, conditional on being admitted, students pay a lower price because each college faces a downward sloping demand curve.

In the low admissions equilibrium, the model predicts that increases in demand results in *decreased* admissions rates, whereas the opposite is true in the high admissions equilibrium. We use this comparative static to take the model to data using Harvard, Yale and Princeton (HYP) as a group of nearly identical colleges. The main result is that HYP schools play the Pareto-dominated low admissions rate equilibrium in which both consumers and producers are worse-off relative to the high admission rate equilibria. Moreover, in the low equilibrium, the social surplus is shared in a ratio that differentially favors producers (colleges) relative to consumers (students) at a ratio of 14.4 to 1, rather than 3.2 to 1 for the high equilibrium.

To put a very *conservative* lower bound on the loss in social surplus, we use the sticker price of tuition as the maximum willingness to pay of students. At the HYP schools, between

⁴We extend our key results to the heterogeneous case in the Appendix.

⁵We show that an equilibrium exists for a given range of parameter values.

30% to 40% of students receive no financial aid, which makes using the sticker price of the college a reasonable lower bound on the maximum willingness-to-pay. Using this value of the maximum WTP, we find that the loss in social surplus from playing the low equilibrium is bounded below by \$73M annually for each of the three universities. The individual schools shoulder \$49M of the loss in social surplus, with students shouldering the remaining \$23M in lost social surplus. To put this loss into context, if HYP placed zero weight on the relative prestige, but acted as a profit maximizing monopolist, the dead weight loss would be \$47M (per school). Consequently, the loss in social surplus due to the low equilibrium play is 55% greater than the deadweight loss.⁶ One important policy implication of our model is that allowing schools to collude on their admissions rates would be Pareto improving.

The analysis in this paper is complementary to the work by Hoxby (2009), who provides a demand-side explanation for the increasing selectivity of elite colleges. Hoxby shows student preferences for colleges have changed over time from preferring local colleges to preferring non-regional schools. This increased student mobility has led to students applying more to elite institutions. This explains why application numbers have increased. Our paper posits an explanation for why supply has not correspondingly increased. In many models in the literature, it is common to assume that the number of spaces at a college are fixed (Chade et. al 2014, Avery and Levin 2010, and Fu 2014). In our model, we do not restrict the supply of spaces, but instead assume a constant marginal cost for each additional student. By endogenizing equilibrium supply, we can test whether preferences over relative prestige can explain the puzzle of stable supply in the face of increasing demand.

2.2. Model

The model consist of a two stage game. In the first stage of the game, students decide whether to apply to a given college and in the second stage of the game colleges decide how

⁶The loss in consumer is particularly important from a policy standpoint because minority students differentially benefit more from attending elite schools (Dale and Krueger, 2002, 2011) yet they are under-represented at many elite schools. Reardon et. al show that whites in the top half of the income distribution are 3 times more likely to attend a selective school than African-Americans of the same socio-economic status.

many students to admit given the college's objective function.

2.2.1. Student Demand for College Admission

To gain traction on modeling the number of students who apply to a given college, we assume that demand for each college is linear, such that a given school i faces an inverse demand function of the form:

$$P_i = P_{0,i} - b_i Q_i, \quad (2.1)$$

where $P_{0,i}$ is the willingness to pay of the student with the highest valuation for college i and b_i is the inverse of the price sensitivity of demand. To determine the number of students applying to a school, we assume that students only apply to a college if the student's willingness to pay is greater than the cost of the application fee, $P_{app,i}$. Since in practice most colleges offer fee waivers for students with financial need, the marginal student is a student whose willingness to pay equals zero. Hence the number of students applying to a school is given by:

$$0 = P_{0,i} - b_i Q_i^* \implies Q_i^* = \frac{P_{0,i}}{b_i}. \quad (2.2)$$

2.2.2. Supply Decision of Colleges

There are $N \geq 2$ distinct colleges. Each school, indexed by i , has utility over its relative prestige and financial earnings. The relative prestige of the college takes the form of the ratio of the college's admissions rate a_i relative to the average admissions rate of the other colleges in its peer group $\bar{a}_{-i} = \frac{1}{N-1} \sum_{k \neq i} a_k$. We define the admissions rate as the ratio of the number of the students who are accepted to college i to the number of student who have applied: $a_i = \frac{q_i}{Q_i}$. The college's financial earnings take the form of total revenue minus total cost, where we assume that the inverse demand for slots is linear and the marginal cost of producing an additional slot at the college is a constant, c_i . The college's utility

function takes the form:

$$V_i = \underbrace{r_i \left[1 - \left(\frac{a_i}{\bar{a}_{-i}} \right) \right]}_{\text{Monetary Value of Relative Prestige}} + \underbrace{(P_{0,i} - b_i q_i) q_i - c_i q_i}_{\text{Profits}}. \quad (2.3)$$

The first term of V_i is the monetary value of relative prestige. If the college has an admissions rate that equals the mean of its peer group then its relative prestige is zero in monetary terms. If, on the other hand, the college has an admissions rate that is twice that of the mean of its peers, then it incurs a prestige penalty equal to $-r_i$. The second term of V_i is the financial profits that the college can earn due to its price setting ability. The college, therefore, faces a trade-off in determining how many students to admit: admitting more students can potentially lead to greater profits, but it potentially lowers the college's relative prestige vis-à-vis that of its peers.

2.2.3. Solving the Model

First Order Conditions

The college chooses q_i to maximize utility. This yields the following first order condition:

$$0 = -\frac{r_i}{\bar{a}_{-i}} + (P_{0,i} - c)Q_i - 2a_i b_i Q_i^2. \quad (2.4)$$

Best Response Functions

We rewrite the first order condition to obtain the best response function, $a_i(\bar{a}_{-i})$, in terms of two economically meaningful quantities (i) the no-prestige monopoly admissions rate:

$$\theta_i = \frac{P_{0,i} - c_i}{2bQ_i}, \quad (2.5)$$

which is the utility maximizing admissions rate for the case where $r_i = 0$ and college i has no utility over relative prestige; and (ii) the prestige ratio:

$$\lambda = \frac{r_i}{\frac{(P_{0,i}-c_i)^2}{2b}}, \quad (2.6)$$

which is the ratio of the prestige weight, r_i , to the social surplus generated from college i setting price equal to marginal cost. Given that its peers have an average admissions rate of $a_i(\bar{a}_{-i})$, college i has a best response of:

$$a_i(\bar{a}_{-i}) = \theta_i - \lambda_i \left(\frac{\theta_i^2}{\bar{a}_{-i}} \right). \quad (2.7)$$

The best response exhibits three important properties. College i 's best response function is: (i). increasing in average admissions rate of its peers \bar{a}_{-i} (ii). decreasing in its own prestige ratio, λ_i (iii). increasing in the no-prestige monopoly admissions rate (at a decreasing rate) if the average admissions rate of its peers is sufficiently large, i.e. $\bar{a}_{-i} > 2\lambda\theta$.

Nash Equilibria in Case of N Identical Colleges

We now use the best response functions to solve for the equilibria of this game in the case of N identical colleges. The equilibrium solution concept that we employ is the Nash equilibrium in which each school is best responding to the admissions decisions of its peers. One important feature of the special case of the model is that in equilibrium each school plays an identical admissions rate because of the symmetry of the problem. Consequently, at the Nash equilibrium $a_i^* = \bar{a}_{-i}^*$. Evaluating the best response function in equation (2.7) at the equilibrium, $a_i^* = \bar{a}_{-i}^*$, we find that the admissions rate in equilibrium takes on one of two values:

$$\boxed{a_{i,\pm}^* = \bar{a}_{-i,\pm}^* = \frac{\theta}{2} (1 \pm \sqrt{1 - 4\lambda})} \quad (2.8)$$

We will refer to $a_{i,+}^*$ as the *high* equilibrium, or \bar{a}_H and we will refer to $a_{i,-}^*$ as the *low* equilibrium, or \bar{a}_L . Both the low and the high equilibrium are increasing in the no-prestige

monopoly admissions rate, θ . Increasing the prestige ratio, λ , however, has opposing effects on the low and high equilibria. In the low equilibrium increasing the prestige ratio increases the equilibrium admissions rate, whereas in the high equilibrium, increasing the prestige ratio decreases the equilibrium admissions rate. The intuition for this result is the following. Changes in the admissions rate affect utility through the prestige component of the utility and through the profit component of the utility function. A one percent change in a college's admissions rate relative to its peers has a constant return of $0.01r$ from the prestige component of utility and a non-linear return coming from the profit component of the college's utility function. Given the concavity of the profit function, the return to increasing the admissions rate is greatest when the admissions rate is low and is smallest when the admissions rate is high.

Conditions for FOCs to Hold

In order for the first order conditions to hold, it must be the case that there is a real valued admissions rate that satisfies equation (2.8). When $\lambda > \frac{1}{4}$, the equilibrium is ill-defined because the FOC generating equation (2.8) is not satisfied for a real number. When $\lambda > \frac{1}{4}$, there is always exists a profitable deviation from playing at an interior point. In particular, it is profitable for a college to undercut its competitors by playing an action $a_i < \left(1 - \frac{1}{4\lambda}\right)\bar{a}_{-i}$. This deviation generates a utility that is greater than one quarter the social surplus generated by perfect competition, which itself is an upper bound on the profits from playing the monopoly admissions rate.⁷ Therefore, in cases where $\lambda > \frac{1}{4}$, colleges compete down towards a corner solution of $a^* = 0$.

2.2.4. Conditions for Interior Solutions

The equilibria in equation (2.8) result in an interior solution $a_{i,+}^* \in (0, 1)$ for the following range of no prestige monopoly admission rates: $0 < \theta < 1$. The no prestige monopoly rate is parametrized by the ratio of the marginal cost c to the maximum willingness to pay P_0 . We

⁷There is no corner solution at $a^* = 1$, since this would require schools to produce below marginal cost

arrive at this parametrization by substituting the expression for the number of applicants in equation (2.2) into the expression for no prestige monopoly admissions rate in equation (2.5). The no prestige monopoly admissions rate is the following function of the parameters of the model:

$$\theta = \frac{1}{2} - \frac{c}{2P_0}. \quad (2.9)$$

The no-prestige monopoly admissions rate is bounded above by $\frac{1}{2}$ because of the assumption of linear marginal cost and the assumption that only students with non-negative willingness to pay apply to the college. In cases where $\theta < 0$, the marginal cost c exceeds the maximum willingness to pay P_0 and colleges set the supply of spaces equal to zero and hence the equilibrium admissions rate is zero. Cases where $\theta > 1$ correspond to either negative marginal cost ($c < 0$) or negative maximum willingness to pay ($P_0 < 0$). Negative marginal cost would assume that colleges receive a subsidy for supplying spaces which would exceed the college's actual cost of supplying the space. Negative maximal willingness to pay would suggest that there is no demand for college except that colleges are willing to pay all students to attend.

Remark: By imposing the condition $0 < \theta < 1$, we rule out negative marginal cost, negative maximum willingness to pay and we also rule out a scenario in which the maximum willingness to pay is less than the constant marginal cost.

Second Order Conditions

The second order condition for a maximum requires that the Hessian matrix of second derivatives, H , be negative semi-definite. For the case of N identical colleges, the Hessian is parametrized by two terms, α , its diagonal entries:

$$\alpha = \frac{\partial^2 V_i}{\partial q_i^2} = \underbrace{-2b}_{<0}. \quad (2.10)$$

and its off-diagonal entries:

$$\beta = \frac{\partial^2 V_i}{\partial q_j \partial q_i} = \frac{1}{\underbrace{(N-1)}_{>0}} \frac{r}{(Q_d \bar{a})^2}. \quad (2.11)$$

This simple parametrization of the Hessian follows from the symmetry of the best response functions. To test that the second order condition is satisfied, we first write the Hessian in compact matrix notation using (α, β) :

$$H_i = \frac{\partial}{\partial \vec{q}} \left(\frac{\partial V_i}{\partial \vec{q}} \right) = (\alpha - \beta) I_{N \times N} + \beta \vec{e}_N \times (\vec{e}_N)^\dagger, \quad (2.12)$$

where $I_{N \times N}$ is the N-by-N identity matrix, \vec{e}_N is a column vector of N ones. For H, to be negative semi-definite, we require that $\vec{x}^\dagger H \vec{x} \leq 0$ for all vectors $\vec{x} = (x_1, x_2, \dots, x_n)$. This condition is satisfied when the following inequality holds:

$$\alpha \sum_{n=1}^N x_n^2 - \beta \underbrace{\left(\sum_{n=1}^N x_n^2 - \left[\sum_{n=1}^N x_n \right]^2 \right)}_{>0} < 0. \quad (2.13)$$

The second order condition is satisfied at the Nash equilibria since we know from equations (2.10) and (2.11) that $\alpha < 0$ and $\beta > 0$.

2.3. Welfare Analysis

The model generates two equilibria $\{a_L, a_H\}$, which are described in equation (2.8). To quantify the efficiency properties of these equilibria, we examine the consumer surplus, the producer surplus and the total social surplus under each equilibrium. Since we observe that colleges play the low admissions rate equilibrium, we benchmark the surplus for both consumers and producers in the low-admissions rate equilibrium, against the counter-factual surpluses under the high-admissions rate equilibrium. We also include welfare calculations for the consumer, producer and social surpluses for the setting in which colleges have no preferences over relative prestige ($\lambda = 0$) and the admissions rate is the monopoly admissions

rate θ . In this setting there is just one equilibrium, the monopoly admissions rate. The monopoly case allows us to tease apart how much of the welfare losses are due to elite schools having preferences for relative prestige.

2.3.1. Consumer Surplus

In a given equilibrium, the admissions rate is $\bar{a} \in \{\bar{a}_L, \bar{a}_H\}$ and Q_d students have applied to each of the N identical schools. The college, which faces a downward sloping demand, sets an equilibrium price $P(Q_d\bar{a})$. Accordingly, the consumer surplus under this equilibrium is the area between the demand curve and the equilibrium price:

$$\begin{aligned} CS &= \frac{1}{2}[P_0 - P(Q_d\bar{a})]Q_d\bar{a} \\ &= \frac{1}{2}(P_0 - P_0 + bQ_d\bar{a})Q_d\bar{a}. \\ &= \frac{1}{2}(bQ_d)Q_d\bar{a}^2. \end{aligned}$$

For convenience, we further simplify the consumer surplus by writing it in units of the maximum willingness to pay P_0 . This simplification comes from imposing the condition that all students with non-negative willingness to pay for college do indeed apply, i.e. $Q_d = \frac{P_0}{b}$. This condition arised in equation (2.2) where we modeled consumer demand. The final expression for the consumer surplus in units of P_0 is:

$$\boxed{CS(\bar{a}) = \left(\frac{\bar{a}^2}{2}\right) Q_d}. \quad (2.14)$$

The consumer surplus increases quadratically in the admissions rate. Consequently, consumer surplus is lower when colleges coordinate on the low admissions rate \bar{a}_L rather than on the high admissions rate \bar{a}_H . We define the difference in consumer surplus that arises when colleges play the low equilibrium instead of the high admissions equilibrium as

$\Delta CS_{L,H} \equiv CS(\bar{a}_L) - CS(\bar{a}_H)$. The difference in consumer surplus is given by:

$$\Delta CS_{L,H} = -\frac{1}{2}Q_d(\bar{a}_H - \bar{a}_L)(\bar{a}_H + \bar{a}_L).$$

The larger the difference between the high and the low equilibrium, the greater the loss in consumer surplus that results from colleges playing the low instead of the high equilibrium. The sum of the high and low equilibrium equals the monopoly admissions rate, i.e. $\bar{a}_H + \bar{a}_L = \theta$, as in equation (2.8). Hence the loss in consumer surplus is:

$$\boxed{\Delta CS_{L,H} = -\left(\frac{\theta(\bar{a}_H - \bar{a}_L)}{2}\right)Q_d}. \quad (2.15)$$

From equation (2.15), the difference in the consumer surplus is also increasing in the monopoly admissions rate. Intuitively, the greater the admissions rate at a college in the absence of preferences over relative prestige, the greater the loss in consumer surplus when the college and its peers coordinate on the low, rather than the high, equilibrium, in the case where they place positive weight on relative prestige.

2.3.2. Producer Surplus

In a given equilibrium, the admissions rate $\bar{a} \in \{a_L, a_H\}$ is the same for all N identical colleges. The relative prestige portion of a college's utility function, therefore, always equals zero in *equilibrium*. The college, which faces a downward sloping demand and operates at a constant marginal cost, c , sets an equilibrium price $P(Q_d\bar{a})$. Accordingly, the producer surplus under this equilibrium is the area between the equilibrium price and the constant marginal cost:

$$\begin{aligned} PS &= [P(Q_d\bar{a}) - c]Q_d\bar{a} \\ &= \frac{1}{2}(P_0 - bQ_d\bar{a} - c)Q_d\bar{a} \\ &= \frac{1}{2}(bQ_d) \left[\frac{P_0 - c}{bQ_d} - \bar{a} \right] Q_d\bar{a}. \end{aligned}$$

As we did with the consumer surplus, we further simplify the producer surplus by writing it in units of the maximum willingness to pay P_0 . The term bQ_d equals P_0 from equation (2.2). Moreover, the term in parentheses $\frac{P_0-c}{bQ_d}$ equals twice the monopoly admissions rate, θ ; hence the producer surplus is:

$$\boxed{PS(\bar{a}) = Q_d(2\theta - \bar{a})\bar{a}.} \quad (2.16)$$

The producer surplus is also a quadratic in the equilibrium admissions rate. The producer surplus equals zero when the equilibrium admissions rate is zero, i.e. $\bar{a} = 0$. Moreover, the producer surplus is maximized when the equilibrium admissions rate equals the monopoly admissions rate, $\bar{a} = \theta$. On the interval $\bar{a} \in [0, \theta)$, the producer surplus is increasing. Consequently, when colleges coordinate on the the low admissions rate equilibrium, rather than the high admissions rate equilibrium, this results in a loss of producer surplus, which we define $\Delta PS_{L,H} \equiv PS(\bar{a}_L) - PS(\bar{a}_H)$:

$$\begin{aligned} \Delta PS_{L,H} &= Q_d [2\theta(a_L - a_H) - a_L^2 + a_H^2] \\ &= Q_d [2\theta(a_L - a_H) - (a_L + a_H)(a_L - a_H)] \\ &= Q_d [2\theta(a_L - a_H) - \theta(a_L - a_H)] \\ &= \theta Q_d(a_L - a_H). \end{aligned}$$

It is convenient to rewrite producer surplus in terms of the difference $a_H - a_L$, which is positive:

$$\boxed{\Delta PS_{L,H} = -\theta Q_d(a_H - a_L).} \quad (2.17)$$

2.3.3. Comparison

The loss in producer surplus from the low equilibrium is exactly twice the loss in consumer surplus. Colleges suffer greater *absolute* losses than students, but students suffer larger

relative losses. In particular, the producer surplus is larger than twice the consumer surplus:

$$PS - 2CS = \underbrace{2\theta Q_d - \bar{a}^2 Q_d}_{PS} - 2 \underbrace{\left[\frac{1}{2} (\bar{a}^2 Q_d) \right]}_{CS} \quad (2.18)$$

$$= 2\bar{a}^2 Q_d \left(\frac{\theta}{\bar{a}} - 1 \right) > 0. \quad (2.19)$$

2.3.4. Social Surplus

In a given equilibrium, the admissions rate is $\bar{a} \in \{a_L, a_H\}$, the total social surplus is the sum of the consumer surplus of equation (2.14) and the producer surplus of equation (2.16):

$$SS(\bar{a}) = \underbrace{\frac{1}{2} \bar{a}^2 Q_d}_{CS} \left(\frac{4\theta}{\bar{a}} - 1 \right). \quad (2.20)$$

Because the monopoly admissions rate θ is always greater than or equal to the equilibrium admissions rate (for an interior solution), the social surplus is always greater than three times the consumer surplus. The difference in the social surplus between the low equilibrium and the high equilibrium $\Delta SS_{L,H} \equiv SS(\bar{a}_L) - SS(\bar{a}_H)$ is equal to one and one half times the difference in the producer surplus between the two equilibria:

$$\Delta SS_{L,H} = -\frac{3}{2} \theta Q_d \underbrace{(\bar{a}_H - \bar{a}_L)}_{-\Delta PS_{L,H}}.$$

The negative value of the difference in the social surplus reflects the reality that both colleges and students are independently worse off under the low admissions rate equilibrium relative to the high admissions rate equilibrium. The low admissions rate equilibrium is therefore Pareto dominated by the high admissions rate equilibrium.

2.4. Equilibrium Selection

The comparative statics of the model provide clear guidance on which equilibrium is played. We will exploit this to show that elite schools play the low admissions equilibrium.

Proposition: If increases in demand for the N colleges result in a decrease in the admissions rate then the colleges are coordinating on the low admissions rate $\bar{a}_L = \bar{a}_-$. If instead, increases in the demand for colleges results in an increase in the admissions rate, then colleges are coordinating on the high admissions rate $\bar{a}_H = \bar{a}_+$.

Proof. The demand function is $Q(P) = \frac{P_0}{b} - \frac{P}{b}$. Increases in P_0 , the intercept of the inverse demand function, shift the demand curve outward, while preserving the price sensitivity of demand, b . The comparative statics for the high and low equilibrium is given by:

$$\frac{d\bar{a}_\pm}{dP_0} = \frac{1}{2} \left(\frac{d\theta}{dP} \right) (1 \pm \sqrt{1 - 4\lambda}) \pm \frac{(-\theta) \left(-\frac{d\lambda}{dP_0} \right)}{\sqrt{1 - 4\lambda}}. \quad (2.21)$$

To simplify the comparative statics in equation (2.21), we rely on the following relationships:

$$\frac{d\lambda}{dP_0} = -\frac{2br}{(P - c)^3} = -\frac{2\lambda}{P_0 - c}, \quad (2.22)$$

$$\frac{d\theta}{dP_0} = \frac{1}{2bQ_d} - \left(\frac{P_0 - c}{2bQ^2} \right) \left(\frac{dQ_d}{dP_0} \right) = \frac{1}{2bQ_d} (1 - 2\theta). \quad (2.23)$$

This simplifies the comparative statics in the following way:

$$\frac{d\bar{a}_\pm}{dP_0} = \left(\frac{\theta}{2(P - c)} \right) \left((1 - 2\theta) [1 \pm \sqrt{1 - 4\lambda}] \pm \frac{4\lambda}{\sqrt{1 - 4\lambda}} \right). \quad (2.24)$$

□

To complete the first part of the proof, we show that for positive monopoly admissions rates ($\theta > 0$), the comparative static on the low admissions equilibrium is always negative, i.e.

$\frac{d\bar{a}_-}{dP_0} < 0$. For $\theta > 0$, we show that we have a contradiction if $\frac{d\bar{a}_-}{dP_0} > 0$:

$$\begin{aligned}
& \frac{da_-}{dP_0} > 0 \\
& \implies \left((1 - 2\theta) [1 - \sqrt{1 - 4\lambda}] - \frac{4\lambda}{\sqrt{1 - 4\lambda}} \right) > 0 \\
& \implies \left((1 - 2\theta) - \frac{4\lambda}{[1 - \sqrt{1 - 4\lambda}] \sqrt{1 - 4\lambda}} \right) > 0 \\
& \implies \underbrace{-\frac{1}{2} \left(\frac{1 - \sqrt{1 - 4\lambda}}{[1 - \sqrt{1 - 4\lambda}] \sqrt{1 - 4\lambda}} \right)}_{< 0} > \theta \\
& \implies -\frac{1}{2} \left(\frac{1}{\sqrt{1 - 4\lambda}} \right) > \theta \\
& \implies 0 > \theta.
\end{aligned}$$

The final part of the proof involves demonstrating that the comparative static for \bar{a}_H is positive. First, we observe that for non-negative marginal cost such that $0 < c < P_0$, it is always the case that the monopoly admissions rate is less than half, i.e. $\theta = \frac{1}{2} - \frac{c}{P_0} < \frac{1}{2}$ from equation (2.5). It is evident from equation (2.24) that the comparative static is positive for the high admissions equilibrium \bar{a}_+ if the monopoly admissions rate, θ , is less than $\frac{1}{2}$.

Remark: Since we observe increasing demand and falling admissions rates at elite colleges, elite colleges must be playing the low admissions equilibrium $a_L = \bar{a}_-$.

2.5. Parameter Estimation

From the sharp comparative statics, we have determined that elite schools play the low equilibrium a_L . In the data we observe the equilibrium admissions rate \bar{a} , and from the data we can estimate the elasticity of supply with respect to demand, $\epsilon_{s,d}$. The model gives us an expression for the two unknown parameters: θ the monopoly admissions rate and λ the prestige ratio, as functions of the two observed quantities – the admissions rate and

the demand elasticity of supply. We solve for the parameter values using the method of elimination given that we have two equations and two unknowns.

2.5.1. The Elasticity of Supply

Proposition: If the demand elasticity of the admissions rate, ϵ_a equals -1, then the number of spaces offered at a college does not change with increases in demand. If $\epsilon_a < -1$ then the supply of slots decreases with increases in student demand, otherwise increased student demand results in an increase supply of spaces.

Proof.

$$q_s = \bar{a}Q_d \quad (2.25)$$

$$\implies \frac{dq_s}{dP_0} = Q_d \frac{d\bar{a}}{dP_0} + \bar{a} \frac{dQ_d}{dP_0} \quad (2.26)$$

$$\implies \frac{dq_s}{dP_0} = bP_0 \frac{d\bar{a}}{dP_0} + \frac{a}{b} \quad (2.27)$$

$$\implies \frac{dq_s}{dP_0} = \frac{a}{b} \left(\frac{P_0}{a} \frac{d\bar{a}}{dP_0} + 1 \right) \quad (2.28)$$

$$\implies \frac{dq_s}{dP_0} = \underbrace{\frac{a}{b}}_{>0} (\epsilon_a + 1) \quad (2.29)$$

$$\implies \frac{b}{a} \frac{dq_s}{dP_0} = \frac{Q_d}{q_s} \frac{dq_s}{dQ_d} = \epsilon_{a,P_0} + 1 \quad (2.30)$$

$$\implies \epsilon_{s,d} = 1 + \epsilon_d \quad (2.31)$$

where $\epsilon_{s,d}$ is the elasticity of supply with respect to demand. The sign of $\frac{dq_s}{dP_0}$ depends on the sign of $(\epsilon_a + 1)$. In particular:

$$\implies \frac{dq_s}{dP_0} = \begin{cases} < 0 & \text{if } \epsilon_a < -1 \\ = 0 & \text{if } \epsilon_a = -1 \\ > 0 & \text{if } \epsilon_a > -1 \end{cases} \quad (2.32)$$

□

The demand elasticity is a function of the parameters, which we can derive from the model:

$$\epsilon_{\bar{a}_{\pm}, P_0} = \frac{P_0}{\bar{a}} \frac{d\bar{a}}{dP_0} \quad (2.33)$$

$$= \frac{1}{2\theta} - 1 \pm \frac{\lambda}{\sqrt{1-4\lambda}} \quad (2.34)$$

The supply elasticity $\epsilon_{s,d}$, and the demand elasticity $\epsilon_{\bar{a}_{\pm}, P_0}$, are connected by an accounting relationship, $\epsilon_{s,d} = 1 + \epsilon_{\bar{a}_{\pm}, P_0}$ in equation (2.31).

$$\epsilon_{s,d} = 1 + \epsilon_{\bar{a}_{\pm}, P_0} \quad (2.35)$$

$$= \frac{1}{2\theta} \pm \frac{\lambda}{\sqrt{1-4\lambda}} \quad (2.36)$$

We will estimate the supply elasticity $\epsilon_{s,d}$ directly from data by regressing the log of the number of admitted students against the log of the number of applicants. We can then rearrange equation (2.35) to solve for λ as a function of the monopoly admissions rate θ , which is a free parameter, and the number of colleges N and the admissions rate \bar{a}_{\pm} , which are both observed. After performing these steps, we arrive at the following expressions for the prestige ratio:

$$\lambda = \frac{-\kappa_1 \pm \sqrt{\kappa_1^2 + \kappa_1}}{2}, \quad (2.37)$$

where $\kappa_1 = \theta^2(2\theta\epsilon_{s,d} - 1)^2(\bar{a}_{\pm})^2$.

2.5.2. Method of Elimination

To eliminate λ , so that we have an expression of just the monopoly admissions rate and the data, we rewrite the expression for the equilibrium admissions rate to obtain λ as a function of the unknown monopoly admissions rate and the observable variables, and we then equate this expression to equation (2.37). This elimination step produces a relationship between a single unknown parameter, the monopoly admissions rate (θ), and the data. The expression

for λ , from rewriting the equilibrium admissions rate is:

$$\lambda = \left(\frac{\bar{a}_{\pm}}{\theta}\right) \left(1 - \frac{\bar{a}_{\pm}}{\theta}\right) \quad (2.38)$$

The result of this final elimination step is that the monopoly admissions rate satisfies the following relationship:

$$\left[2\left(1 - \frac{\bar{a}_{\pm}}{\theta}\right)\right]^2 - \left[(2\theta\epsilon_{s,d} - 1) \left(\frac{2\bar{a}_{\pm}}{\theta} - 1\right)\right]^2 = 0. \quad (2.39)$$

2.5.3. Estimates

Monopoly Admissions Rate (θ)

In the case where the observed admissions rate corresponds to the admissions rate in the high equilibrium $\bar{a}_+ = \bar{a}_H$, the estimated mean monopoly admissions rate obtained from solving equation (2.39) is:

$$\hat{\theta}_H = \left(\frac{3 + 4\bar{a}_H\epsilon_{s,d}}{4\epsilon_{s,d}}\right) \left(1 - \sqrt{1 - \frac{32\bar{a}_H\epsilon_{s,d}}{(3 + 4\bar{a}_H\epsilon_{s,d})^2}}\right). \quad (2.40)$$

In the case where the observed admissions rate corresponds to the admissions rate in the low equilibrium $\bar{a}_- = \bar{a}_L$, the estimated mean monopoly admissions rate obtained from solving equation (2.39) is:

$$\hat{\theta}_L = \left(\frac{1 + 4\bar{a}_L\epsilon_{s,d}}{4\epsilon_{s,d}}\right) \left(1 - \sqrt{1 - \frac{32\bar{a}_L\epsilon_{s,d}}{(1 + 4\bar{a}_L\epsilon_{s,d})^2}}\right). \quad (2.41)$$

For the special case where the elasticity of supply is zero, the estimated value of the monopoly admissions rate is $\hat{\theta} = \frac{4}{3}\bar{a}_H$, for the case in which the high admissions equilibrium is played, and $\hat{\theta} = 4\bar{a}_L$, for the case in which the low admissions equilibrium is

played.⁸

Prestige Ratio (λ)

The prestige ratio is a function of the observed admissions rate and the monopoly admissions rate. To estimate the prestige ratio from the data, we insert the estimated value of the monopoly admissions rate into equation (2.38). In the case of the high admission equilibrium, the estimate prestige ratio is:

$$\hat{\lambda}_H = \left(\frac{\bar{a}_H}{\hat{\theta}_H} \right) \left(1 - \frac{\bar{a}_H}{\hat{\theta}_H} \right), \quad (2.42)$$

where $\hat{\theta}_H$, the monopoly admissions rate in the high equilibrium, is determined by equation (2.40). Similarly, for the case where the observed admissions rate is due to colleges playing the low admission equilibrium (a_L), the estimated prestige ratio is given by:

$$\hat{\lambda}_L = \left(\frac{\bar{a}_L}{\hat{\theta}_L} \right) \left(1 - \frac{\bar{a}_L}{\hat{\theta}_L} \right). \quad (2.43)$$

2.5.4. Computing the Economic Surplus from the Data

If we are in the low admission rate equilibrium, \bar{a}_L , the consumer surplus $CS(\bar{a}_L)$ can now be calculated directly from substituting the observed admissions rate and the observed number of students that applied to the college, Q_d into equation (2.14). There is no estimation error in computing this value. To calculate the counter-factual consumer surplus under the high admissions rate equilibrium, $CS(\bar{a}_H)$, however, we must estimate the model and use the parameter estimates to construct \bar{a}_H , which we then insert into equation (2.14). The computation for the producer surplus and the social surplus follows the same procedure. We compute standard errors on the derived parameters using the delta method.

⁸These results were obtained by setting the elasticity of supply equal to zero in equation (2.39) and solving for the corresponding monopoly admissions rate. We verified that this gave the same result as taking the limit as the elasticity of supply approached zero of equation (2.40) and (2.41).

2.6. Data

We have admissions data on all eight Ivy League Colleges from 2003-2015. This data was collected on an annual basis by the Ivy Coach, a college admissions consulting company that specializes in preparing students to apply to Ivy League colleges. For our purposes, it is important that this data includes both the admissions rate as well as the number of students applying for each college and the number of students granted admission to the college. We merge this data with rankings data from the US News and World Reports covering the same 13 year period.

2.6.1. Summary Statistics

From 2003 to 2015, the average number of students applying to Ivy League colleges increased by 90%, from 17,183 to 32,656 (Table 10). Over the same period that demand for admissions to the Ivy League almost doubled, the average number of students admitted at an Ivy League college increased by 2.4% from 2,773 to 2,840. With the increase in demand outstripping the increase in supply, the average admissions rate has nearly halved, from 16.1% in 2003 to 8.7% in 2015. This trend in the data is consistent with Ivy league colleges playing the low admissions equilibrium in which increasing demand results in decreasing average admissions rates. It is worth noting that the marginal admissions rate is substantially lower than the average admissions rate. Since 2003, there are 15,473 more applicants competing for 67 new spaces, resulting in a marginal admissions rate of 0.4%.

2.7. Empirical Framework

For the purposes of estimating the model, we focus on the three of the eight Ivy League schools: Harvard, Yale and Princeton (HYP). The goal of focusing on these three is that they come close to satisfying the core assumptions of the model: (i) the N schools are identical, (ii) they compete over prestige and (iii) students will apply so long as their willingness to pay is greater than zero. We take each of these assumptions in turn, using

the summary statistics in Table 11 and Table 12 to demonstrate that the HYP trio come closer to satisfying the assumptions of the model than the non-HYP Ivy league colleges: Columbia, Penn, Brown, Dartmouth and Cornell.

2.7.1. Assumptions

Assumption #1: Identical Colleges

In equilibrium, the N schools each coordinate on the same admissions rate and the average admissions rate in the group equals the admissions rate adopted by each of the schools. Over the sample period, 2003-2015, the mean admissions rate of the HYP schools is 8.2%, and each member of the group has an average admissions rate that falls within one percentage point of the group mean. In particular Harvard has an average admissions rate of 7.5%, Yale 8.2% and Princeton 9.1% (Table 10). In contrast, the non-HYP Ivies have average admissions rates that differ by as much as 6.4 percentage from the group mean. Cornell, for example has an average admissions rate of 20.4% and Brown has an admissions rate of 11.6%, while the group mean is 14% (Table 11).

Assumption #2: Competition over Prestige

It is natural to take the model to data using the schools that are most at risk to be harmed by expanding supply due to prestige considerations. Harvard, Yale and Princeton are consistently ranked as the #1, #2 and #3 schools in the country by US New and World Reports College Rankings. In Table 11, we note that the HYP have an average ranking of 1.82. The average ranking is lower than two because Harvard and Princeton often tie for the top spot in the rankings. With each of these three schools at the very top of the rankings, slight movements in the rankings determine which among these three schools is the overall #1 college in America.

In contrast, the stakes are relatively lower at the non-HYP Ivies (Table 12). Among the non-HYP Ivies, where the average ranking is 10.4, movement in the rankings affects which

schools break into the top 10, rather than determining the overall top spot. Columbia and Penn typically oscillate somewhere around a ranking of #8. Cornell and Brown, on the other hand are, on average ranked, #13, with Dartmouth consistently ranked at #11 on the list of Best US Colleges.

Assumption #3: Minimum $WTP = 0$

We assume that the marginal student who applied to a college has a willingness to pay of 0. This assumption pin downs how many students apply to the college as a function of the parameters of the inverse demand function; it relies on a school pursuing a policy of need-blind admissions with both domestic and international students. If a school is need-blind to domestic, but not international students, as is the case with all non-HYP, excluding Dartmouth (as of 2008), the marginal international student is one whose willingness to pay is a non-zero value that is bounded above by the full cost of tuition. The HYP colleges have all practiced need-blind admissions for both domestic and international students since 2001; therefore, the HYP Ivies are the only colleges in the Ivy league where we can reliably impose this assumption on the willingness-to-pay of the marginal student (Table 11).

2.7.2. Parameter Estimates

The two key parameters that we estimate from the model are the monopoly admissions rate (θ) and the prestige ratio (λ). To estimate these two parameters, we use the observed admissions rate, which we know is \bar{a}_L , from the comparative statics of the model and the observation that increased demand has resulted in decreased admissions rates. We also use the demand elasticity of supply, which we estimate from the data.

Estimating the Demand Elasticity of Supply

From Table 11 we observed that across the Ivy League demand increased by 90% but supply increased by a mere 2.4%. This gives us a estimate of around 0.027 for the demand elasticity

of supply. To obtain a more precise measure of this elasticity, we run a regression of the log number of admitted students at each HYP Ivy, against the log of the number of applicants in that year:

$$\log(\text{no. admits}_{i,t}) = \epsilon_{s,d} \log(\text{no. applicants}_{i,t}) + \text{Controls}_{i,t}, \quad (2.44)$$

In our preferred specification (Table 13), in which we include school fixed effects, a dummy that equals 1 for Princeton and Harvard during the years in which they eliminated their early action program (2007-2011), and a variable for lagged yield, we get an elasticity of $\epsilon_{s,d} = 0.074$ (0.026). The approximated value of 0.027, which we read off from the raw data, lies in the confidence interval of the estimated value of 0.074. Including the lagged-yield is important because it allow us to control for a school strategically admitting more students because last year's yield was low. The coefficient on lagged yield is close to -1, which suggests that schools are using the yield of the prior year to forecast the number of acceptances that it should make in order to fill the class. Absent this control, the supply elasticity estimates would be biased upwards.

Monopoly Admissions Rate & Prestige Ratio

Finding the monopoly admissions rate involves simply substituting the values of $\epsilon_{s,d} = 0.074$, which we estimated, and \bar{a}_L , the average admissions rate, into equation (2.41). In the data, we observe admissions rates for Harvard Yale and Princeton, and denote its average value as \bar{a}_L . We use the same supply elasticity of demand in each year.

Once we have estimates for θ and λ , we can derive values for the marginal cost as a fraction of the maximum willingness to pay ($\frac{c}{P_0}$) and the prestige weight r . In particular $r = 2\lambda Q\theta^2$, in units of P_0 and $\frac{c}{P_0} = 1 - 2\theta$.

2.8. Results

2.8.1. Parameter Estimates

In Table 14, we report the summary statistics for the estimated and derived parameters over the 12 year period 2003-2015. In Figure 10, which appears in the appendix, we provide graphs of the annual point estimates with standard errors computed using the delta method. The average monopoly admissions rate ranges from a low of 26% to a high of 44% and the marginal cost as a fraction of the maximum willingness to pay ranges from a low 0.11 to a high of 0.49. Over time, the monopoly admissions rate is declining and the marginal cost is increasing. One of the most striking results in this table is that the prestige ratio, λ is highly stable across the 13 years in the data (Figure 10). It has a maximum of 0.186 and a minimum of 0.0183, which suggests that in each year the ratio of the school's prestige weight to the total social surplus under perfect competition is a shade smaller than one fifth. The prestige weight, r , as measured in units of the maximum willingness to pay ranges from 765 to 1,272.

2.8.2. Decomposition of Economic Surplus

We use these estimated values of the parameters to calculate the Consumer Surplus, Producer Surplus and the Social Surplus in units of P_0 , the maximum willingness to pay. The results of this exercise are recorded in Table 15.

These results confirm what we know from the model, which is that both producers and consumers are worse off in the low equilibrium relative to the high equilibrium. In the low equilibrium the consumer surplus is on average 81 units, and the producer surplus is on average 1,169 units. In the high equilibrium, the average consumer surplus is 9.6 times higher at 782 units, whereas the producer surplus is a little over two times larger at 2,571 units. In addition to delivering 2,104 additional units of social surplus, play in the high equilibrium results in a less skewed sharing of the social surplus between consumers and

producers. In the high equilibrium the sharing rule is 1 to 3.23 in favor of producers, whereas under the low equilibrium the social surplus is shared in a ratio of 1 to 14.4, also in favor of producers.

2.8.3. A Conservative Dollar Estimate of the Loss in Social Surplus

We place a lower bound on the loss in social surplus by using the tuition sticker price to as a conservative estimate of the maximum willingness to pay. We know from the results in Table 11, that the average tuition at HYP is \$34,871; consequently, a loss of 2,104 units of social surplus is equivalent to an annual loss of \$73M for each of the three peer schools. Of this total \$49M in losses are borne by the schools and the balance of \$24M is borne by the students, some of whom lose because they are not admitted, and others because they pay a higher price conditional on being admitted, since there are fewer students enrolled in the college.

2.9. Conclusions

In the past 13 years, the demand for admissions at Ivy league colleges has increased by two-fold while the supply of spaces at Ivy League schools has increased by less than 3%. In this paper we show that this empirical fact is consistent with a model in which schools have utility over relative prestige. The key economic insight of the model is that schools may not increase supply if increasing supply could result in reduced utility due lower relative prestige. There exists, however, a Pareto improving equilibrium of the model in which all schools may expand supply in the face of increasing demand. Based on our estimates, there is the potential for an increase in social surplus that is on the order of \$73M per annum (for each school). The existence of this Pareto-improving equilibrium suggest that there are welfare gains to be had from allowing elite schools to collude on admissions strategies, particularly if this collusion results in schools colluding to play the high admission equilibrium.

CHAPTER 3 : Property Rights and Economic Growth: A Caribbean Natural Experiment

Abstract: I take advantage of a natural experiment to study the effect of property rights on economic growth. I estimate that per capita GDP in the Bahamas grew by an average of 22% (3% per year over 12 years) in response to a law that limited the ability of non-natives to buy and sell land in the Bahamas. Using an instrumental variables approach, I show that this growth occurred in spite of a downturn in net foreign direct investment due to the passage of the law which weakened the property rights of non-native investors. The results of this study highlight the economic importance of distinguishing between the protection of private property for natives and non-natives as separate channels through which institutions cause economic growth. (JEL E02,O11,O43,O54)

3.1. Introduction

I study a natural experiment in which the property rights of non-native investors is varied in a plausibly exogenous way. The results of this study provide evidence that changes in laws securing the property rights of non-natives may have a different effect on the economic growth than changes in laws securing the property rights of natives. This point, although not explicitly made in the literature, is related to the observation by Acemoglu et al. (2002) that property rights protections that spur economic growth ought to protect a “broad cross section of the society.”

The contributions of my work are two-fold. First, I dichotomize property rights into two camps: (i) property rights protecting both native and non-native investors, and (ii) property rights protecting native investors while leaving non-native investors exposed to the risk of expropriation. To my knowledge, this native vs. non-native distinction has not been explored in the property rights literature. Secondly, I measure the responsiveness of non-natives to changes in property rights that differentially affect them.

The native/non-native distinction that I make in this paper is economically meaningful for two reasons. One, natives can use the political process (voting, uprisings and coups) to discipline the government for encroaching on their property rights, whereas non-natives generally do not possess this option. As a result, non-native investors may be more responsive to adverse changes in property rights than native investors because they lack this political mechanism to deter government expropriation. Two, given the personal and familial connection that natives have to the host country, one can imagine that natives and non-natives have different investment horizons, long term and short term respectively. To the extent that both the investors and the government internalize the heterogeneity in investment horizons between native and non-native investors, any strategic interaction between the investors and the government may be affected differentially by changes in the property rights regime. When taken together, these two points suggest that we should care about how changes in property rights affecting non-natives impact the macro economy – non-natives investors being the group most sensitive to these changes on account of having limited political recourse and a shorter time horizon.

The natural experiment that I study is a change in legislation that limited the ability of non-natives to buy land in the Bahamas. In 1981, the government of the Bahamas enacted the Immovable Property Act (IPA_81), which required non-natives to obtain government approval in order to buy land in the Bahamas. In 1993, this law was repealed and replaced with the International Persons Landholding Act (IPLHA_93), which allowed for non-natives to buy land in the Bahamas without government approval. The IPLHA_93, moreover, placed non-natives and natives on equal footing with respect to the assessment of stamp duties on property purchases. Prior to the repeal of the IPA_81, non-natives paid twice the stamp tax rate paid by natives. In the Bahamas stamp taxes range from 2% (on property valued under \$20,000) to 10% (on properties valued over \$250,000).

For the purposes of this paper, I consider the passage of the IPA_81 to have three primary effects. Firstly, by limiting the ability of non-natives to purchase property in the Bahamas,

the IPA_81 effectively restricted the demand for Bahamian land. This demand restriction and the resulting lower prices in the real estate market can be interpreted as an expropriation of the value of the land by the government from actual investors. The second consequence of the IPA_81 is its value as a signal to future investors of the risk of government expropriation in the Bahamas. Third, the IPA_81 increased the cost of acquiring land in the Bahamas in terms of the procedural costs in both time and money associated with obtaining the government approval and paying higher stamp duties. The implementation of the IPLHA_93 in lieu of the IPA_81 has the opposite effect: (i) increasing demand for land in the Bahamas and its market value, (ii) signaling a lower risk of government expropriation to potential investors than before, and (iii) reducing the transaction cost in terms of both time and money of acquiring property in the Bahamas.

Empirical identification depends on my using Barbados as a reference country and constructing a difference-in-difference estimate of the effect of the IPA_81 on the economic growth rate of the Bahamas. Further, I utilize the within country variation coming from the repeal of the IPA_81 and the enactment of the IPLHA_93 to mitigate against omitted variable bias and to instrument for changes in net Foreign Direct Investment. Regressing the log of per capita GDP in the Bahamas on the dummy variable IPA_81, which represents a weakening of property rights for foreign investors in the Bahamas, I obtain an ordinary least squares estimate (OLS) of 0.15 and an instrumental variables (IV) estimate of 0.22, both of which are significant at the one percent level. Weakening the property rights of non-native investors results in a 15% growth in per capita GDP in the Bahamas. To put this into perspective, annual growth in the Bahamas over the sample period is 5% (1976-2005). This difference-in-difference estimate, is robust to a set of political, economic and natural disaster controls.

3.2. Literature Review

An important theme in the Economic Development and Growth literatures is the role of institutions in causing development – chief among institutions, the protection of private

property. In an influential paper, Acemoglu, Johnson, and Robinson (2001) show that countries with strong property rights protection (historically settlement colonies) are, today, better off economically than those with comparatively weaker property rights protections (historically extractive colonies). Acemoglu, Johnson and Robinson (2001) posit that property rights affect growth because “countries with better institutions, more secure property rights, and less distortionary policies will invest more in physical and human capital, and will use these more efficiently to achieve a greater level of income.” Similarly, La Porta et al (2008) note that property rights help to attract foreign capital by signalling a lower risk of government expropriation. In Banerjee and Iyer (2005), an alternate growth channel is proposed. States with bad initial institutions grow at a slower rate because they must first develop the good institutions necessary for growth, before getting on with the business of economic development itself. The work by Dell (2010) provides additional evidence for this framework of institutional overhang, which describes the persistence of colonial institutions in explaining current economic development. Dell (2010) documents the persistent negative effect of weak property rights on both economic and health outcomes in the case of Peruvian miners. Field (2007) and Hornbeck (2010) provide two additional channels through which property rights may affect economic development. In Field (2007), protection of private property increases labor supply for the Peruvian rural poor as they substitute time spent protecting their domicile for time spent working on jobs. In Hornbeck (2010), a technological innovation in the protection of private property - the introduction of barbed wire - reduced losses to agricultural production by mitigating the risk of one farmer’s crops being trampled by another farmer’s cattle.

The natural experiment in the literature most closely related to my work in this paper is the two-country comparison in Henry and Miller (2010), where they compare Jamaica and Barbados, both former British colonies with similar colonial history and post-colonial institutions. In this paper, Henry and Miller (2010) show that the divergence in long-run economic growth between Jamaica and Barbados can be explained by differences in the macroeconomic policies adopted by the two countries in response to the oil crisis of the

1970's. In this paper, I use similar selection criteria to those used in Henry and Miller (2010). In Figure 11, I observe that both the Bahamas and Barbados have similar colonial origins and post-colonial legal institutions (British Common Law). Moreover, the countries are similar in their population size, land mass (when comparing the main islands)¹, and economic structure. Both the Bahamas and Barbados are service-based economies with modest levels of agricultural production and manufacturing. In the results section of the paper, I will further motivate the validity of this cross country comparison by showing that investor confidence in Barbados is unchanged during the time period that the Bahamas experiences shocks to its property rights institutions due to IPA_81 and IPLHA_93.

The country comparison in this paper (Bahamas and Barbados) represents an advance over the comparison in the Henry Miller paper (Jamaica and Barbados) in that the Bahamas and Barbados match more tightly on the following criteria: population size, primary economic activity and currency policy. Whereas Jamaica has a population size that is several times larger than the population size of the Barbados (3 million and 300,000 respectively), the Bahamas and Barbados have population sizes that differ by roughly 50,000 persons. In terms of economic structure, Jamaica has an important mining industry due to its rich deposits of aluminum ore. Barbados and the Bahamas both have a less developed mining sector and derive most of their income from the service industries of tourism and offshore banking (90% and 78% respectively). Finally Barbados and the Bahamas have both maintained a currency peg with the US (2-1 and 1-1, respectively), for over 40 years. Jamaica on the other hand has maintained a floating exchange rate for the past two decades.

The goal of outlining the details of the Bahamas and Barbados comparison is to demonstrate that it meets the benchmarks established in the literature for natural experiments of this kind and in some dimensions it builds on these standards. One limitation of the two-

¹The comparison of the land mass in Barbados to the land mass of the main island in the Bahamas is validated by the fact that close to 70% of the Bahamian population lives on the main island of New Providence. The rest of the islands in the archipelago are sparsely populated. Equally, the total land mass of the Bahamas comparable to total land mass of Jamaica, which was used as a comparison partner to Barbados in Henry and Miller (2010).

country comparison that I propose relative to the Henry and Miller (2010) comparison is the time difference between when the Bahamas and Barbados became independent. Whereas Jamaica and Barbados both gained independence in 1962, the Bahamas earned the right to self-governance in 1967 followed by formal independence in 1973. The time series that I have on net foreign direct investment only begins in 1975. Consequently, I am unable to explicitly model this difference in my data analysis. Any effect of this distinction likely shows up in the dummy variable for the Bahamas in my estimating regressions.

3.3. Addressing Endogeneity

The source of variation in property rights in this natural experiment is the passage of the IPA_81 in the Bahamas and its subsequent repeal. As with any law, it is difficult to assert that the passage of the law is exogenous relative to the outcome being measured. In the spirit of Dell (2010), I provide some reasons for “plausible exogeneity” or endogeneity that would bias against my finding a negative result of property rights for non-natives on economic growth, which is the central result of my paper.

An important identifying assumption in my paper is that the Bahamas and Barbados are similar countries and as such the passage of a law like the IPA_81 was equally likely in both countries, but simply happened to occur in the Bahamas and not Barbados - a type of historical accident which would make exogeneity of the IPA_81 and its subsequent repeal plausible. Some suggestive historical evidence supporting this proposition is the following. In 2008, the opposition Democratic Labor Party of Barbados articulated a land zoning policy of a similar genre to the IPA_81 in the Bahamas:

The Democratic Labor Party believes that land ownership is too important to the Barbadian identity to be left totally to market forces. The time has come for Government to intervene to ensure that every Barbadian gets an opportunity to own a piece of the Rock. This can best be achieved through effective zoning. The East Coast should be zoned for private residence and ‘inland tourism’ controlled

by locals; and most central parishes zoned for farming. Many areas should be zoned for sale only to Barbadians. (Democratic Labor Party Manifesto, pg. 8)

The Democratic Labor Party of Barbados won the January 15, 2008 election unseating, the ruling Barbados Labor Party, who had governed for the previous 14 years. The purpose of this example is to illustrate that the same considerations that led the government in the Bahamas to enact the IPA_81 are present in a viable political party in Barbados, the control country in this natural experiment. While it remains to be seen whether the DLP in Barbados implements a similar legislation to the IPA_81 in the Bahamas, the goal of offering this historical context is to provide suggestive evidence that the natural experiment that I study in this paper is well-suited to address potential endogeneity concerns by showing that the treated and untreated countries are similar with one important difference being the enactment and subsequent repeal of the IPA_81 in the Bahamas.

Another piece of historical evidence supporting the assertion that the IPA_81 can be thought of as a type of historical accident comes from studying the context under which it was implemented in the Bahamas. The governing party which enacted the Immovable Property Act, the Progressive Liberal Party, won the government in 1967 by forming a coalition government with two independent members of the House of Parliament following a 18-18 tie with the United Bahamian Party in the General Election of 1967. Had the two independents instead joined the the United Bahamian Party in forming the government, it is unlikely that the IPA_81 would have been enacted given the differences in political philosophy between the two parties. Moreover, given the closeness of the election, the political balance of power could have plausibly tipped in either direction.

Thirdly, the IPA_81 was enacted to reduce land speculation in the Bahamas by non-natives. Since the reason for enacting the IPA_81 was a prima facie political restriction on a type of investment in Bahamian land, one would expect that any omitted variable bias would serve to overstate the positive effect of property rights on economic development and on investor confidence. Since I find a negative effect of property rights on per capita GDP, the omitted

variable bias is not operational in my favor for the central result of this paper. ²

3.4. Data

My data set includes per capita GDP, net foreign direct investment figures, and the consumer price indices (CPI) for both countries. The per capita GDP time series extends to 1960. The net foreign direct investment data begins in 1976. The consumer price index data begins from 1960. I also have data on aggregate property purchases in the Bahamas covering 33 years during the period 1972-2009. This data is broken down into land purchases by natives and land purchases by non-natives for each year.

In addition to these financial data, I have a time series on population data, election outcomes, and the dates of major hurricanes for both countries. I use the population data to compute net foreign direct investment per capita for each country-year observation. This normalized measure of net foreign direct investment helps me to control for changes in population growth while providing a natural comparison variable between the measured outcome (per capita GDP) and the proposed channel (foreign direct investment). I use the election data to construct control dummies for changes in government and one year forward lags for changes in government. These government controls in my study mitigate against omitted variable bias due to changes in investor decisions driven by differences in preferences between the political parties within a country, investor aversion to the uncertainty accompanying electoral change, and the implementation of additional policies that may work in concert with IPA_81 or IPLA_93 in driving investor confidence or economic growth. I use the hurricane data to construct control a variable for unforeseen natural disasters. With this natural disaster variable, I control for omitted variable bias due to country-specific shocks.

²In "The Economics of Density: Evidence from the Berlin Wall (2001) Ahlfeldt et al. argue that the Berlin Wall provides an exogenous source of variation in the types of businesses on either side of the wall, since the wall boundaries were determined for military rather than economic reasons. One can similarly argue that the political reasons for enacting the IPA_81 would mitigate against possible endogeneity concerns.

3.5. Identification

My identification strategy depends on two methodologies, (i) an event study methodology, and (ii) a difference-in-differences estimator . In the first stage of my analysis, I use these two methodologies to estimate the effect of property rights (IPA_81) on net Foreign Direct Investment (NFDI). Here I use the IPA as a proxy for property rights, i.e., when the IPA_81 is in effect property rights for non-natives are limited. The purpose of the first stage is to establish that IPA_81 is a valid proxy for a change in property rights that affects investor confidence. In the second stage of my analysis, I use the two empirical approaches to estimate the direct effect of IPA_81 on per capita GDP and the indirect effect of IPA_81 on per capita GDP through its effect on net foreign direct investment.

3.5.1. Event Study

It is standard in event studies to observe pre-treatment and post-treatment data. Discontinuous breaks in the data very close to time of the experimental treatment then capture the treatment effect, assuming exogeneity of the treatment. In this study, I track net foreign direct investment and per capita GDP in both the Bahamas (the treated country) and Barbados (the control country), before and after the two treatments. I construct a 5 year event-study window on both sides of the treatment dates (1981 and 1993). This approach has three benefits: (i). I maximally use the data, since the 1976 data are included in the first event study window; (ii). the event study windows for the two treatments are non-overlapping: 1976-1986 and 1988-1998 and; (iii). the event study windows are large enough to capture time-series variation but narrow enough to contain the immediate effect of the two treatments.

3.5.2. Difference-in-Differences Estimate

The second estimation strategy that I employ is a difference-in-differences estimate, using pooled data for both the Bahamas and Barbados across all years. For the first stage of this

analysis, I regress net foreign direct investment per capita (NFDI) on the property rights dummy IPA_81 , a country dummy for the BAH, a continuous time variable ($TIME$) and the political and climate controls (denoted X).

$$NFDI_i = \alpha_0 + \alpha_1 BAH + \theta_1 IPA_81 + \alpha_2 TIME + \alpha_3 BAH * TIME + \gamma_1 X + u_i \quad (3.1)$$

The property rights dummy IPA_81 equals one (1) for country-year data where the country is the Bahamas and the year of observation coincides with the time period when IPA_81 is operative (1982-1993). It is zero (0) otherwise. The country dummy equals (BAH) one (1) for observations of the Bahamas and zero (0) for observations of Barbados.

The second stage results entail running a similar regression to the one in the first stage, with $\log(\text{per capita GDP})$ replacing NFDI as the outcome variable and NFDI entering the right hand side as an explanatory variable. To test whether property rights influence growth through its effect on attracting foreign investment, I interact the property rights dummy with the NFDI explanatory variable.

$$\begin{aligned} \log(GDP_i) = & \beta_0 + \beta_1 BAH + \theta_2 IPA_81 + \lambda_1 NFDI + \lambda_2 IPA_81 \times NFDI + \\ & \beta_2 TIME + \beta_3 BAH \times TIME + \beta_4 \log(CPI) + \gamma_2 X + v_i \end{aligned} \quad (3.2)$$

3.5.3. Contemporaneous Difference Estimate

3.6. Results

3.6.1. Treatment # 1: Suspension of Property Rights for Non-Natives in the Bahamas (IPA_81)

The first treatment that I study is the enactment of the Immovable Property Act in 1981 (IPA_81). I consider this treatment to be an adverse shock to the property rights of foreign investors in the Bahamas. As illustrated in Figure 12, pre-treatment³ net foreign direct investment (NFDI) is positive (on average) in both the Bahamas and Barbados. Further, pre-treatment NFDI in the Bahamas is generally greater than pretreatment NFDI in Barbados. After the first treatment, NFDI in Barbados remains positive and stable relative to pre-treatment values. This fact is consistent with Barbados being a suitable control country for the IPA_81 policy variation in Bahamas. NFDI in the Bahamas, on the other hand, falls following the passage of IPA_81. This decrease in NFDI in the Bahamas suggests that investor confidence in the Bahamas may have declined in response to the change in property rights in the Bahamas.

In Figure #3, I report the difference in per capita GDP between the Bahamas and Barbados above the pre-treatment average, i.e. $\Delta \text{GDP} = (\text{per capita GDP Bahamas} - \text{average pre-treatment per capita GDP Bahamas}) - (\text{per capita GDP Barbados} - \text{average pre-treatment per capita GDP Barbados})$. As time approaches the event date, 1981, the GDP gap between the two countries, (ΔGDP) is steady at around \$600 per capita. After the event date, the GDP gap jumps by \$400, which is equivalent to 6% of the Bahamas's per capita GDP in 1981 or equally 10.5% of Barbados's per capita GDP in 1981. Thereafter, the GDP gap continues an upward trajectory, increasing at a rate of \$1,000 every two years. Based on Figure 13, the average increase in the GDP gap (1981-1986) is \$2,033 or 31% of the Bahamas's GDP in

³Since the Immovable Property Act was passed by the Bahamian government in November of 1981, one can consider the observed value of net FDI in 1981 to be minimally contaminated by the treatment of the act being passed. For consistency purposes, in the second treatment, I also include the data from 1993 in the pre-treatment data, since the Immovable Property Act is repealed on December 31st of that year.

1981. The Bahamas therefore develops an income advantage over Barbados notwithstanding the reduction in foreign direct investment accompanying the passage of the Immovable Property Act, which weakens the property rights of non-natives in the Bahamas.

In the baseline specification, my OLS results yield a difference-in-differences estimate of 15% for the effect of IPA_81 on per capita GDP in the Bahamas (Table 16). This estimate is significant at the one percent level and robust to controls for changes in government, lagged changes in government, lagged changes in net foreign-direct investment and hurricanes.

3.6.2. Treatment #2: Return of Property Rights for Non-Natives in the Bahamas (IPLHA_93)

The second treatment in this natural policy experiment is the repeal of the IPA_81 and the enactment of the IPLHA_93 in its place. Leading up to the General Elections of 1992 in the Bahamas, the opposition party established as a legislative priority “to repeal the intimidating Immovable Property Act and enact an Investment Act designed to promote and enhance economic growth and employment” (FNM Manifesto). After the opposition party won the government, the Immovable Property Act (IPA_81) was repealed on December 31, 1993 and replaced with the International Persons Landholding Act (IPLHA_93). Under the International Persons Landholding, non-natives could purchase land in the Bahamas without government approval.

This second treatment provides an additional event study where I observe the effect of property rights legislation on economic development through the foreign investment channel. Again using net foreign direct investment in Barbados as control, I get an across country comparison for the effect of strong property rights institutions on investment. Secondly, the legislative repeal enables me to do a within country comparison to the initial treatment.

In Figure 14 we observe that prior to the treatment, total net foreign direct investment in the Bahamas and in Barbados is on average \$71 per capita and \$34 per capita, respectively. In the five years following the treatment in 1993, net foreign direct investment in Barbados

remains stable at an average of \$44 per capita.⁴ Per capita net foreign direct investment in the Bahamas, on the other hand, increases almost six-fold during the post treatment period, averaging \$401 per capita. Any concerns that this NFDI increase is due to the change in government in the Bahamas is obviated by the fact that NFDI in the Bahamas is fairly constant from 1992-1994. The government changes in 1992, the Immovable Property Act (IPA.81) is repealed on Dec 31st, 1993, and the IPhLA_93 comes into effect on January 1st, 1994. The time delay between the government change and the repeal of the Immovable Property Act is a nice feature of this natural experiment. I use this feature of the experiment to un-bundle the effect of the change in government from the effect of the repeal of the IPA_81 (enactment of IPhLA _93) on foreign direct investment in the Bahamas. Looking at Figure 14, the steady increase in net foreign direct investment occurs from 1994 onwards. This observation is consistent with a story in which net foreign direct investment in the Bahamas grows because non-natives face less uncertainty.

In Figure 15, I report the difference in per capita GDP between the Bahamas and Barbados above the pre-treatment average, i.e. $\Delta \text{GDP} = (\text{per capita GDP Bahamas} - \text{average pre-treatment per capita GDP Bahamas}) - (\text{GDP per capita GDP Barbados} - \text{average pre-treatment per capita GDP Barbados})$. As the time approaches the event date, 1993, the GDP gap between the two countries is steady at around \$500 per capita. After the event date, the GDP gap narrows by \$250 in 1994 and oscillates about zero over the next four years in my event study window. The average change in the GDP gap between the two countries is \$4.20, or 0.03% of the Bahamas' GDP in 1993. Here we find that property rights do in fact positively affect net foreign direct investment. Net foreign direct investment, however, has very low pass through to economic growth, i.e. a net increase in per capita net foreign direct investment of \$330 is associated with an average increase of \$4.20 in per capita GDP.

The results of this second treatment, motivate the use of the IPLHA_93 as an instrument for net foreign direct investment in the GDP regression. The IPLHA_93 has a significant

⁴This fact that net-FDI in Barbados varies continuously across the discontinuity at 1993 is again reassuring because Barbados is the comparison country which does not receive the treatment.

effect on net foreign direct investment in the Bahamas, the treated country, and no effect on Barbados the control country (Table 17). Additionally, the IPhLA_93 has an insignificant direct effect on per capita GDP, i.e. it affects per capita GDP through its effect on net foreign direct investment. In the OLS regression of log per capita GDP where I include the effects of IPLHA_93 and a time trend of IPLHA_93, the point estimates on these two control variables are insignificant, t-values of -0.35 and 1.5 respectively. Moreover, due to the timing of the IPLHA_93, and the fact that it replaces another law, it is unbundled to some degree from the existing political and legislative background of other laws and policies, which may have worked in concert with the IPA_81 in influencing foreign investment in the Bahamas.

In the first stage of my two stage least squares instrumental variables estimate, I regress net foreign direct investment on the instruments IPLHA_93 and a time trend of IPLHA_93 (TIME_IPLHA_93) as well as on the controls that appear in the baseline model for my regression on log(per capita GDP).

$$\begin{aligned}
 NFDI = & \alpha_0 + \alpha_1 BAH + \theta_1 IPA_{.81} + \alpha_2 TIME + \alpha_3 BAH * TIME + \\
 & \alpha_4 IPLHA_{.93} + \alpha_5 TIME_IPLHA_{.93} + \log(CPI) + \epsilon
 \end{aligned}
 \tag{3.3}$$

Using the coefficients from the first stage, I construct an estimated NFDI, which I denote NFDI_2SLS. In the second stage, I use this constructed value NFDI_2SLS in the place of NFDI in the regression of log(per capita GDP). The second stage regression that I then run is:

$$\begin{aligned}
 \log(GDP) = & \beta_0 + \beta_1 BAH + \theta_2 IPA_{.81} + \lambda_1 NFDI_2SLS + \lambda_2 IPA_{.81} \times NFDI_2SLS + \\
 & \beta_2 TIME + \beta_3 BAH \times TIME + \beta_4 \log(CPI) + \gamma_2 X + v
 \end{aligned}
 \tag{3.4}$$

Using this instrumental variables approach, I obtain an estimate of 22% for the effect of IPA_81 on per capita GDP in the Bahamas (Table 17). This result is significant at the one percent level and is robust to electoral, hurricane and lagged NFDI controls that are used to check the OLS result.

3.7. Robustness Checks

Internal Validity

In this section, I perform three robustness checks on my data. First I check that the results that I obtained are driven by the changes in the law rather than the choice of Barbados as a control country. I do this by running the OLS regression just using the data from the Bahamas, depending on the enactment of IPA_81 and its subsequent repeal as the source of variation for property rights. Under this specification, I obtain an effect size of 17% for IPA_81 (Table 18). This estimate is significant at the one percent level and lies in the confidence interval for the OLS estimate that I previously found using the pooled difference-in-differences estimator. The coefficient on the other variables in the un-pooled regression likewise lie in the confidence intervals of the point estimates from the pooled regression in which I use data from both the Bahamas and Barbados.

Counter-factual & Placebo Tests

Second, I do a counter-factual analysis that consists of the following two hypotheticals: (i) IPA_81 is not enacted in the Bahamas (ii) IPA_81 is enacted in Barbados. Under the first counter-factual experiment, I obtain a negative insignificant point estimate for net foreign direct investment (Table 18). This result is opposite to the positive (and also insignificant) coefficient that I obtained from factual analysis. Under the second counter-factual experiment, I find an insignificant effect of the fictitious IPA_81 on economic growth in Barbados. The point estimate in the second experiment is 0.9% and the t-value is 0.21. Both counter-factual exercises cohere with the results in the previous section. First,

omitting IPA_81 from the growth equation for the Bahamas leads to a spurious negative correlation between net foreign direct investment and GDP growth. Secondly, including fictitious piece of legislation in Barbados parallel IPA_81 in the Bahamas does not help to explain economic growth in Barbados – the comparison country (Table 18).

3.7.1. Contemporaneous Difference Estimator

The third robustness check that I perform uses a contemporaneous difference estimator to correct for unobserved shocks that are common to both the treatment and control country. The contemporaneous difference estimator is obtained by differencing the dependent and independent variables for the Bahamas (BAH) and Barbados (BDOS) for each time observation and then regressing the differenced depend variable on the differenced explanatory variables. This contemporaneous difference estimator mitigates against omitted variable bias driven by time-specific shocks that affect both the Bahamas and Barbados similarly, e.g. the terrorist attacks of 9/11 which affected tourist travel to both islands. The estimating equation looks like:

$$\log(GDP_{BAH,t}) - \log(GDP_{BDOS,t}) = \hat{\theta}_2 IPA.81 + \hat{\gamma}_2 \bar{X}_t + \bar{v}_t \quad (3.5)$$

where \bar{X}_t and \bar{v}_t are the differenced control variables and error terms at time t. The coefficient $\hat{\theta}_2$ captures the effect of IPA_81 on the difference in growth rates between the Bahamas and Barbados. By comparison, θ_1 from the difference-in-differences estimator is the average effect of IPA_81 on GDP growth rates in the Bahamas. In order to convert from $\hat{\theta}_2$ to θ_1 , I compound the growth difference $\hat{\theta}_2$ over the N=12 years that IPA_81 is in effect and compute it's average. Mechanically, after j years of IPA_81, the growth in the Bahamas due to IPA_81 is given by $(1 + r)^j$, where $r = \hat{\theta}_2$, the differential growth due to IPA_81. I let $\hat{IPA}.81$ denote the average effect of IPA_81 computed using $r = \hat{\theta}_2$.

$$I\hat{P}A_{.81} = \frac{1}{N}[(1+r) + (1+r)^2 + \dots + (1+r)^N] - 1 \quad (3.6)$$

The sum in the average growth equation is a standard geometric series, which simplifies to

$$I\hat{P}A_{.81} = \frac{1}{N} \left(\frac{1+r}{r} \right) [(1+r)^N - 1] - 1 \quad (3.7)$$

Estimating the contemporaneous difference model using the explanatory variables that are significant across all specifications from the difference-in-differences estimation procedure, I obtain a point estimate of $\hat{\theta}_1 = 0.03$ and a standard error of 0.01 (Table 20). Computing the average growth from the expression above, I obtain that the average growth is 0.219 or 21.9%, with a standard deviation of 0.086 or 8.6% (calculated using the delta method). For comparison, I found the average effect of IPA_81 to 15% based on the OLS result and 22% based on the IV result. The results of the contemporaneous difference estimator is significant at the one percent level and falls within the confidence interval of both the OLS and IV estimators.

Correcting for Auto-correlation

In an important paper Bertrand, Duflo and Mullianathan (2004) demonstrate that auto-correlation may lead to standard errors for difference-in-differences estimators which are downward biased. As an additional check for auto-correlation, I regress the residuals from the contemporaneous difference estimator on their lagged values out to five lags (Table 19). The positive auto-correlation at one lag is sensitive to the addition of other lags (alternating t-values of 0.84 and -0.90 in a model with four and five lagged residuals, respectively). There is a robust negative correlation at three lags (t-values -2.83 and -3.24 in a model with four and five lagged residuals, respectively). As a result of this robust negative auto-correlation, the standard errors that I obtained from the contemporaneous difference estimator may be

upward biased (too big) as opposed to being downward biased (too small).

To test whether this bias affects the inference that I make, I consider the following counter-factual: I take the positive auto-correlation at one lag to be robust and significant and the negative auto-correlation at four lags to be insignificant. This counter-factual provides a worse-case scenario for the auto-correlation in my model, i.e. the situation in which the standard errors are largest after I would have corrected for auto-correlation. To rid the model of auto-correlation under this set of assumptions, I perform a Cochrane-Orcutt transformation of the data and estimate the parameters of the model (Cochrane and Orcutt, 1949). The Cochrane-Orcutt transformation works when the residuals of the fit are serially correlated at one lag, i.e. $v_t = \rho v_{t-1} + \eta_t$, where the η_t is exogenous and v_t and v_{t-1} are the residuals from the fit at time t and $t-1$. Given an estimating equation $y_t = X_t\beta + u_t$, the Cochrane-Orcutt transformation is $\tilde{y}_t = y_t - \rho y_{t-1} = (X_t - \rho X_{t-1})\beta + \eta_t$. By taking a weighted first difference of the observations, this procedure eliminates the serial correlation ρv_{t-1} from the error term for the observation at time t . The parameter ρ is estimated from regressing the residuals on their first lagged values (estimated to be 0.43 in Table 19).

The results from this counter-factual experiment give $\hat{\theta}_2 = 0.028$ and a standard error 0.014. This point estimate has a p-value of 0.054 (Table 20). Using this value of $\hat{\theta}_2$ in equation (3.7), I find that the average effect of IPA_81 is 20.2% and the standard error is 0.11.⁵ My preferred estimate of the average effect of IPA_81 is 22%, the common result from the contemporaneous difference estimator and the instrumental variables estimator. The standard error on this point estimate is 5.9% for the contemporaneous difference estimator and 8.6% for the IV estimator, making it significant at the five percent level in both specifications. Recall that the negative auto-correlation result in larger standard errors, which makes it likely that the true standard errors are closer to the standard errors obtained from the

⁵Checking, the residuals of the transformed model, I find that the serial correlation at one lag is eliminated, as expected. Extending the model to more lags, I find that the negative auto-correlation at lag four remains significant at the one percent level and there is an additional negative auto-correlation at lag three that is also significant at the one percent level. The results of this counter-factual exercise suggest that main source of auto-correlation is the negative auto-correlation at lag four. As such, the contemporaneous difference estimator provides a more reliable auto-correlation robustness check.

instrumental variables procedure.

3.8. Conclusion

In this natural experiment, I find that restricting foreign investment through weakened property rights for non-natives resulted in economic growth in the Bahamas. Furthermore, restoring property rights for non-natives did not have a significant effect on GDP growth - its appreciable impact on foreign investment in the Bahamas notwithstanding. One limitation of the work in this paper is that in this paper, I take an in-depth look at two countries as opposed to looking at a cross-section of countries, at a given time. This affects the generalizability of the empirical results to other contexts where the history, politics and geography may be different from those of the Bahamas and Barbados. With this important caveat in mind, these results suggest that it may be important to distinguish between protection of private property for natives and non-natives when considering the mechanism through which property rights institutions promote economic growth.

APPENDIX

A.1. Chapter 1: Appendix

A.1.1. Figures and Tables

Table 1: **Relative Marginal Utility of Minority Neighbors (Diff-in-Diff)**

stats	1970	1980	1990	2000	2010
mean	-8.84	-7.29	-6.82	-6.41	-6.11
	(0.79)	(0.57)	(0.10)	(0.14)	(0.10)
p50	-8.75	-7.10	-6.72	-6.48	-6.06
	(0.18)	(0.10)	(0.29)	(0.10)	(0.10)
N	124	124	124	124	124

Table 2: **Census Tract Tipping Points (Diff-in-Diff)**

All	1970	1980	1990	2000	2010
mean	0.06	0.15	0.21	0.29	0.34
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
p50	0.02	0.07	0.11	0.21	0.29
	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
N	38,489	38,489	38,489	38,489	38,489

In Range (Current Year)[†]	1970	1980	1990	2000	2010
mean	0.18	0.26	0.30	0.36	0.41
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
p50	0.09	0.14	0.19	0.28	0.35
	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
N	22,391	26,399	29,039	31,504	32,447

In Range (Census 2010)[†]	1970	1980	1990	2000	2010
mean	0.07	0.17	0.25	0.35	0.41
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
p50	0.03	0.09	0.15	0.27	0.35
	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
N	32,447	32,447	32,447	32,447	32,447

[†] A tract is in range if its tipping point $\in [0, 1]$.

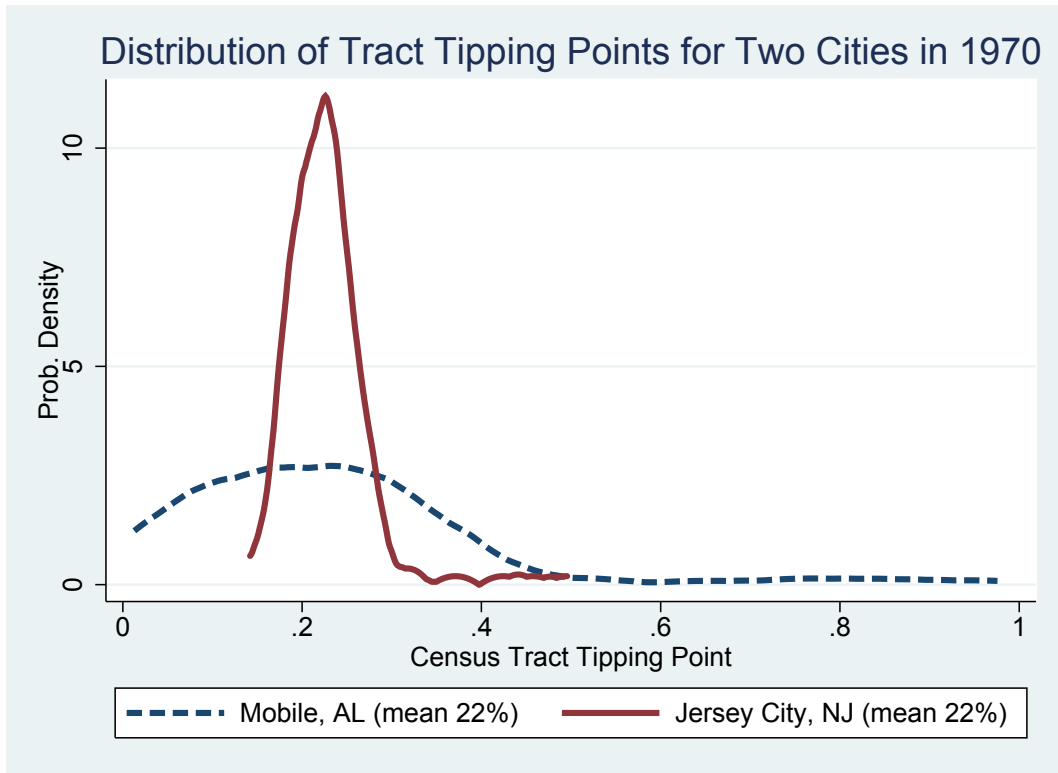


Figure 1: Kernel density plot of estimated census tract tipping points of Jersey City, NJ and Mobile, AL in 1970.

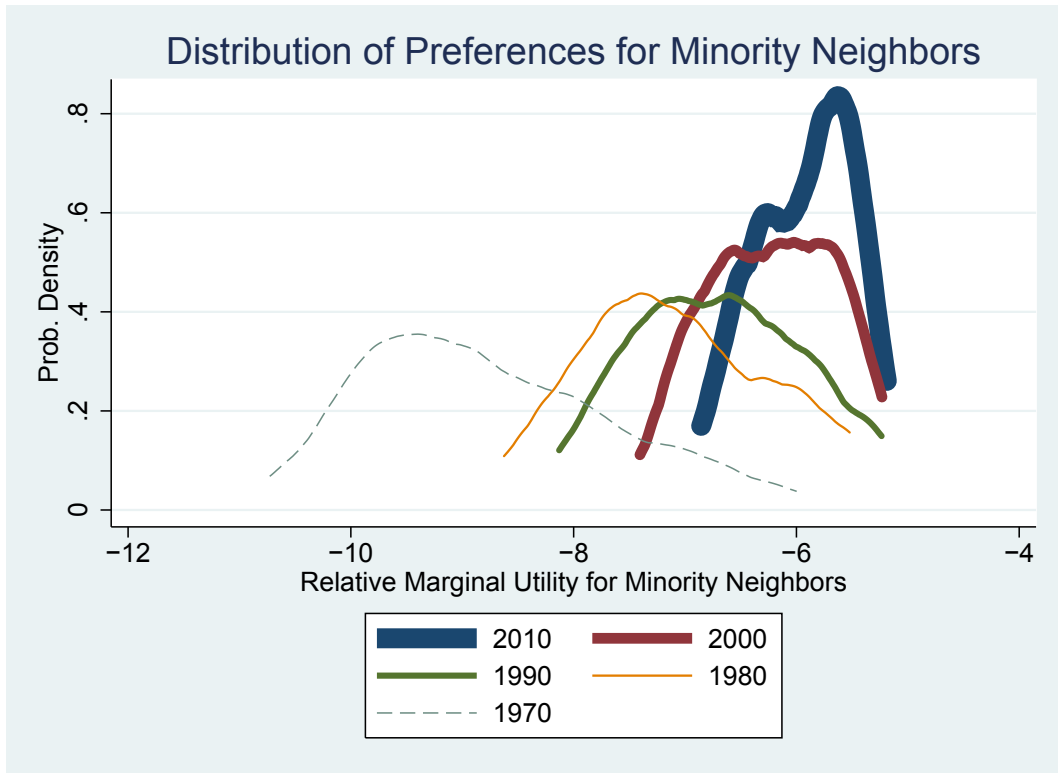


Figure 2: Kernel density plot of the racial preference parameter for each MSA from 1970 to 2010, using the diff-in-diff estimates.

Table 3: **Census Tract Tipping Point (Semi-Parametric)**

All	1970	1980	1990	2000	2010*
mean	0.15	0.22	0.28	0.36	0.42
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
p50	0.13	0.16	0.20	0.28	0.34
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
N	38,489	38,489	38,489	38,489	38,466

In Range (Current Year)	1970	1980	1990	2000	2010
mean	0.15	0.22	0.28	0.36	0.41
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
p50	0.13	0.16	0.21	0.28	0.34
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	38,333	38,085	37,941	37,803	37,694

In Range (Census 2010)	1970	1980	1990	2000	2010
mean	0.16	0.23	0.28	0.36	0.41
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
p50	0.13	0.16	0.21	0.29	0.34
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	37,694	37,694	37,694	37,694	37,694

* I dropped 23 extreme outliers in 2010.

Table 4: **Correlation Tract Tipping Points (Semi-Parametric)**

Year	2010	2000	1990	1980	1970
2010	1.00				
2000	0.71	1.00			
1990	0.58	0.78	1.00		
1980	0.39	0.53	0.73	1.00	
1970	0.18	0.25	0.40	0.77	1.00
# Observations	38,466	38,466	38,466	38,466	38,466

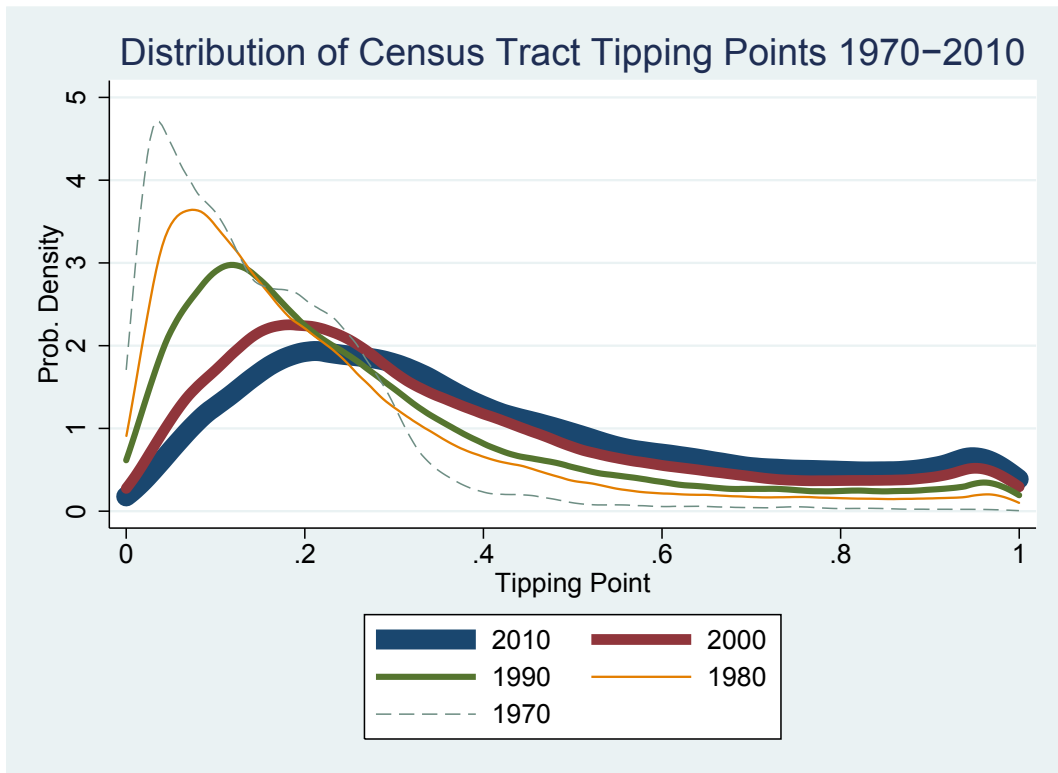


Figure 3: Distribution of census tract tipping points for each census year using the tipping points from the semi-parametric (inverse) method.

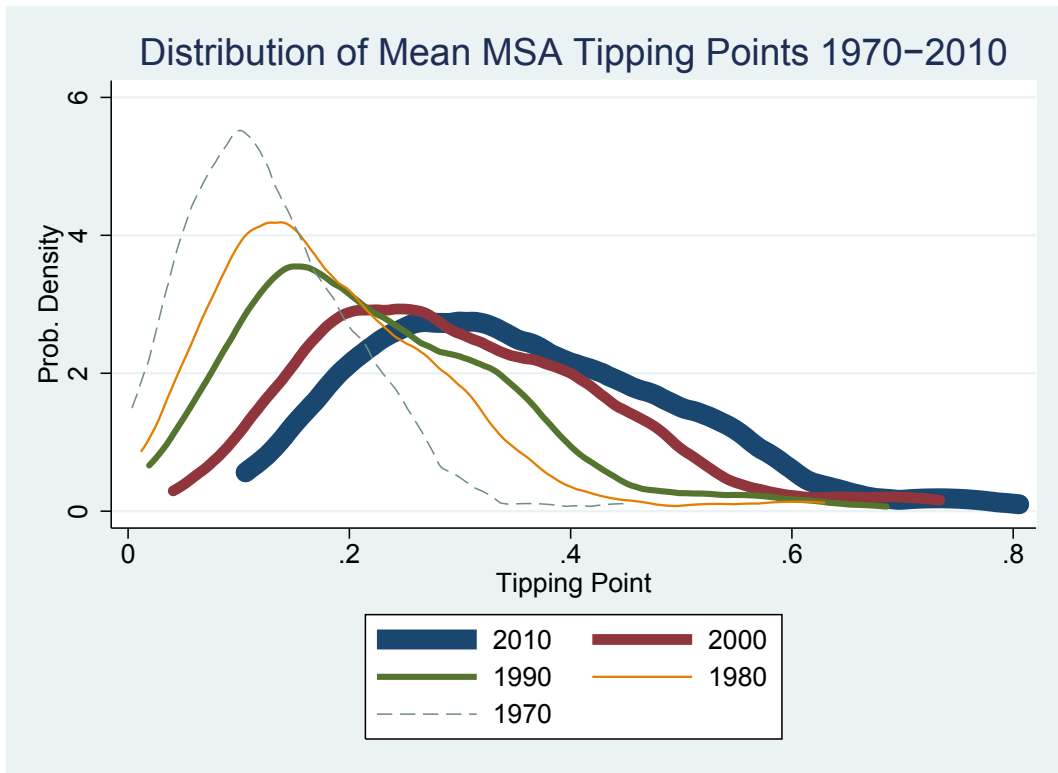


Figure 4: Distribution of mean MSA tipping points for each census year using the tipping points from the semi-parametric (inverse) method.

Table 5: Average MSA Tipping Points

All	1970	1980	1990	2000	2010
mean	0.13 (0.007)	0.18 (0.010)	0.22 (0.011)	0.30 (0.012)	0.35 (0.013)
p25	0.07	0.11	0.13	0.19	0.22
p50	0.12	0.16	0.19	0.26	0.32
p75	0.17	0.24	0.29	0.39	0.46
N	123	123	123	123	123

Northeast	1970	1980	1990	2000	2010
mean	0.09 (0.014)	0.13 (0.019)	0.17 (0.022)	0.24 (0.026)	0.28 (0.028)
p25	0.03	0.06	0.10	0.15	0.18
p50	0.07	0.11	0.15	0.20	0.24
p75	0.11	0.17	0.24	0.30	0.36
N	23	23	23	23	23

Midwest	1970	1980	1990	2000	2010
mean	0.10 (0.010)	0.14 (0.011)	0.16 (0.012)	0.21 (0.012)	0.24 (0.012)
p25	0.07	0.10	0.12	0.17	0.19
p50	0.09	0.14	0.16	0.20	0.22
p75	0.13	0.16	0.19	0.23	0.26
N	27	27	27	27	27

South	1970	1980	1990	2000	2010
mean	0.17 (0.012)	0.22 (0.017)	0.26 (0.020)	0.34 (0.020)	0.41 (0.020)
p25	0.11	0.14	0.18	0.26	0.29
p50	0.15	0.20	0.25	0.32	0.37
p75	0.21	0.27	0.30	0.40	0.49
N	45	45	45	45	45

West	1970	1980	1990	2000	2010
mean	0.12 (0.013)	0.20 (0.021)	0.26 (0.025)	0.36 (0.027)	0.43 (0.028)
p25	0.06	0.10	0.15	0.24	0.32
p50	0.12	0.19	0.28	0.38	0.46
p75	0.17	0.26	0.35	0.45	0.52
N	28	28	28	28	28

Table 6: Median MSA Tipping Points

All	1970	1980	1990	2000	2010
mean	0.11 (0.007)	0.16 (0.010)	0.21 (0.010)	0.28 (0.013)	0.33 (0.013)
p25	0.06	0.10	0.12	0.17	0.22
p50	0.10	0.13	0.17	0.24	0.31
p75	0.16	0.22	0.27	0.37	0.43
N	123	123	123	123	123

Northeast	1970	1980	1990	2000	2010
mean	0.09 (0.013)	0.12 (0.017)	0.15 (0.019)	0.22 (0.025)	0.26 (0.025)
p25	0.03	0.06	0.09	0.14	0.18
p50	0.07	0.10	0.14	0.18	0.23
p75	0.10	0.14	0.21	0.28	0.31
N	23	23	23	23	23

Midwest	1970	1980	1990	2000	2010
mean	0.09 (0.009)	0.12 (0.009)	0.14 (0.09)	0.19 (0.010)	0.22 (0.010)
p25	0.06	0.09	0.11	0.17	0.18
p50	0.08	0.12	0.14	0.17	0.22
p75	0.11	0.14	0.16	0.21	0.24
N	27	27	27	27	27

South	1970	1980	1990	2000	2010
mean	0.14 (0.013)	0.20 (0.017)	0.25 (0.020)	0.32 (0.022)	0.38 (0.022)
p25	0.09	0.13	0.17	0.24	0.30
p50	0.11	0.17	0.22	0.29	0.35
p75	0.19	0.24	0.29	0.37	0.43
N	45	45	45	45	45

West	1970	1980	1990	2000	2010
mean	0.11 (0.013)	0.19 (0.021)	0.25 (0.025)	0.36 (0.027)	0.42 (0.028)
p25	0.05	0.10	0.14	0.24	0.31
p50	0.12	0.17	0.26	0.38	0.43
p75	0.17	0.25	0.34	0.44	0.53
N	28	28	28	28	28

Table 7: Correlation in Mean MSA Tipping Points

Year	2010	2000	1990	1980	1970
2010	1.00				
2000	0.88	1.00			
1990	0.85	0.97	1.00		
1980	0.81	0.93	0.98	1.00	
1970	0.65	0.79	0.84	0.88	1.00
# MSA	123	123	123	123	123

Table 8: Comparison of CMR and Blair Tipping Points

	1990	1980	1970
% Marginal Tracts in MSA (below TP)	0.212 (0.069)**	0.193 (0.067)**	-0.015 (0.055)
% Marginal Tracts in MSA (above TP)	0.311 (0.169)	0.121 (0.157)	0.603 (0.123)**
Constant (Avg. Diff in CMR & Blair MSA TP)	-0.153 (0.021)**	-0.102 (0.020)**	-0.077 (0.017)**
R^2	0.17	0.12	0.26
N	101	100	93

Table 9: Effect of Preferences and Options on Tipping Points

	1970	1980	1990	2000	2010
Minority HHI	-0.011 (0.005)*	-0.034 (0.007)**	-0.070 (0.009)**	-0.097 (0.012)**	-0.144 (0.015)**
Non-minority HHI	0.006 (0.005)	0.026 (0.007)**	0.055 (0.009)**	0.066 (0.012)**	0.112 (0.015)**
Race Prefs.	0.066 (0.007)**	0.076 (0.012)**	0.058 (0.012)**	0.032 (0.011)**	-0.005 (0.008)
Constant	0.128 (0.012)**	0.169 (0.016)**	0.218 (0.018)**	0.293 (0.020)**	0.300 (0.020)**
Region Fixed Effects	Y	Y	Y	Y	Y
R^2	0.55	0.52	0.57	0.55	0.58
N	116	116	118	118	119

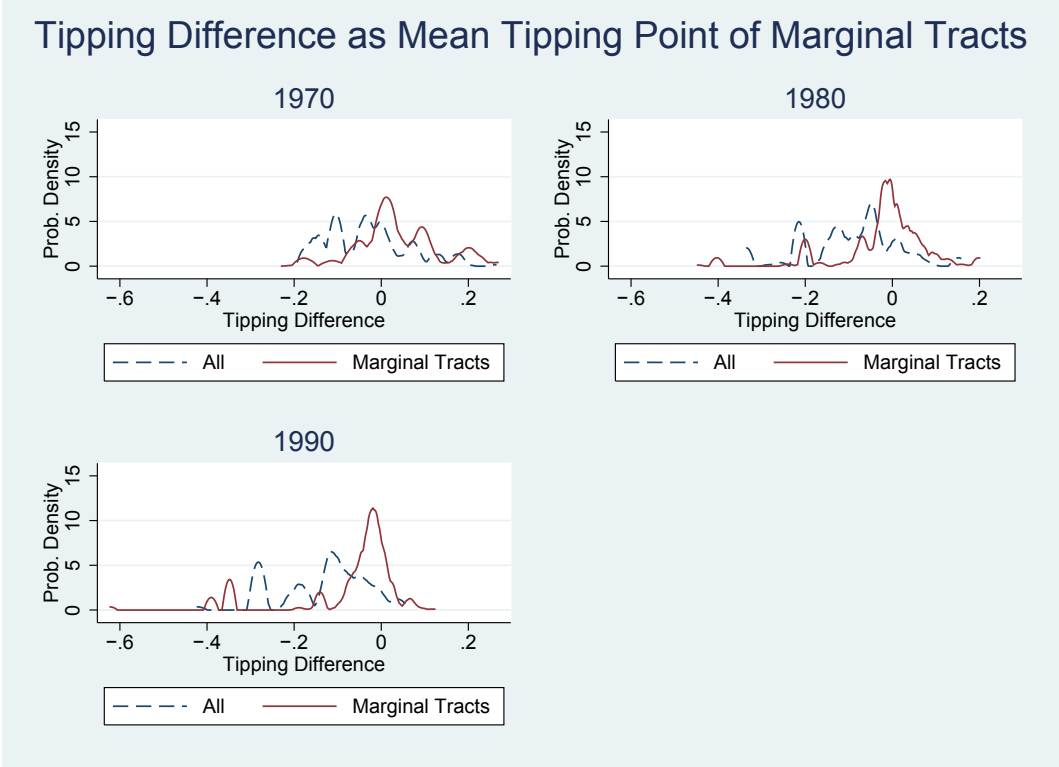


Figure 5: Kernel density plot of the difference between the CMR tipping points and mean MSA tipping points using all tracts, and using the marginal census tracts that are within 5 percentage points of their tipping point.

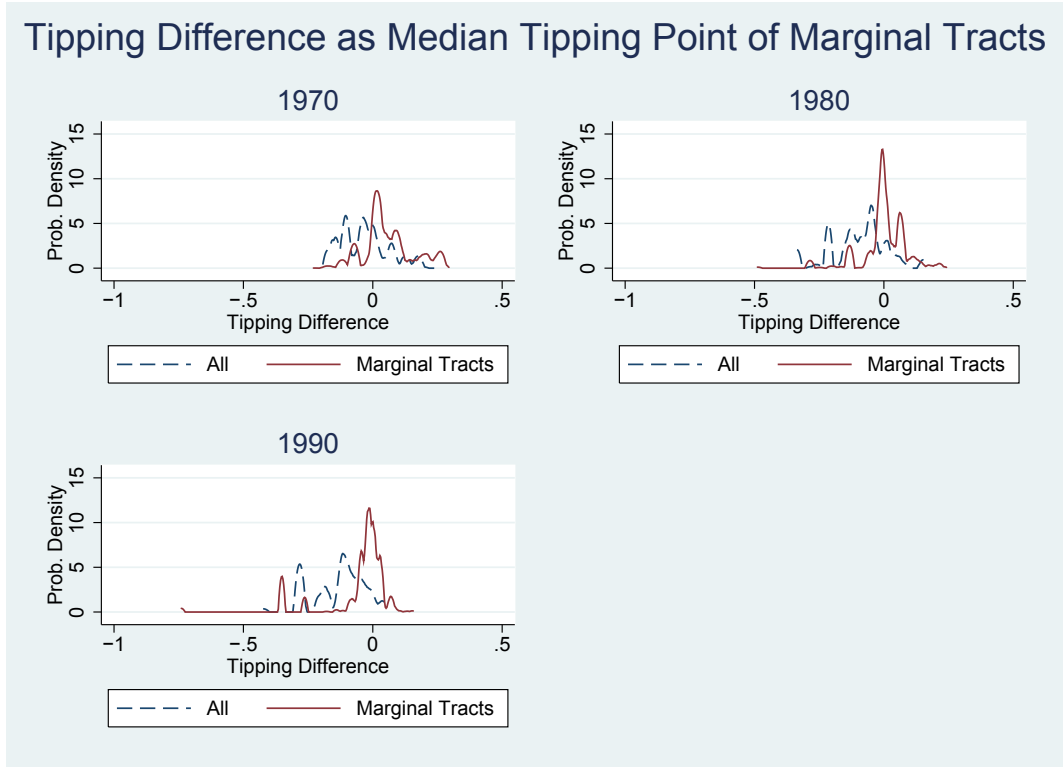


Figure 6: Kernel density plot of the difference between the CMR tipping points and median MSA tipping points using all tracts, and using just the marginal census tracts that are within 5 percentage points of their tipping point.

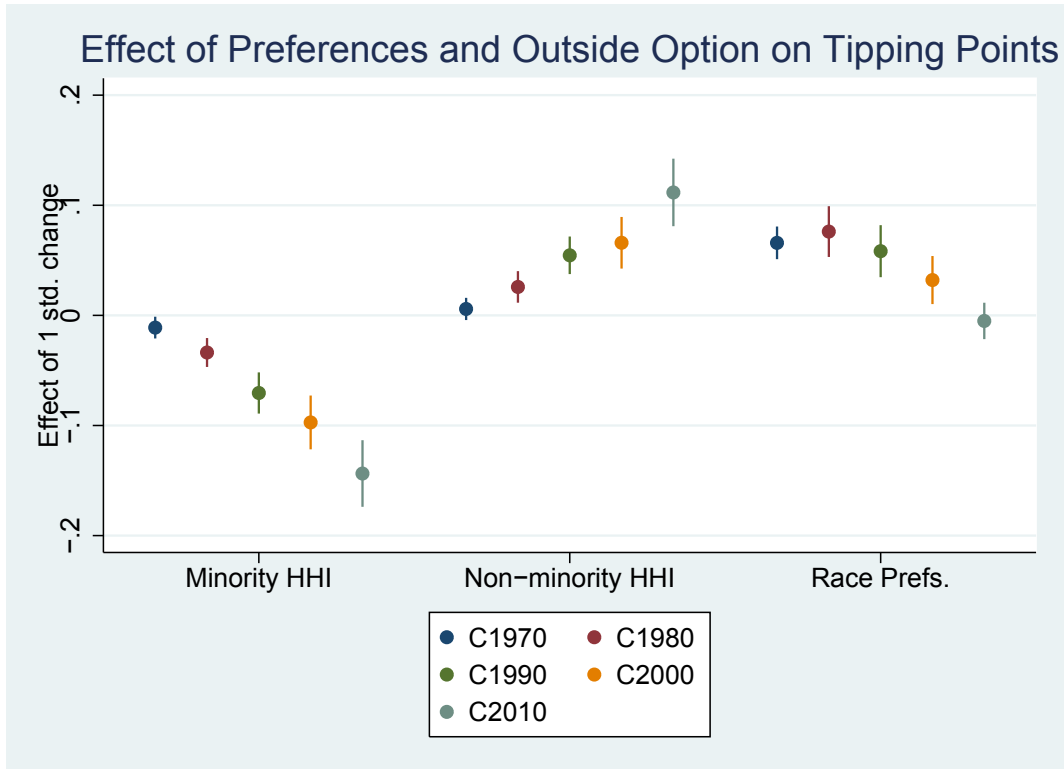


Figure 7: A plot of coefficient estimates from a regression of MSA tipping points on standardized measures of racial preferences and Herfindahl-Hirschman Indices (HHI) of the minority and non-minority concentration of the MSA in census years C1970-C2010.

A.1.2. Computing Tipping Point Standard Errors

Tract Tipping Points

The tipping point of census tract t , in MSA m as a function of the mean utility at the tipping point x_t computed from a 5-th order polynomial:

$$T_{t,m} = \alpha_{0,m} + \alpha_{1,m}x_t + \alpha_{2,m}x_t^2 + \alpha_{3,m}x_t^3 + \alpha_{4,m}x_t^4 + \alpha_{5,m}x_t^5 \quad (\text{A.1})$$

I summarize the tipping points of the N census tracts in the MSA into a single $N \times 1$ vector, \vec{T}_m , using the following matrix equation:

$$\vec{T}_m = M\vec{\alpha}_m. \quad (\text{A.2})$$

In this equation, $\vec{\alpha} = (\alpha_{0,m}, \alpha_{1,m}, \alpha_{2,m}, \alpha_{3,m}, \alpha_{4,m}, \alpha_{5,m})'$ is a column vector of the coefficients on the polynomial used to convert a mean utility into a percent minority at the tipping point, and M is a matrix of powers of the mean utility:

$$M = \begin{pmatrix} 1 & x_{1,m} & x_{1,m}^2 & x_{1,m}^3 & x_{1,m}^4 & x_{1,m}^5 \\ \vdots & \ddots & & & & \vdots \\ \vdots & & \ddots & & & \vdots \\ \vdots & & & \ddots & & \vdots \\ \vdots & & & & \ddots & \vdots \\ 1 & x_{n,m} & x_{n,m}^2 & x_{n,m}^3 & x_{n,m}^4 & x_{n,m}^5 \end{pmatrix}. \quad (\text{A.3})$$

MSA Tipping Points

The MSA tipping point, \bar{T}_m which is an average of the tract tipping point, is given by:

$$\bar{T}_m = \frac{1}{N}(\vec{e}_N)' \cdot \vec{T}_m, \quad (\text{A.4})$$

where \vec{T}_m is the vector of tract tipping point in the MSA and \vec{e}_n is a column vector of N ones used to take the average.

Applying the Delta Method

To compute the covariance matrix of the MSA tipping points, I employ the delta method. The delta method uses a linear approximation of the tipping points in the vicinity of the true parameters to the standard errors of the tipping points as a function of the covariance matrix of the estimated parameters.

Let $\vec{\alpha}^*$ be the true parameter value and $\vec{\alpha}$ be the estimated value of the parameter. Moreover, suppose that $\sqrt{n}(\vec{\alpha} - \vec{\alpha}^*) \rightarrow N(0, \Sigma)$. This implies that $cov(\vec{\alpha}) = \frac{\Sigma}{N}$. Starting with equation (A.2), I approximate the tract tipping points in the vicinity of the true parameters to first order:

$$\vec{T}_m(\vec{\alpha}) \approx \vec{T}_m(\vec{\alpha}^*) + \left(\frac{d\vec{T}}{d\vec{\alpha}} \right) \Big|_{\vec{\alpha}=\vec{\alpha}^*} (\vec{\alpha} - \vec{\alpha}^*) \quad (\text{A.5})$$

$$\implies cov(\vec{T}_m(\vec{\alpha})) \approx \left[\left(\frac{d\vec{T}}{d\vec{\alpha}} \right) \Big|_{\vec{\alpha}=\vec{\alpha}^*} \right] cov(\vec{\alpha}) \left[\left(\frac{d\vec{T}}{d\vec{\alpha}} \right) \Big|_{\vec{\alpha}=\vec{\alpha}^*} \right]' \quad (\text{A.6})$$

$$\implies cov(\vec{T}_m(\vec{\alpha})) \approx M cov(\vec{\alpha}) M' \quad (\text{A.7})$$

$$\implies cov(\vec{T}_m(\vec{\alpha})) \approx \frac{M \Sigma M'}{N} \quad (\text{A.8})$$

The standard errors for the tract tipping points are the diagonal entries of this matrix. For example the standard error on the tipping point of tract t is given by:

$$se(T_{tm}(\vec{\alpha})) \approx \sqrt{\frac{(\vec{i}_t)' M \Sigma M' (\vec{i}_t)}{N}} \quad (\text{A.9})$$

where \vec{i}_t is an $N \times 1$ column vector of zeros in all rows except for the t -th row, which has an

entry of 1.

Standard Error for MSA Tipping Points

The exercise for computing the standard errors for the MSA average tipping point relies on similar implementation of the delta method. Applying the delta method to equation (A.4) I obtain the following result:

$$\bar{T}_m(\vec{\alpha}) \approx \bar{T}_m(\vec{\alpha}^*) + \left(\frac{d\bar{T}}{d\vec{\alpha}} \right) \Big|_{\vec{\alpha}=\vec{\alpha}^*} (\vec{\alpha} - \vec{\alpha}^*) \quad (\text{A.10})$$

$$\implies \text{var}(\bar{T}_m(\vec{\alpha})) = \left[\left(\frac{d\bar{T}}{d\vec{\alpha}} \right) \Big|_{\vec{\alpha}=\vec{\alpha}^*} \right] \text{cov}(\vec{\alpha}) \left[\left(\frac{d\bar{T}}{d\vec{\alpha}} \right) \Big|_{\vec{\alpha}=\vec{\alpha}^*} \right]' \quad (\text{A.11})$$

$$\implies \text{var}(\bar{T}_m(\vec{\alpha})) = \frac{1}{N^2} [(\vec{e}_N)' M] \text{cov}(\vec{\alpha}) [(\vec{e}_N)' M]' \quad (\text{A.12})$$

$$\implies \text{var}(\bar{T}_m(\vec{\alpha})) = \frac{1}{N^3} [(\vec{e}_N)' M] \Sigma [(\vec{e}_N)' M]' \quad (\text{A.13})$$

$$\implies \boxed{\text{se}(\bar{T}_m(\vec{\alpha})) = \sqrt{\frac{[(\vec{e}_N)' M] \Sigma [(\vec{e}_N)' M]'}{N^3}}} \quad (\text{A.14})$$

Standard Error for Average MSA Tipping Point

To compute the average MSA tipping point in a given year, I perform a pooled regress of the minority share of all tracts in all years on coefficients, which are allowed to vary by both MSA and by year. In practice, the MSA tipping points computed above for a given year form the building blocks of this estimation procedure.

A.1.3. Model of Housing Supply

Housing in each neighborhood n is supplied by a single, profit-maximizing firm. The firm sets a price p_n to maximize profits π_n generated from new housing construction:

$$\pi_n = \underbrace{(p_n - MC_n)}_{\text{Price Mark-up}} \times \underbrace{(Q_n - Q_{e,n})}_{\text{Qty. New Housing}} . \quad (\text{A.15})$$

The quantity of housing demanded, $Q_n = Q_n^w + Q_n^m$, is endogenous to the firm's choice of housing price, p_n , as given by the demand functions for both whites (Q_n^w) and minorities (Q_n^m) in equation (1.10). The quantity $Q_{e,n}$ is the existing housing stock in neighborhood n from the previous period. This quantity is pre-determined. The difference $Q_n - Q_{e,n}$ represents the new construction in the neighborhood that is supplied by the local monopolist.

The marginal cost of producing housing, MC_n , is assumed to be a function of the density of households, $\frac{Q_n}{A_n}$, where Q_n is the total number of households in neighborhood n and A_n is the total area of land in neighborhood n . The marginal cost of new construction also depends on the elasticity of supply in the neighborhood, η_n , and on the cost of construction, C_0 , which is assumed to be constant across the city:

$$MC_n = C_0 \left(\frac{Q_n}{A_n} \right)^{\frac{1}{\eta_n}}. \quad (\text{A.16})$$

Given the functional form in equation (A.16), the marginal cost of new construction is higher in denser, more built-up neighborhoods, and likewise higher in cities where construction costs are higher. In the case of a perfectly elastic housing supply, $\eta_n \rightarrow \infty$, and marginal cost equals the cost of construction (i.e., $MC_n = C_0$). In the case of perfectly inelastic housing supply, $\eta_n \rightarrow 0$, and the marginal cost of constructing new housing becomes arbitrarily large since developable land is unavailable. From this limiting behavior, it is evident that the function form in equation (A.16) is consistent with our intuition on the reduced-form relationships between marginal cost and the relevant inputs (construction cost, housing density, and housing supply elasticity).

A.1.4. Solving the Firm's Problem

Firm FOC

To solve the firm's problem, I insert the marginal cost function from equation (A.16) into the profit function of equation (A.15) and take the derivative of profits with respect to price,

p_n . This yields the following first-order condition:

$$\underbrace{\left(\frac{Q_n - Q_{e,n}}{Q_n}\right)}_{\text{Frac. New Supply: } \Delta Q_n} + \underbrace{\frac{P_n}{Q_n} \left(\frac{dQ_n}{dp_n}\right)}_{\text{Demand Elasticity: } \epsilon_{n,d}} \underbrace{\left[1 - \frac{MC_n}{P_n} \left(1 + \frac{1}{\eta_m} \left(\frac{Q_n - Q_{n,e}}{Q_n}\right)\right)\right]}_{\% \text{ Price Mark-up}} = 0. \quad (\text{A.17})$$

The price elasticity of housing demand, $\epsilon_{d,n}$, in equation (A.17), is computed using the demand function in equation (1.10):

$$\epsilon_{d,n} = - \underbrace{\frac{\beta_w Q_n^w (Q_{tot}^w - Q_n^w)}{Q_n Q_{tot}^w}}_{\text{Contribution from Whites}} - \underbrace{\frac{\beta_m Q_n^m (Q_{tot}^m - Q_n^m)}{Q_n Q_{tot}^m}}_{\text{Contribution from Minorities}}, \quad (\text{A.18})$$

$$\implies \epsilon_{d,n} = -\beta_w(1 - f_n)s_{-n}^w - \beta_m f_n s_{-n}^m. \quad (\text{A.19})$$

The term $s_{-n}^w = \frac{(Q_{tot}^w - Q_n^w)}{Q_{tot}^w}$ is the share of whites who live in neighborhoods other than n , and $s_{-n}^m = \frac{(Q_{tot}^m - Q_n^m)}{Q_{tot}^m}$ is the share of minorities who live in neighborhoods other than n . The share of whites (minorities) in other tracts reflects the desirability of those tracts for whites (minorities) relative to the inside option of neighborhood n . As such, the magnitude of the elasticity of demand is increasing in the share of whites (minorities) in the outside option. As previously introduced, f_n is the fraction of minorities in neighborhood n , and β_w and β_m are the price coefficients for whites and minorities in the utility function. For the case where the share of minorities is close to zero, the elasticity of demand is approximately equal to $-\beta_w$, which is the price elasticity of demand for whites ($s_{-n}^w \approx 1$ for large N). The same is the case in the limit as $f_n \rightarrow 1$, where the price elasticity of demand simplifies to $-\beta_m$ ($s_{-n}^m \approx 1$ for large N).

Equilibrium Prices and Mark-ups

Solving this first-order condition of equation (A.17) for housing price as a function of the elasticity of housing demand, $-|\epsilon_{d,n}|$, the elasticity of supply, η_m , marginal cost, MC_n , and

the fraction of the housing supply that is new construction, $\Delta Q_n = \frac{Q_n - Q_{e,n}}{Q_n}$, I get that:

$$p_n = MC_n \left(\frac{1 + \frac{\Delta Q_n}{\eta_n}}{1 - \frac{\Delta Q_n}{|\epsilon_{d,n}|}} \right). \quad (\text{A.20})$$

The term $\frac{\Delta Q_n}{\eta_n}$ is the component of the price mark-ups due to supply constraints, and the term $\frac{\Delta Q_n}{|\epsilon_{d,n}|}$ captures the component of the price mark-ups due to demand pressure. This decomposition is particularly salient when the fraction of new construction is much less than the magnitude of the elasticity of demand and the elasticity of supply, $\frac{\Delta Q_n}{|\epsilon_{d,n}|} \ll 1$ and $\frac{\Delta Q_n}{\eta_n} \ll 1$. In this limit, the denominator in the price equation (A.20) is approximated in the following way:

$$\frac{1}{\left(1 - \frac{\Delta Q_n}{|\epsilon_{d,n}|}\right)} \approx \left(1 + \frac{\Delta Q_n}{|\epsilon_{d,n}|}\right). \quad (\text{A.21})$$

Accordingly, the expression for equilibrium price, p_n , in equation (A.20) is approximated by:

$$p_n \approx MC_n \left(1 + \frac{\Delta Q_n}{\eta_n} + \frac{\Delta Q_n}{|\epsilon_{d,n}|}\right). \quad (\text{A.22})$$

A lower elasticity of supply, η_n , leads to higher prices. The more scarce developable land is, the higher prices are in equilibrium. A lower elasticity of demand, the higher prices are in equilibrium. If households are unresponsive to price increases, then demand pressure bids up prices. Equilibrium prices are also proportional to marginal cost. Higher marginal cost leads to higher prices. Rearranging equation (A.22) to solve for the price mark-up as a fraction of the marginal cost, I find that the mark-up as a percentage of the marginal cost equals the sum of the percentage increase in price due to the demand pressure and the percentage increase in price due to the supply constraints:

$$\text{Mark-up \%} = \frac{p_n - MC_n}{MC_n} \approx \underbrace{\frac{\Delta Q_n}{\eta_n}}_{\text{Supply Constraints}} + \underbrace{\frac{\Delta Q_n}{|\epsilon_{d,n}|}}_{\text{Demand Pressure}}. \quad (\text{A.23})$$

A lower elasticity of supply, η_n , leads to higher price mark-ups. Similarly, a lower elasticity of demand (in magnitude), leads to higher price mark-ups. According to the model, housing

prices, as well as both absolute and relative price mark-ups, should be greater in large, dense cities like New York and San Francisco than in cities like Phoenix or Dallas, which are less dense and more supply and demand inelastic. Moreover, within cities, price and mark-up differentials are higher in neighborhoods near the city center and coast, where available land is scarce and demand is great and more price inelastic.

A.2. Chapter 2: Appendix

A.2.1. Tables & Figures

Table 10: **Ivy League Admissions Numbers (2003-2015)**

Year	No. Applied	No. Admitted	No. Entering
2003	17,183	2,773	1,669
2004	17,059	2,775	1,694
2005	18,717	2,865	1,685
2006	20,260	2,902	1,690
2007	21,100	2,813	1,703
2008	23,404	2,939	1,702
2009	25,293	3,004	1,723
2010	28,071	3,009	1,685
2011	30,717	3,019	1,726
2012	30,328	2,922	1,737
2013	30,910	2,876	1,741
2014	31,682	2,824	1,750
2015	32,656	2,840	.
Average	25,183	2,889	1,709

Table 11: **Characteristics of HYP Ivies: Harvard, Yale, & Princeton (2003-2015)**

College	Rank	Admit Rate	Early Decision	Need-blind	Tuition (\$)
Harvard	1.38	.075	0	1	34,298
Princeton	1.15	.091	0	1	34,388
Yale	2.92	.082	0	1	35,928
Average	1.82	.082	0	1	34,871

Table 12: **Characteristics of non-HYP Ivies (2003-2015)**

College	Rank	Admit Rate	Early Decision	Need-blind	Tuition (\$)
Brown	12.8	.116	1	0	37,199
Columbia	7.77	.094	1	0	38,384
Cornell	12.9	.204	1	0	36736
Dartmouth	10.9	.132	1	.54	36,929
Penn	7.54	.154	1	0	35,099
Average	10.4	.140	1	.108	36,869

Table 13: **Estimating the Demand Elasticity of Supply at HYP Ivies**

	Base	FX	No EA	Lagged Yield
Log No. Applicants	0.180 (0.045)**	0.158 (0.053)**	0.144 (0.046)**	0.074 (0.026)**
No Early Action (H & P)			0.055 (0.024)*	0.044 (0.017)*
Lagged Yield				-0.977 (0.153)**
Constant	5.769 (0.462)**	5.988 (0.533)**	6.117 (0.470)**	7.530 (0.317)**
R^2	0.37	0.40	0.49	0.75
N	39	39	39	36

Table 14: **Estimated Parameters of the Model (2003-2015)**

	mean	sd	max	min
Monopoly Admissions Rate (θ)	0.335	0.067	0.443	0.256
Prestige Ratio (λ)	0.184	4e-07	0.185	0.183
Scaled Marginal Cost ($\frac{c}{P_0}$)	0.33	0.134	0.49	0.11
Prestige Weight ($\frac{r}{P_0}$)	1007	173.7	765.0	1272

Table 15: Welfare Results for HYP Ivies (2003-2015)

Equilibrium Admissions Rate	mean	sd	max	min
a_L	0.08	0.02	0.11	0.06
a_H	0.25	0.05	0.34	0.19

Consumer Surplus	mean	sd	max	min
$CS(a_L)$	81.03	13.41	101.35	62.13
$CS(a_H)$	782.33	140.42	999.05	588.72
$\Delta CS_{L,H}$	-701.31	127.01	-526.59	-897.70

Producer Surplus	mean	sd	max	min
$PS(a_L)$	1,169.12	200.53	1,475.52	889.26
$PS(a_H)$	2,571.73	454.54	3,270.91	1,942.44
$\Delta PS_{L,H}$	-1,402.61	254.02	-1,053.18	-1,795.39

Social Surplus	mean	sd	max	min
$SS(a_L)$	1,250.14	213.94	1,576.87	951.39
$SS(a_H)$	3,354.06	594.95	4,269.95	2,531.16
$\Delta SS_{L,H}$	-2,103.92	381.03	-1,579.77	-2,693.09

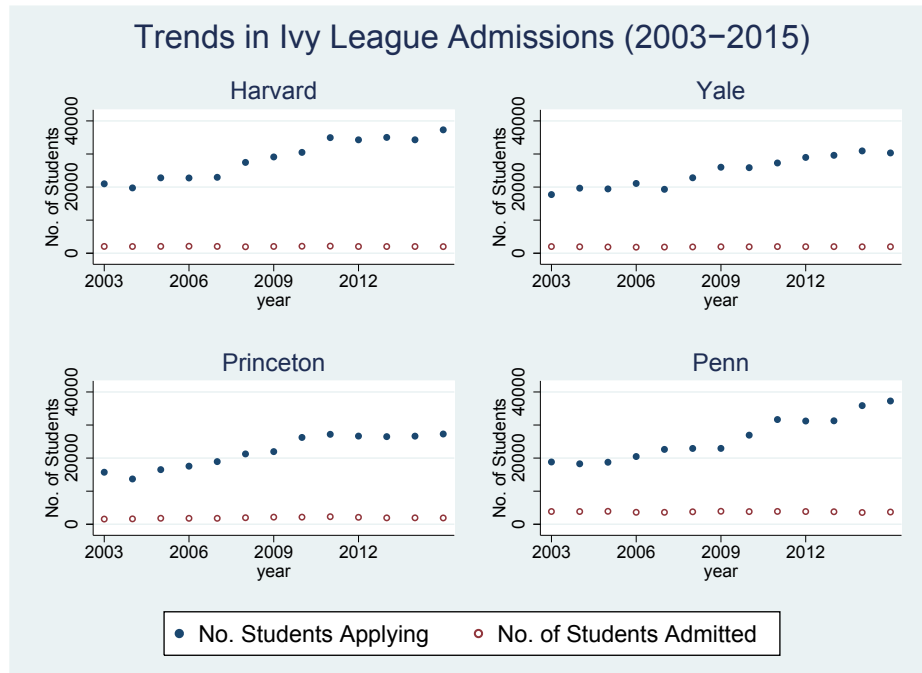


Figure 8: Number of College Applicants and Admits at Harvard, Yale, Princeton and Penn (2003-2015)

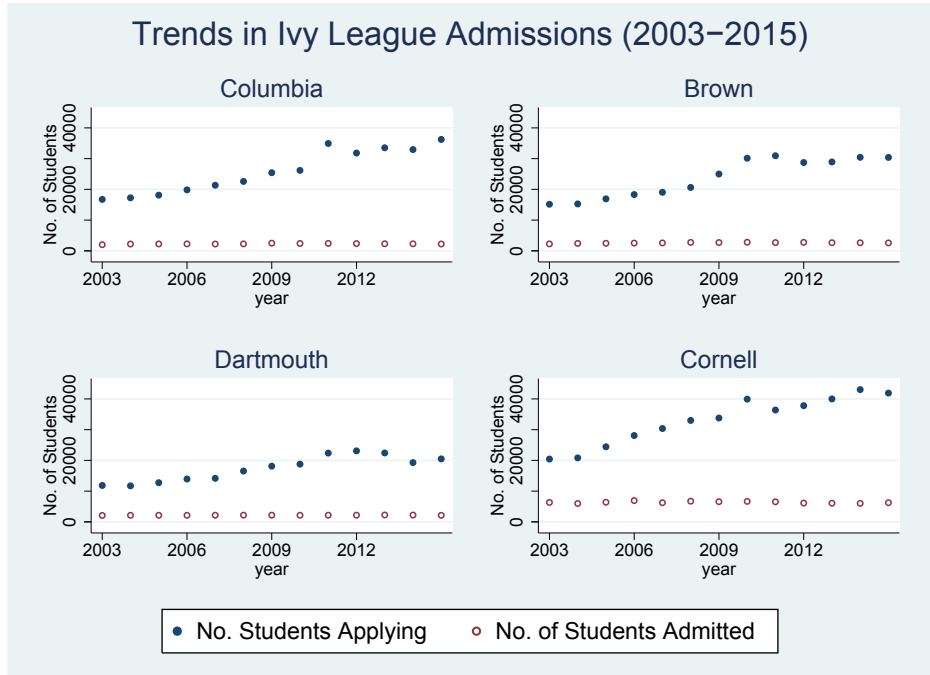


Figure 9: Number of College Applicants and Admits at Columbia, Brown, Dartmouth, Cornell (2003-2015)

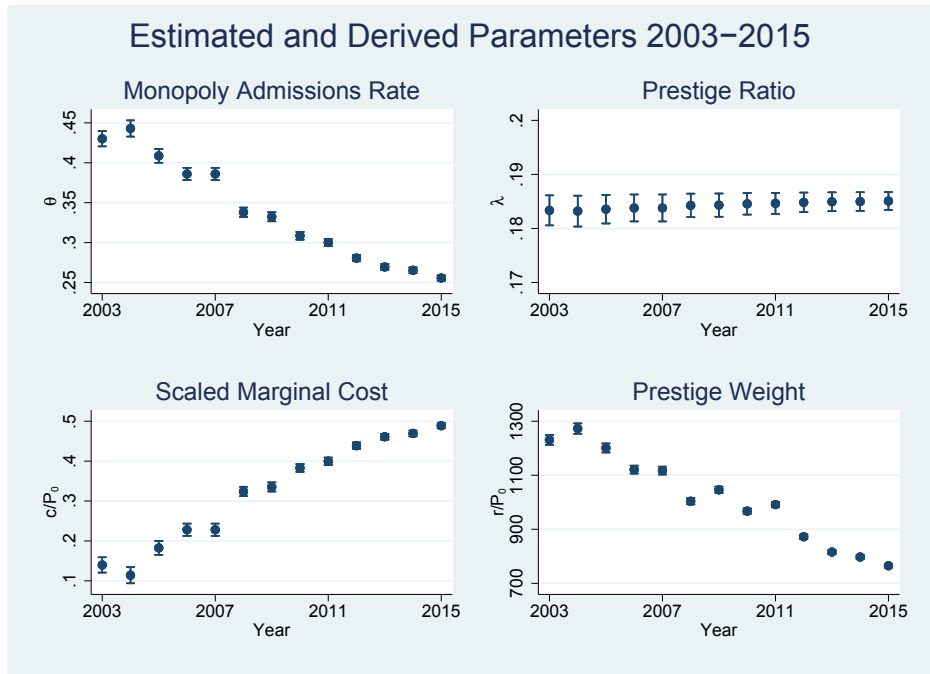


Figure 10: Estimated Parameters from College Prestige Game (2003-2015)

A.2.2. Heterogeneity: Two Types of Colleges

We considered the case of N identical colleges. We now consider the case of two colleges where the colleges differ in their evaluation of relative prestige and the demand and supply parameters. As we found previously the best response functions for colleges is given by:

$$a_1 = \theta_1 - \frac{\gamma_1}{2a^2}a_2 \quad (\text{A.24})$$

$$a_2 = \theta_2 - \frac{\gamma_2}{2a^2}a_1 \quad (\text{A.25})$$

Summing equation (A.24) and (A.25) and dividing by:

$$(a_1 + a_2) = (\theta_1 + \theta_2) - \frac{1}{2a^2}(\gamma_1 a_2 + \gamma_2 a_1) \quad (\text{A.26})$$

$$\implies \bar{a} = \theta_m - \frac{1}{4a^2}(\gamma_1 a_2 + \gamma_2 a_1) \quad (\text{A.27})$$

where $\theta_m = \frac{1}{2}(\theta_1 + \theta_2)$ and $a = \frac{1}{2}(a_1 + a_2)$. We now use the best response functions (A.24) and (A.25) to solve for $\frac{1}{2}(\gamma_1 a_2 + \gamma_2 a_1)$ in terms of the parameters and the average admissions rate. To do this we multiply equation (A.24) by γ_2 and equation (A.25) by γ_1 and take the sum:

$$\frac{1}{2}(\gamma_1 a_2 + \gamma_2 a_1) = \frac{1}{2}(\gamma_2 \theta_1 + \gamma_1 \theta_2) - \frac{\gamma_1 \gamma_2}{2a^2} \frac{1}{2}(a_1 + a_2) \quad (\text{A.28})$$

$$= \rho_m - \frac{\gamma_1 \gamma_2}{2a} \quad (\text{A.29})$$

$$(\text{A.30})$$

where $\rho_m = \frac{1}{2}(\gamma_2\theta_1 + \gamma_1\theta_2)$. Substitute equation (A.29) into equation (eqn0:het2) to express the average admissions rate as a function of the parameters:

$$a = \theta_m - \frac{1}{2a^2} \left[\rho_m - \frac{\gamma_1\gamma_2}{2a} \right] \quad (\text{A.31})$$

$$a = \theta_m - \frac{1}{4a^3} [2\rho_m a - \gamma_1\gamma_2] \quad (\text{A.32})$$

This simplifies to:

$$\boxed{4a^4 - 4\theta_m a^3 + 2\rho_m a - \gamma_1\gamma_2 = 0} \quad (\text{A.33})$$

Generically, there are four possible equilibria in the case where there is heterogeneity and two colleges. First, we will show that this quartic reduces to the expected results when the two colleges are identical. Then we will show how to solve this quartic analytically for a special type of heterogeneity both for $N = 2$ and then for general N .

Claim: In the limit as $\theta_1 = \theta_2 = \theta$ and $\gamma_1 = \gamma_2 = \gamma$, we recover the same solutions as we would from the homogeneous case.

Proof. In this case equation (A.33) reduces to:

$$4a^4 - 4\theta a^3 + 2\gamma\theta a - \gamma^2 = 0 \quad (\text{A.34})$$

$$(4a^4 - \gamma^2) - 2\theta a(2a^2 - \gamma) = 0 \quad (\text{A.35})$$

$$(2a^2 - \gamma)(2a^2 + \gamma) - 2\theta a(2a^2 - \gamma) = 0 \quad (\text{A.36})$$

$$(2a^2 - \gamma)(2a^2 + \gamma - 2\theta a) = 0 \quad (\text{A.37})$$

$$(2a^2 - \lambda\theta^2)(2a^2 + \lambda\theta^2 - 2\theta a) = 0 \quad (\text{A.38})$$

There are four solutions:

$$a_1 = \frac{\theta}{2} (1 + \sqrt{1 - 2\lambda}) \quad (\text{A.39})$$

$$a_2 = \frac{\theta}{2} (1 - \sqrt{1 - 2\lambda}) \quad (\text{A.40})$$

$$a_3 = \theta\sqrt{\lambda} \quad (\text{A.41})$$

$$a_4 = -\theta\sqrt{\lambda} \quad (\text{A.42})$$

□

Notice that a_1 and a_2 are the conjugate pair that we would get from evaluating the the homogeneous case of equation (??) at $N = 2$.

A.2.3. Special Case of Heterogeneous Colleges with $N = 2$

When $\theta_1 \neq \theta_2$ but $\gamma_1 = \gamma_2 = \gamma$, there are tractable closed-formed solutions to the game of $N = 2$ heterogeneous schools, which we will explore¹. This case corresponds to a situation in which the school with the least inelastic demand places a greater weight on prestige. To understand better what this case tells us, recall the definition of γ_i and express it as a function of the model primitives:

$$\gamma_i \equiv \lambda_i \theta_i^2 \quad (\text{A.43})$$

$$= \frac{2b_i r_i}{(P_0 - c_i)^2} \frac{(P_0 - c_i)^2}{(2Q_i b_i)^2} = \frac{\frac{r_i}{2}}{2b_i Q_i^2} \quad (\text{A.44})$$

$$= \frac{\frac{r_i}{2}}{Q_i P_0} \quad (\text{A.45})$$

$$= \frac{\frac{r_i b_i}{2}}{P_0^2} \quad (\text{A.46})$$

¹We can show that for $N = 2$, and more general heterogeneity, $\gamma_1 \neq \gamma_2$ and $\theta_1 \neq \theta_2$, one can use the quartic formula to obtain closed-form solutions that are less tractable than this special case.

This last step comes from the assumption that the marginal applicant got a fee waiver and applied for free. Therefore, at the total number of applications Q_i , the inverse demand equals zero, i.e. $P(Q_i) = 0 = P_0 - b_i Q_i \implies b_1 Q_1 = P_0$. Therefore, gamma is the ratio of the monetary equivalent of a school having an admissions rate that is $\frac{1}{2}$ that of its peer group, to the college's total maximum total revenue, i.e. the revenue if all students who applied were accepted and paid the same prices as the candidate with the highest willingness to pay. Since we assumed that $\gamma_1 = \gamma_2 = \gamma$, it follows from equation (A.46) that:

$$\frac{1}{2}r_1b_1 = \frac{1}{2}r_2b_2 \equiv R. \quad (\text{A.47})$$

$$\implies r_1 = r_2 \frac{b_2}{b_1}. \quad (\text{A.48})$$

If $|b_2| > |b_1|$, then College 2 faces a more elastic demand curve and has a lower economic value of relative prestige than College 2, which faces a more inelastic demand curve. In other words, the case $\gamma_1 = \gamma_2 = \gamma$ correspond to a setting in which the most desirable elite school, i.e. the one that faces the most inelastic demand curve, values relative prestige the most.

Solutions

In a Nash equilibrium, the average admissions rate between the two competing schools must solve the following quartic equation:

$$a^4 - \frac{1}{2}(\theta_1 + \theta_2)a^3 + \frac{\gamma(\theta_1 + \theta_2)}{4}a - \frac{1}{4}\gamma^2 = 0 \quad (\text{A.49})$$

$$\left(a^4 - \left(\frac{\gamma}{2}\right)^2\right) - \left(a^2 - \frac{\gamma}{2}\right)\frac{1}{2}(\theta_1 + \theta_2)a = 0 \quad (\text{A.50})$$

$$\left(a^2 + \left(\frac{\gamma}{2}\right)\right)\left(a^2 - \left(\frac{\gamma}{2}\right)\right) - \left(a^2 - \frac{\gamma}{2}\right)\frac{1}{2}(\theta_1 + \theta_2)a = 0 \quad (\text{A.51})$$

$$\left(a^2 - \left(\frac{\gamma}{2}\right)\right)\left(a^2 + \left(\frac{\gamma}{2}\right) - \frac{1}{2}(\theta_1 + \theta_2)a\right) = 0 \quad (\text{A.52})$$

$$\left(a^2 - \left(\frac{\gamma}{2}\right)\right)\left(a^2 + \left(\frac{\gamma}{2}\right) - \theta_m a\right) = 0 \quad (\text{A.53})$$

There are four solutions to equation (A.53) which characterize the average admissions rate:

$$a_{1,2} = \frac{\theta_m}{2} \left(1 \pm \sqrt{1 - \frac{2\gamma}{\theta_m^2}} \right) \quad (\text{A.54})$$

$$a_{3,4} = \pm \sqrt{\frac{\gamma}{2}} \quad (\text{A.55})$$

The solutions $a_{3,4}$ only hold for the degenerate case where $\theta_1 = \theta_2$. For these cases, the best response functions are $a_1 = \theta_1 - a_2$ and $a_2 = \theta_2 - a_1$ which $\implies \theta_1 = \theta_2$, violating the maintained assumption of heterogeneity that $\theta_1 \neq \theta_2$. Moreover, the average admissions rate in $a_4 < 0$, which is unphysical. This leaves us with two possible average admissions rates in equilibrium $a_{1,2}$. To solve for the action of an individual college, a_i as a function of the average admissions rate, a , we use the best response function in equation (A.24):

$$a_1 = \theta_1 - \frac{\gamma}{2a^2} a_2 \quad (\text{A.56})$$

$$a_1 = \theta_1 - \frac{\gamma}{2a^2} (2a - a_1) \quad (\text{A.57})$$

$$\implies a_1 \left(1 - \frac{\gamma}{2a^2} \right) = \theta_1 - \frac{\gamma}{a} \quad (\text{A.58})$$

$$\implies a_1 = \frac{\theta_1 - \frac{\gamma}{a}}{\left(1 - \frac{\gamma}{2a^2} \right)} \quad (\text{A.59})$$

Similarly, we can show that:

$$a_2 = \frac{\theta_2 - \frac{\gamma}{a}}{\left(1 - \frac{\gamma}{2a^2} \right)} \quad (\text{A.60})$$

A.3. Comparative Static for Heterogeneous Case

To compute the comparative statics of the actions of colleges 1 and 2, we take derivatives of the FOCs:

$$\frac{da_1}{dP_0} = \frac{d\theta_1}{dP_0} - a_2 \frac{d}{dP_0} \left(\frac{\gamma}{2a^2} \right) - \left(\frac{\gamma}{2a^2} \right) \frac{da_2}{dP_0} \quad (\text{A.61})$$

$$\frac{da_2}{dP_0} = \frac{d\theta_2}{dP_0} - a_1 \frac{d}{dP_0} \left(\frac{\gamma}{2a^2} \right) - \left(\frac{\gamma}{2a^2} \right) \frac{da_1}{dP_0} \quad (\text{A.62})$$

$$(\text{A.63})$$

Subtracting equation (A.62) from equation (A.61):

$$\left(\frac{da_1}{dP_0} - \frac{da_2}{dP_0} \right) = \frac{d}{dP_0} (\theta_1 - \theta_2) + (a_1 - a_2) \frac{d}{dP_0} \left(\frac{\gamma}{2a^2} \right) + \left(\frac{\gamma}{2a^2} \right) \left(\frac{da_1}{dP_0} - \frac{da_2}{dP_0} \right) \quad (\text{A.64})$$

$$\implies \left(\frac{da_1}{dP_0} - \frac{da_2}{dP_0} \right) = \frac{\frac{d}{dP_0} (\theta_1 - \theta_2) + (a_1 - a_2) \frac{d}{dP_0} \left(\frac{\gamma}{2a^2} \right)}{\left(1 - \frac{\gamma}{2a^2} \right)} \quad (\text{A.65})$$

To simplify the expression in equation (A.65), we need to find values for three quantities; $\frac{d}{dP_0} (\theta_1 - \theta_2)$, the comparative static on relative demand, $\frac{d}{dP_0} \left(\frac{\gamma}{2a^2} \right)$ the comparative static on the adjustment factor, and the sign of $\left(1 - \frac{\gamma}{2a^2} \right)$, the denominator.

Comparative Static on Relative Demand

We assume that the intercept of the inverse demand curves, P_0 , is the same for both types of schools, but that the schools differ in their marginal cost, $c_1 \neq c_2$, and differ in the price sensitivity of demand, $b_1 \neq b_2$. The price sensitivity of demand and the total number of applicants applying to the schools are related in the following way, we assume that the marginal applicant received an application fee waiver. Since there a total Q_i applications, this implies $P(Q_i) = 0 = P_0 - b_i Q_i \implies b_1 Q_1 = P_0^2$. The comparative static on relative

²This expression also holds approximately if the marginal applicant paid an application fee of P_{app} , as long as $P_0 \gg P_{app}$

demand can be expressed as:

$$\theta_i = \frac{P_0 - c_i}{2b_i Q_i} = \frac{P_0 - c_i}{2P_0} = \frac{1}{2} - \frac{c_i}{2P_0} \quad (\text{A.66})$$

$$\implies \theta_1 - \theta_2 = -\frac{(c_1 - c_2)}{2P_0} \quad (\text{A.67})$$

$$\implies \frac{d\theta_1 - \theta_2}{dP_0} = \frac{(c_1 - c_2)}{2P_0^2} \quad (\text{A.68})$$

Comparative Static on Adjustment Factor

First, let us recall the definition of γ_i and express it as a function of P_0 and other parameters.

Then we will impose $\gamma_1 = \gamma_2 = \gamma$ to solve for the comparative static of the adjustment factor

$$\left(1 - \frac{\gamma}{2a^2}\right).$$

$$\gamma_i \equiv \lambda_i \theta_i^2 \quad (\text{A.69})$$

$$= \frac{2b_i r_i}{(P_0 - c_i)^2} \frac{(P_0 - c_i)^2}{(2Q_i b_i)^2} = \frac{\frac{r_i}{2}}{2b_i Q_i^2} \quad (\text{A.70})$$

$$= \frac{\frac{r_i}{2}}{Q_i P_0} \quad (\text{A.71})$$

$$= \frac{\frac{r_i b_i}{2}}{P_0^2} \quad (\text{A.72})$$

This last step comes from the assumption that the marginal applicant got a fee waiver and applied for free. There for at the total number of applications Q_i , the inverse demand equals zero, i.e. $P(Q_i) = 0 = P_0 - b_i Q_i \implies b_i Q_i = P_0$. Therefore, gamma is the ratio of the monetary equivalent of a school having an admissions rate that is $\frac{1}{2}$ that of its peer group to the total maximum total revenue if all students who applied were accepted and paid the same prices as the candidate with the highest willingness to pay. Since we assumed that $\gamma_1 = \gamma_2 = \gamma$, it follows from equation (A.72) that:

$$\frac{1}{2} r_1 b_1 = \frac{1}{2} r_2 b_2 \equiv R. \quad (\text{A.73})$$

and hence γ is inversely proportional to the square of P_0 , with R being the constant of proportionality:

$$\implies \gamma = \frac{R}{P_0^2} \quad (\text{A.74})$$

$$\implies \frac{d\gamma}{dP_0} = -\frac{2\gamma}{P_0} \quad (\text{A.75})$$

From this expression, we can compute the comparative static on the adjustment factor as a function of the parameters and the demand elasticity of the average admissions rate, ϵ_a ,

$$\frac{d}{dP_0} \left(\frac{\gamma}{2a^2} \right) = \frac{d\gamma}{dP_0} \left(\frac{1}{2a^2} \right) - \frac{\gamma}{a^3} \frac{da}{dP_0} \quad (\text{A.76})$$

$$= -\frac{2\gamma}{P_0} \left(\frac{1}{2a^2} \right) - \frac{\gamma}{a^3} \frac{da}{dP_0} \quad (\text{A.77})$$

$$= -\left(\frac{\gamma}{a^2 P_0} \right) \left(1 + \frac{P_0}{a} \frac{da}{dP_0} \right) \quad (\text{A.78})$$

$$\frac{d}{dP_0} \left(\frac{\gamma}{2a^2} \right) = -\left(\frac{\gamma}{a^2 P_0} \right) (1 + \epsilon_a) \quad (\text{A.79})$$

The comparative static on the adjustment factor is smaller the larger P_0 is and smaller the larger the average admissions rate is. If the demand elasticity is positive, the magnitude of the adjustment factor increases with the average admissions rate elasticity. If the demand elasticity is negative, as is the case of the low equilibrium, then the adjustment factor increases in magnitude only if the demand elasticity is smaller than -1.

Sign of Denominator

To sign the difference $\left(\frac{da_1}{dP_0} - \frac{da_2}{dP_0} \right)$, we label the college with the larger monopoly admission rate as College 1 and the college with the lower monopoly admissions rate as College 2, i.e. $\theta_1 > \theta_2$.

Claim: For the low equilibrium, $a, 0 > 2a^2 - \gamma$ and hence the denominator $\left(1 - \frac{\gamma}{2a^2} \right) < 0$.

Proof. Proof by contradiction. The admissions rate in the low equilibrium is given by $a_- = \frac{\theta_m}{2} \left(1 - \sqrt{1 - \frac{2\gamma}{\theta_m^2}}\right)$. We will show that if we assume $0 < 2a^2 - \gamma$, instead of $0 > 2a^2 - \gamma$, that this will violate the condition for multiple equilibria $1 - \frac{2\gamma}{\theta_m^2} > 0$, which we have also assumed to hold in order for the low admissions rate equilibrium to exist.

$$0 < 2a^2 - \gamma \tag{A.80}$$

$$\implies 0 < \frac{\theta_m^2}{2} \left(1 - 2\sqrt{1 - \frac{2\gamma}{\theta_m^2}} + 1 - \frac{2\gamma}{\theta_m^2}\right) - \gamma \tag{A.81}$$

$$\implies 0 < \theta_m^2 - 2\gamma - \theta_m^2 \sqrt{1 - \frac{2\gamma}{\theta_m^2}} \tag{A.82}$$

$$\implies \theta_m^4 \left(1 - \frac{2\gamma}{\theta_m^2}\right) < (\theta_m^2 - 2\gamma)^2 \tag{A.83}$$

$$\implies \theta_m^4 - 2\gamma\theta_m^2 < \theta_m^4 - 4\gamma\theta_m^2 + 4\gamma^2 \tag{A.84}$$

$$\implies 2\gamma\theta_m^2 \left(1 - \frac{2\gamma}{\theta_m^2}\right) < 0 \tag{A.85}$$

$$\implies \left(1 - \frac{2\gamma}{\theta_m^2}\right) < 0 \tag{A.86}$$

□

Claim: If in low equilibrium and $\theta_1 > \theta_2$, then $a_1 - a_2 < 0$ and the school with the higher monopoly admissions rate has the lower competitive admissions rate.

Proof. Solving the best response function we know:

$$a_i = \frac{\theta_i - \frac{\gamma}{a}}{1 - \frac{\gamma}{2a^2}} \tag{A.87}$$

$$\implies a_1 - a_2 = \frac{\theta_1 - \theta_2}{1 - \frac{\gamma}{2a^2}} \tag{A.88}$$

$$\implies a_1 - a_2 = -\frac{(\theta_1 - \theta_2)}{\left|1 - \frac{\gamma}{2a^2}\right|} \tag{A.89}$$

$$\implies a_1 - a_2 < 0 \tag{A.90}$$

□

A.3.1. Heterogeneous Responses to Demand Shocks

Claim: Given two colleges, 1 & 2, where College 2 has a lower monopoly admissions rate than College 1, i.e. $\theta_1 > \theta_2$, if the colleges coordinate on the low average admissions rate, \bar{a}_L , and if the elasticity of the average admissions rate is sufficiently negative, i.e. $\epsilon_d < -1 + \frac{\bar{a}_L^2}{\gamma} \left| \frac{\theta_1 - \theta_2}{a_1 - a_2} \right|$, then College 2 decreases its admissions, a_2 rate by more than College 1, a_1 when faced with a common demand shock.

Proof. In the low admissions equilibrium, $2\bar{a}_L^2 - \gamma < 0$, and $\theta_1 - \theta_2 > 0$, which together $\implies a_1 - a_2 < 0$. Going back to equation (??) we can simplify the expression for $\frac{da_1}{dP_0} - \frac{da_2}{dP_0} > 0$ using the results of equation (A.68) and (A.79):

$$\left(\frac{da_1}{dP_0} - \frac{da_2}{dP_0} \right) = \frac{\frac{(c_1 - c_2)}{2P_0^2} - (a_1 - a_2) \left(\frac{\gamma}{\bar{a}_L^2 P_0} \right) (1 + \epsilon_a)}{\left(1 - \frac{\gamma}{2\bar{a}_L^2} \right)} \quad (\text{A.91})$$

$$= - \left(\frac{\theta_1 - \theta_2}{P_0} \right) \frac{\left[1 + \left(\frac{a_1 - a_2}{\theta_1 - \theta_2} \right) \left(\frac{\gamma}{\bar{a}_L^2} \right) (1 + \epsilon_a) \right]}{\left(1 - \frac{\gamma}{2\bar{a}_L^2} \right)} \quad (\text{A.92})$$

$$= - \left(\frac{\theta_1 - \theta_2}{P_0} \right) \frac{\left[1 - \left| \frac{a_1 - a_2}{\theta_1 - \theta_2} \right| \left(\frac{\gamma}{\bar{a}_L^2} \right) (1 + \epsilon_a) \right]}{\left(1 - \frac{\gamma}{2\bar{a}_L^2} \right)} \quad (\text{A.93})$$

$$= - \left(\frac{\theta_1 - \theta_2}{P_0} \right) \frac{\left[2\bar{a}_L^2 - \gamma \left(2 \left| \frac{a_1 - a_2}{\theta_1 - \theta_2} \right| (1 + \epsilon_a) \right) \right]}{(2\bar{a}_L^2 - \gamma)} \quad (\text{A.94})$$

$$\left(\frac{da_1}{dP_0} - \frac{da_2}{dP_0} \right) = \underbrace{- \frac{1}{2\bar{a}_L^2 - \gamma} \left(\frac{\theta_1 - \theta_2}{P_0} \right) \left(2\bar{a}_L^2 - \gamma \left[2 \left| \frac{a_1 - a_2}{\theta_1 - \theta_2} \right| (1 + \epsilon_a) \right] \right)}_{> 0} \quad (\text{A.95})$$

$$\implies \frac{da_1}{dP_0} - \frac{da_2}{dP_0} > 0 \quad (\text{A.96})$$

$$\implies 0 < 2\bar{a}_L^2 - \gamma \left[2 \left| \frac{a_1 - a_2}{\theta_1 - \theta_2} \right| (1 + \epsilon_a) \right] \quad (\text{A.97})$$

$$\implies \epsilon_a < -1 + \frac{\bar{a}_L^2}{\gamma} \left| \frac{\theta_1 - \theta_2}{a_1 - a_2} \right| \quad (\text{A.98})$$

□

A.4. General N Heterogeneous Case

We also have analytically tractable results for the case of N colleges, where there are $\frac{N}{2}$ type 1 colleges (θ_1) and $\frac{N}{2}$ low type 2 colleges (θ_2). While $\theta_1 > \theta_2$, we still maintain $\gamma_1 = \gamma_2 = \gamma$, to preserve analytic tractability while allowing for heterogeneity.

Claim: For this case of N colleges of two types, there exist multiple equilibria if $\frac{4\gamma}{\theta_m^2} \left(\frac{N-1}{N}\right) < 1$, where $\theta_m = \frac{1}{2}(\theta_1 + \theta_2)$.

Proof. By symmetry, the type 1 colleges will follow an identical action a_1 and the type 2 colleges will follow an identical a_2 action. The best response functions evaluated at a_1 and a_1 give a relationship between these optimal actions and the average admissions rate $\bar{a} = \frac{1}{2}(a_1 + a_2)$:

$$a_2 = \theta_2 - \frac{\gamma}{\bar{a}^2} \frac{1}{N} \left[\left(\frac{N}{2} - 1\right) a_2 + \left(\frac{N}{2}\right) a_1 \right] \quad (\text{A.99})$$

$$\implies a_2 = \theta_2 - \frac{\gamma}{\bar{a}} + \frac{1}{N} \frac{\gamma}{\bar{a}^2} a_2 \quad (\text{A.100})$$

$$a_1 = \theta_1 - \frac{\gamma}{\bar{a}^2} \frac{1}{N} \left[\left(\frac{N}{2} - 1\right) a_1 + \left(\frac{N}{2}\right) a_2 \right] \quad (\text{A.101})$$

$$\implies a_1 = \theta_1 - \frac{\gamma}{\bar{a}} + \frac{1}{N} \frac{\gamma}{\bar{a}^2} a_1 \quad (\text{A.102})$$

Summing the two best response functions in equation (A.100) and (A.102):

$$(a_1 + a_2) = (\theta_1 + \theta_2) - 2\frac{\gamma}{\bar{a}} + \frac{1}{N} \frac{\gamma}{\bar{a}^2} (a_1 + a_2) \quad (\text{A.103})$$

$$\implies \bar{a} = \theta_m - \frac{\gamma}{\bar{a}} + \frac{\gamma}{N\bar{a}} \quad (\text{A.104})$$

$$\implies \bar{a}^2 - \theta_m \bar{a} + \gamma \left(\frac{N-1}{N}\right) \quad (\text{A.105})$$

In equilibrium, the average admissions, \bar{a} , rate takes on one of two values; a high value \bar{a}_H

and a lower value \bar{a}_L :

$$\bar{a}_H = \frac{\theta_m}{2} \left(1 + \sqrt{1 - \frac{4\gamma}{\theta_m^2} \left(\frac{N-1}{N} \right)} \right) \quad (\text{A.106})$$

$$\bar{a}_L = \frac{\theta_m}{2} \left(1 - \sqrt{1 - \frac{4\gamma}{\theta_m^2} \left(\frac{N-1}{N} \right)} \right) \quad (\text{A.107})$$

In the high average admissions rate equilibrium, \bar{a}_H , the two colleges have equilibrium actions (a_{1H}, a_{2H}) given by:

$$(a_{1H}, a_{2H}) = \left(\frac{\theta_1 - \frac{\gamma}{\bar{a}_H}}{1 - \frac{\gamma}{N\bar{a}_H^2}}, \frac{\theta_2 - \frac{\gamma}{\bar{a}_H}}{1 - \frac{\gamma}{N\bar{a}_H^2}} \right) \quad (\text{A.108})$$

In the low average admissions rate equilibrium, \bar{a}_L , the two colleges have equilibrium actions (a_{1L}, a_{2L}) given by:

$$(a_{1L}, a_{2L}) = \left(\frac{\theta_1 - \frac{\gamma}{\bar{a}_L}}{1 - \frac{\gamma}{N\bar{a}_L^2}}, \frac{\theta_2 - \frac{\gamma}{\bar{a}_L}}{1 - \frac{\gamma}{N\bar{a}_L^2}} \right) \quad (\text{A.109})$$

□

Country Comparison

<u>BAHAMAS:</u>	<u>BARBADOS:</u>
<ul style="list-style-type: none"><input type="checkbox"/> Colonial Origin: British<input type="checkbox"/> Legal System: British Common Law<input type="checkbox"/> Population: 309,156<input type="checkbox"/> Area (Main Island): 207 km²<input type="checkbox"/> Per capita GDP<ul style="list-style-type: none">✓ \$3,330.79 (1976)<input type="checkbox"/> GDP composition by sector:<ul style="list-style-type: none">✓ agriculture: 3%✓ industry: 7%✓ services: 90%<input type="checkbox"/> Property Rights<ul style="list-style-type: none"><input type="checkbox"/> Natives: protected<input type="checkbox"/> Foreigners: protected, except during 1981-1993: "Immovable Property Act" (IPA)	<ul style="list-style-type: none">❖ Colonial Origin: British❖ Legal System: British Common Law❖ Population: 284,589❖ Area : 433 km²❖ Per capita GDP<ul style="list-style-type: none">➢ \$1,767.94 (1976)❖ GDP composition by sector:<ul style="list-style-type: none">➢ agriculture: 6%➢ industry: 16%➢ services: 78%❖ Property Rights<ul style="list-style-type: none">➢ Natives: protected➢ Foreigners: protected

Figure 11: Country Comparison.

Table 16: **Effect of Law Change on GDP Growth**

	(1)	(2)	(3)	(4)
IPA81	0.2790 (0.0502)	0.1528 (0.0424)	0.1499 (0.0423)	0.1531 (0.0429)
BAH	0.4889 (0.0395)	0.5270 (0.0464)	0.5121 (0.0472)	0.5284 (0.0481)
BAH x TIME		-0.0013 (0.0033)	0.0000 (0.0034)	-0.0015 (0.0035)
NFDI		0.0001 (0.0001)	0.0000 (0.0001)	0.0001 (0.0001)
IPA81 x NFDI		-0.0006 (0.0003)	-0.0005 (0.0003)	-0.0006 (0.0003)
NFDI LAG				
log(CPI)		1.0360 (0.1118)	1.0460 (0.1115)	1.0380 (0.1142)
GOV			-0.0392 (0.0440)	
GOV LAG			-0.0620 (0.0441)	
HURR				0.0044 (0.0360)
TIME	0.0547 (0.0020)	0.0111 (0.0051)	0.0105 (0.0051)	0.0110 (0.0052)
Const.	7.8557 (0.0378)	4.0840 (0.4082)	4.0570 (0.4067)	4.0750 (0.4184)
Adjusted R	0.9468	0.9794	0.9796	0.9790
F-stat	350.80	402.60	316.20	345.60

Table 17: Using Legal Reform as Instrument for Net FDI

	Excl. Restriction	First Stage	Second Stage
IPA81	-145.82 (89.92)	0.143 (0.076)	0.223 (0.058)
BAH	41.27 (60.64)	0.554 (0.049)	0.526 (0.046)
TIME	7.07 (6.61)	0.008 (0.005)	0.006 (0.006)
IPLHA93	-38.42 (140.17)	-0.047 (0.123)	
TIMExIPLHA93	42.64 (13.02)	0.019 (0.013)	
log(CPI)	-99.02 (148.57)	1.109 (0.118)	1.135 (0.127)
BAHxTIME	9.02 (8.56)	-0.004 (0.008)	-0.007 (0.006)
NFDI		0.000 (0.000)	
IPA81xNFDI		0.000 (0.000)	
$NFDI_{2SLS}$			0.000 (0.000)
IPA81x $NFDI_{2SLS}$			-0.001 (0.001)
intercept		3.82 (0.43)	3.72 (0.46)
Adjusted R	0.8306	0.98	0.9809
F-stat	42.32	321.4	432.8

Table 18: **Robustness Checks, Counter-factuals, and Placebo Tests**

	Bahamas Only	Bahamas No IPA	Barbados IPA
IPA81	0.1650 (0.0534)		0.0097 (0.0460)
NFDI	0.0000 (0.0001)	-0.0001 (0.0001)	0.0001 (0.0003)
TIME	0.0172 (0.0145)	0.0061 (0.0149)	0.0107 (0.0084)
log(CPI)	0.8807 (0.2929)	1.1835 (0.2851)	1.0440 (0.1784)
IPA81xNFDI	-0.0005 (0.0004)		
Intercept	5.1680 (1.0560)	4.1303 (1.0365)	4.0460 (0.6284)
Adjusted R	0.9617	0.9488	0.9780
F-stat	146.7000	180.3000	322.6000

Table 19: **Robustness Checks: Auto-correlation**

	1	2	3
Residual (t-1)	0.43 (0.18)	0.15 (0.18)	-0.17 (0.19)
Residual (t-2)		0.03 (0.15)	0.15 (0.16)
Residual (t-3)		-0.42 (0.15)	-0.42 (0.13)
Residual (t-4)		-0.35 (0.17)	-0.52 (0.16)
Residual (t-5)			-0.25 (0.16)
Adjusted R	0.15	0.37	0.51
F-stat	6.07	4.68	6.05

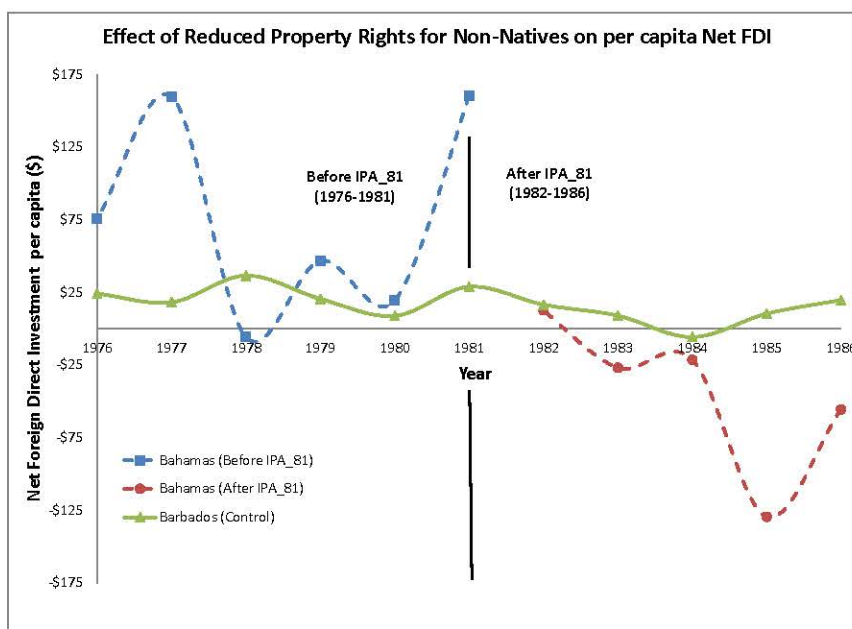


Figure 12: Effect of Reduced Property Rights for Non-Natives on per capita Net FDI.

Table 20: Contemporaneous Difference Estimates

	Contemp. Diff	Cochrane Orcutt
IPA81	0.030 (0.011)	0.023 (0.014)
TIME	-0.001 (0.001)	-0.001 (0.001)
log(CPI)	0.085 (0.111)	0.137 (0.176)
Intercept	0.266 (0.014)	0.153 (0.011)
Adjusted R	0.443	0.250
F-stat	8.678	4.177

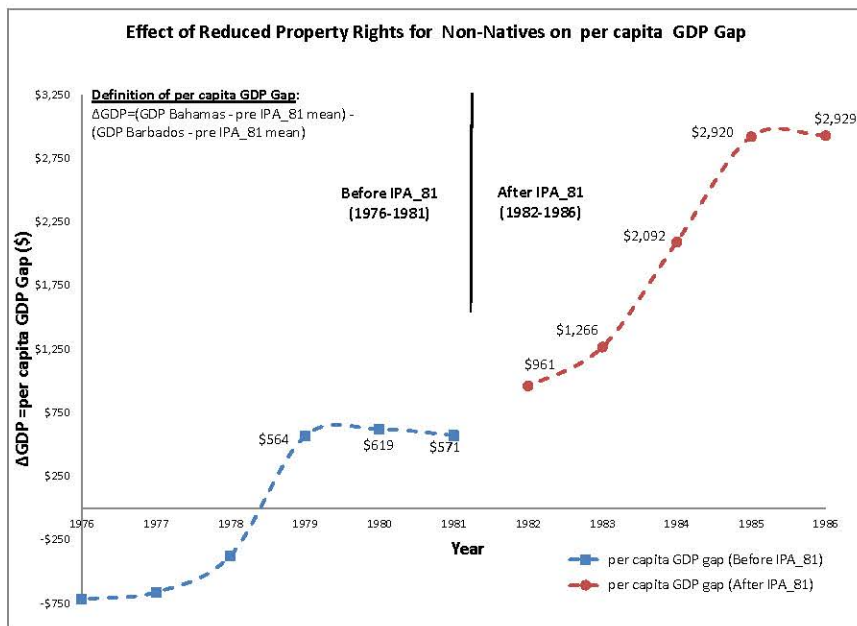


Figure 13: Effect of Reduced Property Rights for Non-Natives on per capita GDP Gap.

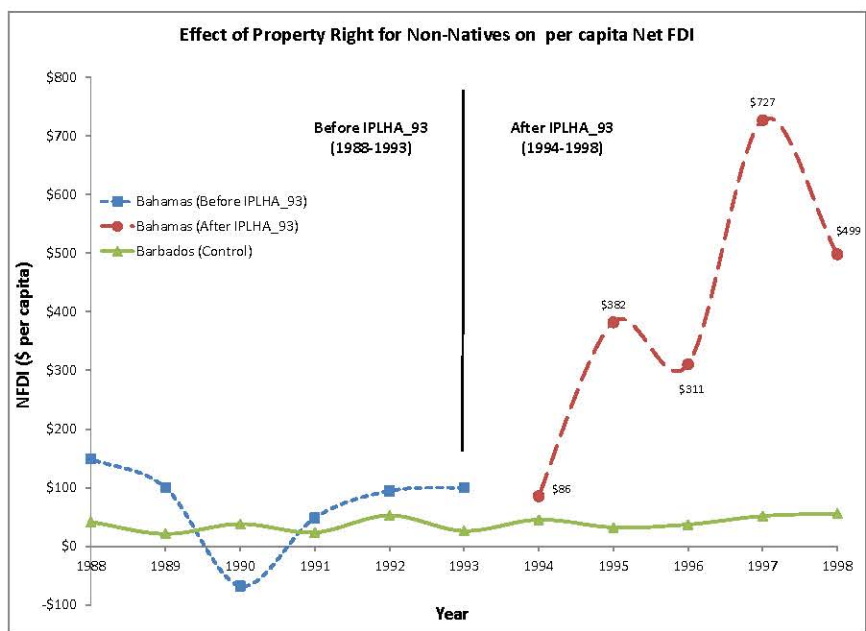


Figure 14: Effect of Property Rights for Non-Natives on per capita Net FDI.

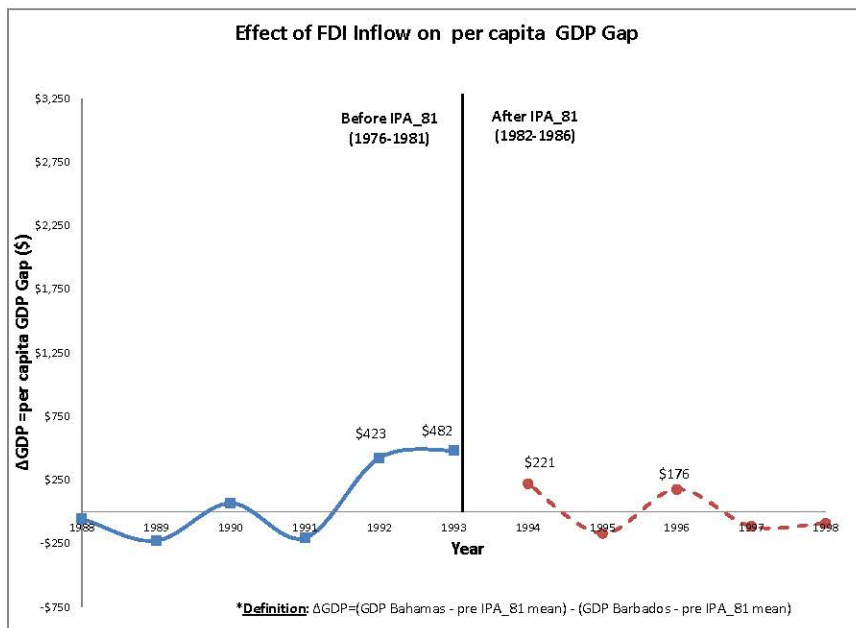


Figure 15: Effect of FDI Inflow on per capita GDP Gap.

BIBLIOGRAPHY

1. Acemoglu, Daron, Simon Johnson and James A. Robinson. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics* 107 (2002): 1231-1294.
2. Acemoglu, Daron, Simon Johnson and James A. Robinson. "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review* 91 (2001): 1369-1401.
3. Ahlfelds, Redding, Sturm and Wolf. "The Economics of Density: Evidence from the Berlin Wall." *Working Paper* (2010).
4. Anderson, Simon, Andre de Palma, and Jacques-Francois. "Discrete Choice Theory of Product Differentiation." PhD diss., MIT, 1992.
5. Anwar, Shamena, and Hanming Fang. "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence." *American Economic Review* 96 (2006): 127-151.
6. Avery, Chris, Jonathan Levin. "Early Admissions at Selective Colleges." *American Economic Review* 100: 2125-2156.
7. Banerjee, Abhijit and Lakshmi Iyer. "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India." *American Economic Review, American Economic Association* 95 (2005): 1190-1213.
8. Baum, S., Ma, J., and Payea, K. "Education Pays: The Benefits of Higher Education for Individuals and Society." *The College Board* (2010).
9. Bayer, Patrick, Robert McMillan, and Fernando Ferriera. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." *Journal of Political Economy* 115 (2007): 588-638.
10. Bayer, Patrick, and Christopher Timmins. "Estimating Equilibrium Models of Sorting Across Locations." *Economic Journal* 117 (2007): 353-374.
11. Bayer, Patrick, Hanming Fang, and Robert McMillan. "Separate When Equal? Racial Inequality and Residential Segregation." *Journal of Urban Economics* 82 (2014): 32-48.
12. Bayer, Patrick, Robert McMillan, and Kim S. Rueben. "Residential Segregation in General Equilibrium." *NBER Working Paper No. 11095* (2005).
13. Becker, Gary S. "Nobel Lecture: The Economic Way of Looking at Behavior." *Journal of Political Economy* 101 (1993): 385-409.

14. Becker, Gary, and Kevin M. Murphy. "Social Economics: Market Behavior in a Social Environment." Harvard University, 2000.
15. Berry, Steven, James Levinsohn, and Ariel Pakes. "Automobile Prices in Market Equilibrium." *Econometrica* 63 (1995): 841-890.
16. Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics* 119, n.1 (2004): 249-275.
17. Black, Sandra. "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics* 114 (1999): 577-599.
18. Blackburn, M.L. and D. Neumark. "Omitted Ability Bias and the Increase in the Return to Schooling." *Journal of Labor Economics* 11, No. 3 (1993).
19. Blume, Lawrence. "The Statistical Mechanics of Strategic Interaction." *Games and Economic Behavior* 5 (1993): 387-424.
20. Boustan, Leah. "Was Postwar Suburbanization 'White Flight' ? Evidence from the Black Migration." *Quarterly Journal of Economics* 125 (2010): 417-443.
21. Brock, William, and Steven Durlauf. 2001. "Discrete Choice with Social Interactions." *Review of Economic Studies* 68 (2001): 235-260.
22. Caetano, Gregorio, and Vikram Maheshri. "School Segregation and the Identification of Tipping Behavior." *Working Paper* (2013).
23. Card, David and Alan Krueger, "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania." *American Economic Review* 84, n. 4 (1994), 774-775.
24. Card, David, Alexandre Mas, and Jesse Rothstein. "Tipping and the Dynamics of Segregation." *Quarterly Journal of Economics* 123 (2008): 177-218.
25. Casey, Marcus. "Price Impacts of Black Entry: Evidence from Transactions Data." *Working Paper*.
26. Chade, Hector, Gregory Lewis, Lones Smith. "Student Portfolios and the College Admissions Problem." *Review of Economic Studies* 8 (2014): 971-1002.
27. Clairvaux, Bernard. *On Loving God*. Cistercian: 1995.
28. Cochrane, D and G. H. Orcutt. "Application of Least Squares Regression to Relationships Containing Auto-Correlated Error Terms." *Journal of the American Statistical Association* 44 (1949), 32-61.

29. Conlon, Christopher T., and Julie Holland Mortimer. "Demand Estimation Under Incomplete Product Availability." *American Economic Journal* 5 (2013): 130.
30. Dale, Stacey, Krueger, Allen. "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables." *Quarterly Journal of Economics* 117 (2002): 1491-1528.
31. Dale, Stacey, Krueger, Allen. "Estimating the Return to College Selectivity Over the Career Using Administrative Earnings Data." *Princeton University Working Paper* (2011).
32. Dell, Melissa. "The Persistent Effects of Peru's Mining Mita." *Econometrica* 78 (2008): 1863-1903.
33. Diamond, Jared and James Robinson. "Natural Experiments of History." Harvard University, 2010.
34. Easterly, William. "Empirics of Strategic Interdependence: The Case of the Racial Tipping Point." *NBER Working Paper No. 15069* (2009).
35. Field, Erica. "Entitled to Work: Urban Property Rights and Labor Supply in Peru." *Quarterly Journal of Economics* 122 (2007): 1561-1602.
36. Chao, Fu. "Equilibrium Tuition, Applications, Admissions and Enrollment in the College Market." *Journal of Political Economy* 122 (2014): 225-281.
37. Glaeser, Edward L.. "The Political Economy of Hatred." *The Quarterly Journal of Economics* 120 (2005): 45-86.
38. Glaeser, Edward L., Bruce Sacerdote, and Jose A. Scheinkman. 1996. "Crime and Social Interactions." *Quarterly Journal of Economics* 111 (1996): 507-548.
39. Glaeser, Edward, Rafael La Porta, Florencio Lopez de Silanes and Andrei Shleifer. "Do Institutions Cause Growth?" *Journal of Economic Growth* 9 (2004): 271- 303.
40. Grodzins, Morton. "The Metropolitan Area as a Racial Problem." University of Pittsburgh, 1958.
41. Hennessy, John. "Should Stanford Expand the Freshman Class?" *Stanford Alumni Magazine* (2007).
42. Henry, Peter and Conrad Miller. "Institutions versus Policies: A Tale of Two Islands." *American Economic Review, American Economic Association* 99 (2010): 261-67.
43. Hornbeck, Richard. "Barbed Wire: Property Rights and Agricultural Development." *Quarterly Journal of Economics* 125, No. 2 (2010): 767-810.

44. Hoxby, Caroline, "The Return to Attending a More Selective College." *Unpublished Manuscript*.
45. Hoxby, Caroline, "The Changing Selectivity of American Colleges." *Journal of Economic Perspectives* (2009).
46. Jullien, Bruno. "Participation Constraints in Adverse Selection Models." *Journal of Economic Theory* 93 (2000): 147.
47. Katz, L. and K. Murphy. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *Quarterly Journal of Economics* 107, No. 1 (1992): 35-78.
48. Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence Katz. "Experimental Analysis of Neighborhood Effects." *Econometrica* 76 (2007): 83-119.
49. Knowles, John, Nicola Persico, and Petra Todd. 2001. "Racial Bias in Motor Vehicle Searches: Theory and Evidence." *Journal of Political Economy* 109 (2001): 203-229.
50. Kreps, David and Robert Wilson. "Reputation and Imperfect Information." *Journal of Economic Theory* 27, n. 2 (1982): 253-279.
51. La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer, 2008. "The Economic Consequences of Legal Origins." *Journal of Economic Literature*, American Economic Association, vol. 46(2), pages 285-332, June.
52. Long, M.C.. "College Quality and Early Adult Outcomes." *Economics of Education Review* 27 (2007): 588.
53. MacKinlay, Craig A. "Event Studies in Economics and Finance." *Journal of Economic Literature* 35, n. 1 (1997): 13-39.
54. McCall, John J.. "Economics of Information and Job Search." *Quarterly Journal of Economics* 84 (1970): 113-126.
55. McFadden, Daniel. "Modeling the Choice of Residential Location." (1978): 75-96.
56. Mortensen, Dale, and Christopher Pissarides. "Job Creation and Job Destruction in the Theory of Unemployment." *Review of Economic Studies* 61 (1994): 397-415.
57. Reardon, Sean, Rachel Baker and Daniel Klasik. "Race, income, and enrollment patterns in highly selective colleges, 1982-2004 (Unpublished manuscript)."
58. Schelling, Thomas. "Models of Segregation." *American Economic Review Papers and Proceedings* 59 (1969): 488-493.
59. Schelling, Thomas. "Dynamic Models of Segregation." *Journal of Mathematical Sociology* 1 (1971): 143-186.