



Center for Population Dynamics

School of Social and Family Dynamics
Arizona State University

Tempe, Arizona, 85287-3701, USA www.asu.edu/cepod

CePoD Working Paper # 09-114

Can't Get There From Here: Investigating The Intersection of Race, Class, Sport, and Social Mobility with Newspaper Data

Joshua Kjerulf Dubrow

jimi adams

2009

Can't Get There From Here: Investigating The Intersection of Race, Class, Sport, and Social Mobility with Newspaper Data*

Joshua Kjerulf Dubrow
Polish Academy of Sciences

jimi adams
Arizona State University

ABSTRACT

The purpose of this study is to empirically show that race, social class and family structure origins, and their intersection especially, influence membership in the National Basketball Association (NBA), one of the greatest achievements in professional sports. Drawing on classic stratification research, we expect that African-American players from disadvantaged social origins have lower chances to be in the NBA than similarly disadvantaged white players, or than African American and white players from advantaged social backgrounds. Information on race, and on social class and family structure background for a subpopulation of NBA players (N = 155) comes from 245 articles published in local, regional and/or national newspapers between 1994 and 2004. Findings support our research hypotheses: The basic relationship between social origin and destination remains unchanged after accounting for selection bias, a methodological problem common in newspaper data.

* A version of this article was presented in 2004 under the title, "Media Portrayals of Social Class Origin and Family Background of the Modern Professional Basketball Player," at the North Central Sociological Association annual meeting in Cleveland, OH. The authors gratefully acknowledge the helpful comments made by the participants of that session, as well as by Ben Cornwell, Tim Curry, Alex Heckert, Steven Ortiz, Dean Purdy, Stephen K. Sanderson, Kazimierz M. Slomczynski, and Irina Tomescu-Dubrow. All mistakes herein are the sole property of the authors. The research leading to this article was conducted at the Department of Sociology, The Ohio State University. Please direct all correspondence to Joshua Kjerulf Dubrow, Institute of Philosophy and Sociology, Polish Academy of Sciences, Room 211, 72 Nowy Swiat, 00-330 Warsaw, Poland, dubrow.2@osu.edu.

Introduction

The purpose of this paper is to empirically demonstrate that the intersection of race and, social class and family structure background influences basketball players' odds of attaining a position in the National Basketball Association (NBA), an achievement that can be considered "the ultimate in sports status attainment" (Leonard 1996). We build on the theoretical assertion that class and family background, in conjunction with race, represent a key component in the sports-upward mobility relation. On a unique dataset constructed by coding 245 articles published in local, regional and/or national newspapers in the period 1994-2004, we examine whether African American players from disadvantaged social origins have lower odds of being in the NBA than (a) white players of similar disadvantage, and (b) African-American and white players from relatively well-off families.

Problem Statement

The relationship between race and social mobility is central to research in sports and society, and is often discussed of in terms of odds of achievement. According to one theoretical approach, sport provides historically marginalized groups, racial ones included, a pathway for upward mobility via college scholarships and psychological intangibles that can be translated into success in the labor market (Eide and Ronan 2001). This perspective is heavily criticized by researchers who contend that "the notion that sport is a readily accessible avenue to upward social mobility is a myth that continues to fuel the American Dream" (Washington and Karen 2001: 189). In their view, racial minorities in the US, African Americans especially, are locked into lower social positions, a problem for which involvement into professional sports does not offer a straight-forward solution, as very low odds of attainment supposedly indicate.

What is common to these otherwise divergent views on the role of sport for social success is the limited attention major stratification variables, social class of origin and family structure background especially, receive (but see Washington and Karen 2001). The omission is particularly troublesome as both types of arguments employ odds to sustain their positions, yet odds estimated in this manner will not account for differences in social class background and in family structure background *across* groups, and especially *within* groups.

Failure to consider how within-race heterogeneity influences players' chances of success occurs frequently both in academic and non-academic reports. For example, Eitzen (1999) in "Upward Mobility through Sport? The Myths and Realities" claims that "...while the odds of African American males making it as professional athletes are more favorable than is the case for whites (about 1 in 3,500 African American male high school athletes, compared to 1 in 10,000 white male high school athletes) these odds remain slim. Of the 40,000 or so African Americans boys who play high school basketball, only 35 will make the NBA and only 7 will be starters." The National College Athletic Association (NCAA) puts the odds of high school players being drafted by an NBA team at 3 in 10,000 (NCAA website). Taking race into account, the odds are reported to be as low as .000002 for whites and .000006 for Blacks (Leonard 1996: 296).

Few have directly tested the claim that most professional basketball players grew up in inner city ghettos and other impoverished social environments. A recent ESPN Magazine report (Craggs 2008) used U.S. Census 2000 data to examine the hometowns -- their measure of social origins -- of 158 American-born players from the NBA drafts between 1998 and 2008. They found that the majority of these players are from middle class hometowns: "The median household income of our draftees' hometowns was \$38,127, which tracks closely with the national average of \$41,994." Underscoring this finding is that the NBA draftees hailed from hometowns that had about the same percentage of people who graduated from high

school as the national average. While a rather insensitive measure of social origins, it does suggest further inquiry into the class and family structure background of NBA players (Dubrow 2002).

Due to the understudied problem of the intersection of main stratification variables, the basic question of how social origin – expressed through race, as well as through advantaged vs. disadvantaged class and family structure background -- influences professional athlete status attainment still needs examination. In this paper we draw on classic social stratification research and sports literature, and a unique dataset from newspaper articles published between 1994 and 2004, to examine the relationship between basketball players' race and social origin on one hand, and their likelihood of being in the NBA, on the other.

We proceed as follows: first, we briefly discuss the theoretical framework underlying this study, and present a set of statistically testable hypotheses about the role of race, class, and family structure background for professional basketball athlete status attainment. We then discuss the data and methods that allow us to test the posited relationships. The third section of the paper presents the empirical results of our analyses. We conclude with a discussion of the implications of our findings for theory and policy.

Theoretical Background and Research Hypotheses

American studies in sport and mobility devote great attention to the relation between race and athletic involvement, particularly so within professional basketball, given African Americans' overwhelming presence at all levels of competition, and the NBA level especially. Some researchers argue that sport provides historically marginalized groups a pathway to upward mobility via college scholarships and psychological intangibles that can be translated into success in the labor market; empirical findings suggest that sport participation has a positive effect on the educational attainment and earnings of African American males (Eide and Ronan

2001). Studies also indicate that the pathway to fame and fortune, especially with regards to the NBA, goes through “big-time” athletic programs where graduation rates for college basketball athletes are lower than the rest of the university population (Eitzen 1996: 101).

To explain why race matters for athletic performance, the current literature focuses on debunking genetic theories, and positing theories of the social environment (Coakley 2004: 292-97; Dubrow 2002; Edwards 1973: Chapter 7; Hartmann 2000; Ogden and Hilt 2003). Genetic arguments stipulate dominance as a result of realizing genetic predispositions in athletic prowess (Entine 2000; Goldberg 1991). In this theory, choice of a sports career “is affected by the reality of one’s basic fitness for a sport or event,” assuming race-based differences in athletic task performance capability (Goldberg 1991: 124). In contrast, social environment theorists argue that this form of dominance is rooted in social facts, including differential labor market access and cultural orientations. In positing race as influencing the odds of attainment, social environment theorists draw on the vast literature of the social implications of physiognomic traits (for a review, see Winant 2000). It is not clear exactly how the process of racial concentration into particular sports occurs, though school experiences with physical education gatekeepers is a plausible mechanism (Harrison Jr. et al. 2004). Others debate the intersection of race with micro and macro level social environment mechanisms, such as role model behavior, cultural proclivities, media influences, and limited occupational opportunities (Ogden and Hilt 2003; Edwards 1973; Sellers and Kuperminc 1997; Pascarella and Smart 1991; Sailes 1998; Harris 1998; Spreitzer 1994; Harrison Jr. et al. 2004). In sum, there is little disagreement that race affects sport mobility; what needs further exploration are the means by which this happens.

In contrast to the significant work on the role of race, empirical studies on the influence of social class and family structure background on professional sport attainment have received less attention (Washington and Karen 2001; Spreitzer 1994). Social mobility

research consistently shows that individuals' social origin, measured in terms of social class affects up-ward mobility: other things considered, people from disadvantaged social background fare less well in the stratification hierarchy than those of privileged social origin.¹ Sport mobility research posits a likely class-based mechanism for professional athlete status attainment. Measuring disadvantage in terms of family income, Spreitzer (1994) argues that freshman high school athletes from disadvantaged backgrounds were less likely to continue school athletics in their sophomore year and were more likely to have dropped out of athletics by their senior year (368). Since high school is a necessary step to playing in college or, more recently, a leap into the professional leagues, disadvantaged background should play a substantial role in influencing the odds of becoming a professional athlete.

Theories of the social environment prompt us to consider -- next to race and social class -- the role of players' family structure background. Braddock et al (1991) find that African American male 8th graders from two-parent homes are more likely to participate in interscholastic and intramural sports than those from single parent homes (119), which may indicate that family structure disadvantage influences the pathway to professional athlete status. While in statistical sense non-two-parent status determines later achievements, it is important to keep in mind that social class origin strongly influences the family effect (Milne et al. 1986; McLanahan and Sandefur 1994).

The above-mentioned theoretical considerations point to the complex picture of relative disadvantage that the intersection of race with class and family structure background can produce in the American stratification system. For example, whether measured as lower social class position or growing up in a single parent household, African Americans are more likely to come from disadvantaged backgrounds than whites. Race is generally considered to

¹ Both social stratification and mobility research acknowledge the importance of social origin for social destination (i.e. where people end up in the stratification system). Disagreements within the field pertain to measurement issues (whether to measure social origin and destination in terms of social status or social class), and to the magnitude of the role of social origin. For a basic review, see Kerbo 2003.

be a valid indicator of disadvantage (economic, but not only), but one should remember that class and family structure background differentiation *within* racial categories also exists. To date however, very few studies examine the issue of intersectionality in relation to sport participation in general, and especially to professional athlete status attainment, mainly because adequate data are lacking.²

In this paper we rely on primary data to analyze the relation of *within-race* heterogeneity in terms of social class and family structure background, to sport mobility. Our main research hypothesis states that *race – in conjunction with social class and family structure background - influences the odds of attaining the status of professional basketball players*, in this case, membership in the NBA. This translates into the following expectations:

a) Due to the relatively high level of resources and social capital required to achieve a professional athletics career, *African-American players from disadvantaged social class should be less likely to become professional basketball athletes* than white players of similar disadvantage, or African American and white players from relatively well-off families. In other words, they should have lower odds of being in the NBA.

b) The same mechanism leads us to expect that *African Americans from disadvantaged family background are less likely to populate professional basketball leagues*.

² Currently, no large N data exist that have professional athletes as the units of analysis and include basic variables of race, class, and family structure background. This limits professional athlete status attainment studies to quantitative analyses of macro-level data that cannot account for the intersection of race, class, and family structure (e.g. Eitzen 1999; Leonard 1996; NCAA website), and small N qualitative and case studies that account for this intersectionality but are not generalizable (e.g. the movie *Hoop Dreams* 1994).

c) Given our argument that non-white race and disadvantaged social origin intersect, we expect *African Americans from disadvantaged class and family backgrounds to be the least likely to be present in the NBA.*

In relation to this last hypothesis, however, it is important to note the following: while in statistical sense non-two-parent status determines later achievements, social class origin strongly influences the family effect (Milne et al. 1986; McLanahan and Sandefur 1994). Given this causal chain, we argue that even if the majority of professional basketball players were raised in non-two-parent families, most should come from the middle class and above. In consequence, among professional basketball players *compounded disadvantage, i.e. both low social class and non-two parent family structure background, should be rare, yet more common among African-Americans than among whites.*

Data and Methods

For our analyses, we follow data collection techniques researchers of political phenomena have long employed (see Earl et al 2004; Oliver and Meyers 1999; Woolley 2000), and construct a unique dataset on professional basketball players' social origin through newspaper coding. Specifically, 245 articles published in local, regional and/or national papers between 1994 and 2004 provide biographical information on a subpopulation of 155 National Basketball Association (NBA) players. Lexis-Nexis, an electronic database of newspaper articles deposited by participating newspaper organizations, served as our bibliographic search engine, or on-line index (see Appendix C for the detailed description of our search process).

Our data collection strategy is closest to what Woolley (2000: 157) describes as “record-coding”: coding to create an event record, or “count.” In practice, instead of

gathering every article that referred to social origin for a particular player, we stopped searching for news articles if (a) the same information was repeated in multiple articles and (b) after exhausting all possible word searches. Readers will note that we explicitly sample based on characteristics of the dependent variable, that is, membership in the NBA. We discuss the limitations this imposes on our causal model and the possibilities for constructing valid comparison groups in Appendix B.

Due to the lack of quantitative data containing mobility variables for professional athletes, primarily because professional athletes in sports such as basketball, football, baseball and hockey, similar to elites in other sectors of society (e.g. CEO's and politicians), are not acquiescent to traditional survey research methods (e.g. face to face interviews and mail-out questionnaires), using primary data from secondary sources to address our research question is crucial.³

Sampling Frame

Our sampling frame comes from a complete listing of NBA players who were in the league and appeared on the official NBA website as of July 15, 2002.⁴ Of the total 475 NBA players listed, 465 had information that pertained to birth date and race. From this number, we selected those African American and white players who were born in or after 1977 in the

³On this score, elite studies are telling: despite the resources and best efforts in elite populations research, survey response rates are typically between 30 and 40 percent (for statistics on response rates where top managers are the target population, see Baruch 1999: 431; for a discussion on problems with securing interviews with politicians, see Maisel and Stone 1998).

⁴ Since July represents the period after the NBA draft and before training camp, it allows for the maximum number of potential NBA players, including those who would eventually be cut from their team before the first official league game. Thus, one could argue that a percentage of our players may not have been in the league long enough for reporters to write human interest stories about them. This possible newspaper based selection mechanism is not likely to bias our results, as our pool of articles spans 1994-2004 including those written about the player when they were in college, when they were selected by the NBA team, or when they reappeared in the league before the end of 2004.

United States and Canada (N = 155).⁵ We chose 1977 as the boundary for year of birth to (a) focus our attention on a manageable subpopulation and (b) control for (to the extent possible) selection bias based on changing newspaper reporting styles over time. We eliminated players who were born outside of the U.S. and Canada because of possible country-level variations in the differences in, and effects of, class and family-structure background that would be impossible to explain with the limited number of cases. The number of Canadians in our subpopulation is two.

Race is a crucial variable and was difficult, if often times impossible, to determine from reading players' biographies in the newspaper. Following the strategy likely employed by Murrell and Edward (1994), we determined race by examining available pictures of the athletes; pictures come from the official NBA website.⁶ The racial composition of our subpopulation is roughly equivalent to the total NBA population in 2002: eighty-six percent of our sample are African American and 14 percent are white, compared with 77 percent African American and 21 percent white for the entire league. Overall, our percentages are substantively similar to those reported in other sources (e.g. Coakley 2004: 318), indicating that our method of identifying players' racial background is reliable.⁷

We limit our population to NBA players. Thus, our research question should read:
Among those who have attempted to become NBA players, did the intersection of race, class, and family structure background influence their odds of attainment?

⁵Most data on birthplaces come from the official NBA website; where the NBA website did not list the birthplace, we gathered the information from newspaper articles.

⁶ Murrell and Edward (1994) did not specify how they coded race. We examined as many of the most recent studies that we could find that used newspaper data with race as a variable and where professional athletes were the units of analysis; none posited an alternative procedure to ours in determining the race of an athlete.

⁷ A substantial proportion of white NBA players come from outside the U.S. and Canada; thus the differences between our subpopulation and that reported in other sources likely stems from the case-selection logic we employed.

Coding

Classic sociological definitions of social class and family structure background guided our coding process. To identify social class background, we examined words and phrases indicating parental occupation, description of neighborhood, and general statements of class position. Our definition of social class follows that of Weber (1946), referring to social groups whose position in society differ according to their relative access to economic resources. We define class primarily in economic terms (Myles and Turegun 1994), relying heavily on parental occupation as the best single indicator of class background (Breiger 1995). However, we used parental occupation only in cases where it was clear that the parents contributed to the player's upbringing. Two examples illustrate journalistic phrases: "Father is an assistant principal...while his mother is a teacher" and "[player's mother] had worked her way up the corporate ladder at [company name]... Because of her success at [company name], [player's mother] could transfer to just about anywhere in the country."

Economic resources influence residential situation. Thus, we took neighborhood descriptions as a further marker of class position.⁸ Neighborhood description consists of two possible indicators. The first is physical appearance of housing: "[Player] would run the 14 flights of stairs in his project" or "...home in a two-car-garage corner of suburbia." The second refers to perceived level and types of crime: "The dangers of the big city came to the [player's childhood family] doorstep... a car drove past, shots were fired," and "In the neighborhood we grew up in, you look around. Everybody is selling drugs."

General statements of class position refer to players' identification of relative economic standing, e.g. "I was neither rich nor poor" indicating, in this case, middle class

⁸ We note, however, that African Americans tend to reside in lower socioeconomic status neighborhoods than comparable whites (Villemez 1980; Iceland et al. 2005). Thus, this indicator could bias our results towards finding more African Americans from lower social class backgrounds. Because we combined as many indicators of parental occupation, neighborhood descriptions, and general statements of class position as possible, this bias is unlikely to change the substantive conclusions. Specific examples of indicators are provided in Appendix A.

status. The other type was journalists' direct reference to class position, e.g. "he was raised in an upper middle class suburb" to indicate above middle class status.

Using information on parental occupation, neighborhood descriptions, and general statements of class position, our coding yielded three categories of social origin: (1) lower class, (2) middle class, and (3) upper middle and upper class.

To identify family background, we examined words and phrases indicating presence or absence of mother, father, or both. A non-grandparent two-parent family corresponds to two parents having raised the respondent. For this category, it did not matter if father and mother were biologically related to the respondent or if they were living together. A necessary condition was that two parents were present and contributing during much of the respondent's childhood upbringing. For most cases we found direct indicators (e.g. "raised by a single mother"), but some were indirect (e.g. "I never had a father around"). If player or reporter indicated "father was never around," we assumed that the biological or step-father was either *de jure* or *de facto* separated from, and not contributing to, the household.

Considering the source of data, it was virtually impossible to determine the exact length of time a respondent was with or without two parents. Hence, we relied on media reports that clearly indicated who was present during a substantial portion of the pre-college or pre-NBA (depending on method of NBA entry) years. We treated step-parent families as two parents unless it was clear that the intervening single parent time spanned most of the respondent's childhood.

We coded responses to family structure background as falling into four categories: (1) raised by single father, (2) raised by single mother, (3) raised by grandparent(s) or social welfare institution, and (4) raised in non-grandparent two-parent family.⁹

⁹ Due to the lack of detailed income and sibship information in our data sources, we could not examine effects of family size on class position (e.g. Downey 1995). Recognizing the role of resource dilution (as a product of family size), we use information on social class that is as detailed as possible to determine discrete class categories (see Appendix A).

Since not all newspaper articles provide the same type of information, we have different sample sizes for different variables. “Response rates” for social class origin are 52 percent, and for family structure background 68 percent. For 46 percent of our respondents, we have information on both social origin variables.

Methodological Issues in Using Newspaper Data: Selection Bias and Description Bias

The benefits of using primary data from secondary sources have been recognized in various disciplines, from research of political phenomena (Earl et al 2004: 65; Woolley 2000: 156) to sport and society (Farquharson and Majoribanks 2006; Hardin et al 2004; Hartmann 2000; Murrell and Edward 1994; Wilson 1997). Yet, newspaper data also incur severe methodological problems, selection bias and description bias especially (Earl et al 2004).

Newspaper Based and Indexing Procedure Selection Bias

Selection bias stems from the contexts and practices of the news agencies in choosing which events are worthwhile reporting on and which not. It creates “the difficulty ...that one risks confounding the substantive phenomenon of interest with the selection process” (Berk 1983: 391). In the context of our data – newspaper articles on professional basketball players – selection bias could stem from newspapers selectively reporting on (a) some players’ biographies’ but not on others, and/or (b) on some players’ social class and/or family structure background but not on others.

Whether the old media adage “if it bleeds, it leads” is the mechanism, or the media’s higher propensity of over-representing African Americans as poor (Gilens 1996), there is reason to suspect that newspapers would be more likely to feature articles on players who faced extreme hardships, deliberately leaving out players who had very little news-worthy tragedy in their lives. In our data, almost all players had some aspect of their life history told

in the newspaper, but variation in the length and depth of the story produces differences in the frequency of valid indicators of class and family structure.¹⁰

To account for this possible problem, our total of 245 articles comes from the widest array of newspapers, ranging from local and regional (e.g. Sacramento Bee, Atlanta-Constitution Journal) to national (e.g. USA Today, New York Times). Research using newspaper data to count collective action events demonstrates that the form and magnitude of selection bias varies by newspaper constituency: a regional newspaper would report different information than a national newspaper (Earl et al 2004: 70; Oliver and Meyers 1999). We also followed the argument that mixing newspaper types yields better data: “all things equal, the more local the focus of attention, the more credible the use of newspapers as a source of event data” (McAdam and Su 2002).

The practices of the on-line search index, in our case Lexis-Nexis, could also introduce selection bias, one of the potential problems being a reduction in the total number of possible available articles (see Woolley 2000).¹¹ However, as Woody points out, when used effectively, on-line indexing can be superior to consulting indexes created by newspaper publishers or private organizations (Woolley 2000: 165). In the next section of the paper, we perform a statistical test to determine whether newspaper-based selection bias is indeed a problem in our analysis.

¹⁰ Newspaper articles containing basketball player biographies are formulaic; a typical biographical story starts with a brief description of some happening that occurred in the player’s life: a tragic death of a loved one, a remembrance of the long hours spent practicing, a particular lesson learned while training that unlocked a hidden potential, a cute story from childhood, and so on. This description is usually supplemented with quotes from parents, friends, mentors, or the athlete himself. Next is a description of the athlete’s playing statistics and current on-the-court performance, either for the year or the career to date. After this introduction, the middle part delves into how the happening described at the beginning of the article provides insights into the player’s personality or on-the-court performance. The end of the article discusses possible futures for the athlete: how the player’s experiences will shape his continued performance or the player’s assessment of his own future and the future of his team.

¹¹ Woolley (2000: 164) also points to an inconsistency between newspapers on the editions they send to Lexis-Nexis: “Newspapers do not always consistently store the same edition with online databases”. We do not know of a credible argument for how variations in newspaper editions would systematically and in a non-random fashion influence the likelihood of finding indicators of class and family structure in players’ biographies.

Description Bias

Description bias refers to a situation in which the newspaper, for a variety of reasons including its own and/or its reporters' prejudices, reports information incorrectly.

Researchers of newspaper accounts of protest events note that description bias is less of a problem with hard news – i.e. who, what, when, where, and why – where facts are more dutifully recorded and with less overt prejudice, and more of a problem with soft news: editorializing of hard facts (Earl et al 2004).

NBA players' biographies can be either hard news or soft news, depending on the newspaper article author's purposes. Players' biographies have media value inasmuch as they are stories of "human interest and human drama" (Oliver and Myers 1999). To accentuate news-worthiness, authors may provide their own judgments and embellish the story, thereby turning hard news into soft news; or, even worse, they may simply lie. With the type of data collection this paper employs, it is impossible to determine the authors' purposes. This is, perhaps, the greatest limitation of our methodological approach.

How serious is the threat of description bias for our analyses? We have no statistical means to detect incorrect information in newspaper authors' depictions of NBA players' biographies. What we do instead is look at the most likely sources of potential misrepresentation of players' social background.

Trouble with reliability across indicators of social class and family structure background could occur if reporters apply sociological concepts *ad verbatim*, because they may use them in the wrong context. In our data, sociological terms rarely appeared as such: for example, out of the 245 analyzed articles, only five contained the term "class" in reference to economic position or life chances. Usually, "class" referred to descriptions of behavior (e.g. "class act"), or to school. In coding we were aware of potential misuse of sociological jargon, but noticed that little of it was present. Furthermore, our data collection approach

relies on journalists' strong-suit, that is, detailed observations of the neighborhoods they visit and the people they meet, which are akin to hard news and can be successfully used in sociological work (Earl et al 2004).

Even if journalists are biased in the indicators that they report, the forms of bias are likely to be contradictory, in a sense canceling each other out. For example, research suggests that Americans tend to classify themselves as middle class when more objective measures indicate otherwise, and journalists may exhibit this behavior in their reporting (Jackman and Jackman 1983; Luo and Brayfield 1996). If this is so, our results would over-represent middle class NBA players. On the other hand, given the media's tendency to show African Americans in poverty, journalists may report more indicators of low class position, bending our results toward an overrepresentation of lower class African American NBA players (Gilens 1996). Moreover, "rags-to-riches" stories are popular American fare, compelling journalists to seek out these stories whenever possible (McManus 1994; Shoemaker and Reese 1996). If this is the case, then the number of NBA players who rose from poverty to NBA wealth and fame would be overrepresented, rather than underrepresented, in our data. In conclusion, while we cannot statistically prove that description bias does not seriously affect our findings, we have strong logical arguments that this is the case.

Findings

Our first interest is in the social milieu NBA basketball players are most likely to come from. Confirming our hypothesis, the data indicate that most players come from the middle class and above, and not from the lower social classes (Table 1). Looking at the combined categories of upper and middle class of origin reveals that 66 percent of African Americans and 93 percent of whites have advantaged social background. One can see, indeed, that across racial groups, the proportion of African Americans coming from a disadvantaged class

is significantly larger than that of whites ($z = 2.76, p < 0.01$). Yet, irrespective of race, the vast majority of professional basketball players in the NBA enjoy a relatively advantaged social origin.

--Table 1 about here--

With regards to family structure background composition, the picture is more complex. Results in Table 2 show that the majority of white players come, as we expected, from two-parent families (81 percent). For African Americans, this type of family background makes up only 43 percent of the sample. The majority of African American players, 57 percent, grew up in families with only one parent. Single-mother households dominate this form of disadvantage, as they constitute well over one third (44 percent) of the environment African American NBA players were raised in.

--Table 2 about here--

To determine whether the differences in family structure background composition between whites and African Americans are not due solely to chance, we performed the test of differences between proportions. Results confirm that the race difference is statistically significant ($z = 2.80, p < 0.01$).

Next, we were interested if professional basketball players who made it into the NBA display compounded disadvantage of low socioeconomic origin and disadvantaged family structure background. To find out, we first calculated the percent that had either lower class or non-two parent family structure backgrounds, and then the combination of both (see Table 3). Among white players, 21 percent has either form of disadvantage, and no one has both. In the case of African Americans, well over half have either form of disadvantage (59 percent), and 28 percent has both. The race difference is statistically significant ($z = 3.01, p < .01$). However, we stress that for both race groups the proportion of players with

compounded disadvantage is relatively low (see discussion below for comparisons to the general population), supporting our expectation that among professional basketball players the presence of both low social class and non-two parent family structure background would be rare.

-- Table 3 about here --

We used data from the U.S. Census 2000 and the General Social Survey to calculate odds ratios that assist in understanding how the intersection of race and social class and family structure background influences the chances of NBA attainment (see Appendix B for how we measured class and family structure using these other data sets).

In terms of odds ratios, we estimate that among African Americans, a child raised in low class family has 37% less chances to become an NBA player than a child raised in middle or upper class family. An analogous figure for whites is much higher (75%) but revealing the same tendency. In the case of non-two-parent families, the coefficients for African American and whites are more similar than in the case of social class. Among African Americans a child from a non-two-parent family has 18% less chance to become an NBA player than a child raised in two-parent family. For whites the figure is less than twice as large (33%).

Assessing the Impact of Newspaper-Based Selection Bias

In terms of statistical modeling, selection bias occurs when the distribution of studied variables among cases included in the analysis (non-missing cases) differs from the distribution of the same variables among cases that potentially belong to the sample (missing cases). To determine whether our data bears this problem, we need to know if the included cases, that is, NBA players for whom information on social origin is available, are

significantly different from NBA players for whom such information is lacking. Specifically, considering our hypotheses, we need to ask: “Did newspaper-based selection bias result in richer data on players from advantaged social origins?” For if that were the case, this form of bias would have affected our results.

To answer the question above, we must identify the circumstances under which journalists would report on advantaged NBA players more often than on disadvantaged ones. Considering that newspapers are in the business of attracting customers, it is reasonable to expect that the media most likely focuses their human interest stories on players who are among the best and, hence, the most popular. Selection bias, then, would be toward ‘eminent’ NBA players, i.e. those who are among the best in the league: more detailed information would be given about them, including their social class and family structure background.

According to our theory, players from advantaged social origins are more likely to have made it into the NBA. Extending this assumption to ‘eminence’, namely that the more advantaged the player’s social origins, the more likely he is to acquire the skills necessary to be among the best in the NBA, we formulate the following hypotheses for the mechanism that could lead newspaper-based selection bias to affect our data:

Hypothesis 1: The more eminent an NBA player, the more likely news would report detailed information on his social class and family structure background.

Hypothesis 2: NBA players from more advantaged origins are more likely to be eminent players.

Hypothesis 3: In so far as both the above two expectations are met, newspaper articles would cover in greater detail information on social class and family structure background of NBA players from more advantaged origins.

If Hypothesis 3 were to have empirical support, our data would be biased. While we cannot test it directly, we can determine whether the first two expectations of the newspaper-based selection mechanism hold. Specifically, for selection bias to occur, advantaged social origins *must be related* to being eminent (Hypothesis 2), which in turn should be related to being included in our data (Hypothesis 1). If, on the other hand, the relationship between advantaged origins and being an *eminent* NBA player is not significant, then the news articles have not reported selectively on players from advantaged backgrounds (although they may have selected more often on eminent ones). In other words, as long as Hypothesis 2 does not find support, we can conclude that NBA players for whom detailed social class and family structure background information is available are not more likely to come from advantaged origins than NBA players for whom we do not have similar information (missing cases).

To perform this test, we first determined how the 155 players in our subpopulation rank on eminence. We measured eminence as position in the *NBA draft*, where the higher the score, the higher the draft pick.¹² These scores obtained from draft position we use as independent variable in logistic regression in which the dependent variable is existence of data on social class and family-structure background (1=yes, 0 = no).

-- Table 4 about here --

Table 4 presents the logistic regression results. Supporting Hypothesis 1, more eminent players are significantly more likely to have information on social class, on family structure, and on both variables combined. For each dependent variable, the model fit is satisfactory.

To assess whether advantaged social origins are positively and significantly related to players' eminence (Hypothesis 2), we employed ordinary least squares regression of player's eminence separately on three independent variables: advantaged social class (middle to upper

social class origins), advantaged family structure background (two-parent family), and compounded advantage (both advantaged social class and family structure background). Results (see Table 5) show that that none of the advantaged origins variables have a significant effect on eminence.

-- Table 5 about here --

Substantively, having failed to find support for Hypothesis 2 allows us to conclude that our data is not biased toward advantaged NBA players: in terms of social class and/or family background, players for whom this information is available are not substantively different from players on which such information is lacking (i.e. missing cases).

Summary and Discussion

In this paper we argued that major stratification variables, namely race, social class and family structure, and especially their intersection, as key components in the sports-upward mobility relationship, will affect membership in the NBA. Following data collection techniques researchers of political phenomena employ (see Earl et al 2004; Oliver and Meyers 1999; Woolley 2000), we constructed a unique dataset on professional basketball players' social origin through coding 245 newspaper articles published in local, regional and/or national papers over a ten years time span. Given the lack of quantitative data containing mobility variables for professional athletes, using primary data from secondary sources to address our research question was crucial. However, this approach raises a set of methodological issues that could have affected our results: bias due to sampling on the dependent variable, newspaper-based selection bias, and newspaper-based description bias. We performed a number of statistical tests to determine if sampling on the dependent variable and newspaper-based selection bias are a serious threat to our substantive conclusions. Their

¹² In particular, those who were the first pick overall in a given year's NBA Draft score highest. Since not all players were drafted, non-draftees are coded with the lowest possible score (6.5% of 155).

results, which we discuss extensively in the paper (see Appendix B, and Table 4 and 5, respectively) allow us to conclude that this is not the case.

We have no statistical means, however, to detect if there is incorrect information in newspaper authors' depictions of NBA players' biographies, that is, to assess if description bias is present, and to what extent. Thus, we examine theoretically the most likely sources of potential misrepresentation of NBA players' social background, such as reporters' inaccurate usage of sociological concepts, and/or bias in the social origins indicators that they report. In coding we were aware of potential misuse of sociological jargon, but noticed that little of it was present; we draw on literature on journalistic practices suggesting that bias could lead in multiple directions simultaneously, in a sense canceling itself out. For example, our results could over-represent middle class NBA players given Americans' tendency to classify themselves as middle class, a behavior journalists may exhibit in their reporting too (Jackman and Jackman 1983; Luo and Brayfield 1996). On the other hand, given the media's tendency to show African Americans in poverty, journalists may report more indicators of low class position, bending our results toward an overrepresentation of lower class African American NBA players (Gilens 1996). Moreover, "rags-to-riches" stories are popular American fare, compelling journalists to seek out these stories whenever possible (McManus 1994; Shoemaker and Reese 1996). If this is the case, then the number of NBA players who rose from poverty to NBA wealth and fame would be overrepresented, rather than underrepresented, in our data. In conclusion, we cannot statistically prove that description bias does not seriously affect our findings, but we have strong logical arguments that it does not.

In line with our research hypotheses, analyses in this paper show that, similarly to other achievement paths, membership in the US professional basketball league is significantly related to 'classic' stratification variables. Among NBA players on whom we have social

origin information, 93 percent of whites and 66 percent of African Americans come from the middle class and above. Twenty one percent of white players have either disadvantaged social class or disadvantaged family background, while the corresponding percentage for African American players is 59. Not one single white player displays compounded disadvantage of low socioeconomic origin and disadvantaged family structure background, whereas 28 percent of African American players do. All racial differences are statistically significant. In terms of odds of achievement, we find that African Americans from disadvantaged social origins have lower odds of being in the NBA than both whites of similar disadvantage and African Americans and whites of relative advantage. For example, among African Americans, a child raised in low class family has 37% less chances to become an NBA player than a child raised in middle or upper class family, while the analogous figure for whites is much higher (75%).

Overall, these results suggest that NBA attainment is not an entirely meritocratic process, as the intersection of race, class and family structure background, presents unequal pathways into the league. We would like to stress here that meritocracy, or the lack thereof, cannot be gleaned just from the pronouncement of “the odds,” as academic and non-academic reports tend to do. To place this study in proper theoretical perspective, the most anyone can say about the odds is that the intersection of race and class and family contextualizes the opportunity structure for attaining an NBA position. Thus, before invoking “the odds” for or against the merit of sports for upward mobility of disadvantaged groups, one should specify the ultimate goal of attainment: if it is the NBA, then referring to the odds as an indicator of the unlikelihood to succeed through basketball playing is useful, as clearly the league is virtually unattainable for disadvantaged youth, particularly African Americans. If, on the other hand, the goal is ‘just’ upward mobility through sport participation (i.e. attaining other ends than the NBA), then sport could still represent a beneficial means, as previous research

indicates (Eide and Ronan 2001). In any case, the odds are influenced by a broad range of factors, most basically by relative disadvantage of social origins, which should be built into their calculation.

Refining the odds by examining the intersection of race and social class and family structure background should also refine the message of activists who advocate sport as a good channel for upward mobility. There is evidence that disadvantaged youth's career aspirations can be unrealistic, focusing on sport participation as a means to the NBA rather than as a means to college and improved human capital (see Sellers and Kuperminc 1997). Thus, activists' message should emphasize sport participation because it develops life skills and enhances human capital rather than as a path to a sports career.

There are many excellent foci for future research to further the understanding of temporal, spatial, and social aspects of sport mobility into the professional leagues. Expanding our subsample to other generations of players, both younger and older, would allow us to see if the effect of social origin diminishes or not over time. As the NBA increasingly draws from players outside the U.S. and Canada, geographic reach could also extend to a greater range of international players. In addition, other sports with relatively similar concentrations of African American and white players (such as the National Football League or Major League Baseball) would provide contexts to analyze the intersection of race, class, and family structure background. Researchers should also examine Women's National Basketball Association players to determine whether the intersection of gender with our variables influences the sport mobility process. Finally, we encourage other scholars to replicate our results. Replication of scientific endeavors is not as common as should be (King 1995); we provided a detailed methodological road-map to enable replication of this study and expansion of our data collection approach.

References

- Baruch, Yehuda. 1999. "Response Rate in Academic Studies – A Comparative Analysis." *Human Relations* 54 (4), 421-438
- Berk, Richard. 1983. "An Introduction to Sample Selection Bias in Sociological Data." *American Sociological Review* 48:386-98.
- Braddock, J. 1991. Race, Athletics, and Educational Attainment: Dispelling the Myths. *Youth and Society* 12: 335-350.
- Breiger, Ronald L. 1995. "Social Structure and the Phenomenology of Attainment." *Annual Review of Sociology* 21:115-36.
- Coakley, Jay. 2004. *Sports in Society: Issues and Controversies*. New York: McGraw-Hill.
- Craggs, Tommy. 2008. "Where They Come From." ESPN Magazine On-Line: <http://sports.espn.go.com/espnmag/story?id=3243679>. Accessed October 5, 2009.
- Downey, Douglas B. 1995. "When Bigger Is Not Better: Family Size, Parental Resources, and Children's Educational Performance." *American Sociological Review* 60(5):746-61.
- Dubrow, Joshua. 2002. "Last Bus Out of the Ghetto: Towards a Theory as to Why African Americans Dominate Modern Professional Basketball." Paper presented at the annual meeting of the American Sociological Association, Chicago, IL, August.
- Earl, Jennifer and Andrew Martin, John D. McCarthy, Sarah A. Soule. 2004. "The Use of Newspaper Data in the Study of Collective Action." *Annual Review of Sociology* 30:65-80.
- Edwards, Harry. (1973). *Sociology of Sport*. New York: Dorsey Press.
- Eide, Eric R. and Nick Ronan. 2001. Is Participation in High School Athletics an Investment or a Consumption Good? Evidence from High School and Beyond. *Economics of Education Review* 20, 431-442.
- Eitzen, D. Stanley. 1996. „Classism in Sport: The Powerless Bear the Burden.” *Journal of Sport and Social Issues* 20(1): 95-105.
- Eitzen, D. Stanley 1999. "Upward Mobility through Sport? The Myths and Realities." Retrieved April 16, 2007. (<http://www.zmag.org/zmag/articles/mar99eitzen.htm>)
- Eitzen, D. Stanley. 1999. *Fair and Foul: Beyond the Myths and Paradoxes of Sport*. Lanham, MD: Rowan & Littlefield.
- Entine, Jon. (2000). *Taboo: Why Black Athletes Dominate Sports and Why We Are Afraid to Talk About It*. New York: Public Affairs.
- Gilens, Martin. 1996. "Race and Poverty in America: Public Misperceptions and the American News Media." *Public Opinion Quarterly* 60:515-41.
- Goldberg, Steven.1991. *When Wish Replaces Thought*. Buffalo, NY: Prometheus Press.

- Hardin, Marie and Julie E. Dodd, Jean Chance, Kristie Walsdorf. 2004. „Sporting Images in Black and White: Race in Newspaper Coverage of the 2000 Olympic Games.“ *Howard Journal of Communication* 15(4): 211-227.
- Harris, Othello. 1998. "The Role of Sport in the Black Community." *African Americans in Sport: Contemporary Themes*, edited by Gary Sailes. New Brunswick, NJ: Transaction Publishers.
- Harrison Jr., Louis, Laura Azzarito, and Joe Burden Jr. 2004 "Perceptions of Athletic Superiority: A View From the Other Side." *Race, Ethnicity, and Education* 7(2):149-66.
- Hartmann, Douglas. 2000. "Rethinking the Relationships between Sport and Race in American Culture: Golden Ghettos and Contested Terrain." *Sociology of Sport Journal* 17: 229-253.
- Hoberman, John. 1997. *Darwin's Athletes: How Sport Has Damaged Black America and Preserved the Myth of Race*. Boston: Houghton Mifflin.
- Iceland, John, Cicely Sharpe, and Erika Steinmetz. 2005. "Class Differences in African American Residential Patterns in US Metropolitan Areas: 1990–2000." *Social Science Research* 34(1):252-66.
- Jackman, Mary R. and Robert W. Jackman. 1983. *Class Awareness in the United States*. Berkeley: University of California Press.
- Karen Farquharson and Timothy Majoribanks. 2006. "Representing Australia: Race, the Media, and Cricket." *Journal of Sociology* 42(1): 25-41.
- Kerbo, Harold R. 2003. *Social Stratification and Inequality*. New York: McGraw Hill.
- King, Gary. 1995. "Replication, Replication." *PS: Political Science and Politics*, Vol. 28(3): 443-499.
- Knoppers, Annelies and Agnes Elling. 2004. "'We Do Not Engage in Promotional Journalism': Discursive Strategies Used by Sport Journalists to Describe the Selection Process." *International Review for the Sociology of Sport* 39: 57-73.
- Leonard, Wilbert M. 1996. "The Odds of Transiting From One Level of Sports Participation to Another." *Sociology of Sport Journal* 13(3):288-99.
- Luo, Ye and April Brayfield. 1996. "Women's and Men's Subjective Class Identification From the 1970s to the 1990s." *Sociological Spectrum* 16(1):83-107.
- MacLeod, Jay. 1995. *Ain't No Makin' It: Aspirations and Attainment in a Low-Income Neighborhood*. Boulder, CO: Westview Press.
- Maisel, Sandy and Walter Stone. 1998. "The Politics of Government Funded Research: Notes from the Experience of the Candidate Emergence Study." *PS: Political Science and Politics* 31:811-817.

- Martin, John L. and King-To Yeung. 2003. "The Use of the Conceptual Category of Race in American Sociology, 1937–99." *Sociological Forum* 18(4):521-43.
- McAdam, Doug and Yang Su. 2002. "The War at Home: The Impact of Anti-War Protests, 1965-1973." *American Sociological Review* 67:696-721.
- McLanahan, Sara and Gary Sandefur. 1994. *Growing Up With a Single Parent: What Hurts, What Helps*. Cambridge: Harvard University Press.
- McManus, John. 1994. *Market-Driven Journalism: Let the Citizen Beware?* Thousand Oaks: Sage Publications.
- Milne, Ann M., David E. Myers, Alvin S. Rosenthal, and Alan Ginsburg. 1986. "Single Parents, Working Mothers, and the Educational Achievement of School Children." *Sociology of Education* 59(3):125-39.
- Murrell, Audrey J. and Edward M. Curtis. 1994. "Causal Attributions of Performance for Black and White Quarterbacks in the NFL: A Look at the Sports Pages." *Journal of Sport and Social Issues* 18(3): 224-233.
- Myles, John and Adnan Turegun. 1994. "Comparative Studies in Class Structure." *Annual Review of Sociology* 20:103-24.
- National College Athletic Association. "Estimated Probability of Competing in Athletics Beyond the High School Interscholastic Level." Retrieved October 26, 2007. (<http://www.ncaa.org>).
- Ogden, David C. and Michael L. Hilt. 2003. "Collective Identity and Basketball: An Explanation for the Decreasing Number of African Americans on America's Baseball Diamonds." *Journal of Leisure Research* 35(2):213-27.
- Oliver, Pamela E. and Daniel J. Myers. 1999. "How Events Enter the Public Sphere: Conflict, Location, and Sponsorship in Local Newspaper Coverage of Public Events" *American Journal of Sociology* 105(1): 38-87.
- Pascarella, Ernest T. and John C. Smart. 1991. "Impact of Intercollegiate Athletic Participation for African American and Caucasian Men." *Journal of College Student Development* 32:123-30.
- Sailes, Gary. 1998. "Betting Against the Odds: An Overview of Black Sports Participation." *African Americans in Sport: Contemporary Themes*, edited by Gary Sailes. New Brunswick, NJ: Transaction Publishers.
- Sellers, Robert M. and Gabriel P. Kuperminc. 1997. "Goal Discrepancy in African American Male Student-Athletes' Unrealistic Expectations for Careers in Professional Sports." *Journal of Black Psychology* 23(1):6-23.
- Shoemaker, Pamela J. and Stephen D. Reese. 1996. *Mediating the Message: Theories of Influences on Mass Media*. Allyn & Bacon. New York.

- Spreitzer, Elmer. 1994. "Does Participation in Interscholastic Athletics Affect Adult Development? A Longitudinal Analysis of an 18-24 Age Cohort." *Youth & Society* 25(3):368-87.
- Stolzenberg, Ross M. and Daniel A. Relles. 1997. "Tools for Intuition about Sample Selection Bias and Its Correction." *American Sociological Review*. 62(3): 494-507.
- Villemez, Wayne J. 1980. "Race, Class, and Neighborhood: Differences in the Residential Return on Individual Resources." *Social Forces* 59(2):414-30.
- Washington, R. E., & Karen, D. (2001). Sport and Society. *Annual Review of Sociology* 27, 187-212.
- Weber, Max. 1946. "Class, Status, Party." Pp. 180-195 in *From Max Weber: Essays in Sociology*, edited by Hans Gerth and C. Wright Mills. New York: Oxford University Press.
- Wilson, Brian. 1997. "'Good Blacks' and 'Bad Blacks': Media Constructions of African-American Athletes in Canadian Basketball." *International Review for the Sociology of Sport* 32(2): 177-189.
- Wilson, William J. 1996. *When Work Disappears: The World of the New Urban Poor*. New York: Knopf.
- Winant, Howard. 2000. "Race and Race Theory." *Annual Review of Sociology* 26:169-85.
- Winship, Christopher and Robert D. Mare. 1992. "Models for Sample Selection Bias." *Annual Review of Sociology* 18: 327-350.
- Woolley, John T. 2000. "Using Media-Based Data in Studies of Politics." *American Journal of Political Science* 44(1):156-173.

Table 1 Social Class Background by Race (%)

	Race	
	African American	White
Upper Middle Class and Above	8.96	35.71
Middle Class	56.72	57.14
Lower Class	34.33	7.14
N	67	14

Table 2 Family Structure Background by Race (%)

	Race	
	African American	White
Single Parent		
Mother	44.44	18.75
Father	3.33	0
Grandparent(s), Other	8.89	0
Two Parent	43.33	81.25
N	90	16

Table 3 Compounded Social Origin Disadvantage by Race

Form of Disadvantage	Race	
	African American	White
Lower Class <i>or</i> Non-Two-Parent Family	58.76%	21.05%
N	97	19
Lower Class <i>and</i> Non-Two Parent Family	28.33%	0%
N	60	11

Table 4 Logistic Regression of Existence of Data on Social Class and Family-Structure Background on Players' Eminence

	Existence of Data (Yes = 1, No = 0)								
	Social Class			Family Structure			Both		
	b	Standard Error	Exp(B)	b	Standard Error	Exp(B)	b	Standard Error	Exp(B)
Players' Eminence	0.028***	0.010	1.029	0.038***	0.010	1.038	0.032***	0.010	1.033
Constant	-0.916***	0.384	0.400	-0.522	0.378	0.593	-1.333***	0.410	0.264
-2 Log Likelihood	205.541			180.732			202.666		
Model Chi Square	9.019***			14.196***			11.119***		
Cox and Snell R ²	0.057			0.088			0.069		
N	155			155			155		

*** p < 0.01

Table 5 Ordinary Least Squares Regression of Players' Eminence on Advantaged Social Origin Variables

	B (standard error)
Advantaged Social Origins	
Middle to Upper Social Class Background	4.340 (4.161)
Constant	36.292*** (3.490)
R ²	0.014
N	80
Two Parent Family Structure	-2.673 (3.231)
Constant	40.500*** (2.263)
R ²	0.007
N	105
Both Forms of Advantage	2.772 (3.982)
Constant	39.103*** (2.673)
R ²	0.007
N	70

*** p < 0.01

Appendix A: Examples of Coding Categories for Social Origin

Social Class

Lower

- “years filled with difficult memories such as poverty...”
- “There was a lot of crime in my neighborhood...”
- “...a single-parent home in the urban projects.”
- “[Player] saw his mother struggle...Such financial struggles are now over.”
- “He grew up...in a housing project.”
- “[Player] would run the 14 flights of stairs in his project...I lived in poverty all my life, too.”

Middle

- “she had worked her way up the corporate ladder...because of her success [there, she] could transfer to just about anywhere in the country.”
- “My dad was in the Air Force. He was a staff sergeant.”
- “His family was neither rich nor poor.”
- “becoming a South Side business man with his own cab company...”
- “his father, a computer analyst...”
- “Father is an assistant principal...while his mother is a teacher.”
- “...from a two-parent home in a two-car-garage corner of suburbia”

Upper

- “as the son of [an NBA] coach...”
- “[Mother and father], who own an insurance agency...considered [\$12,500 per year tuition per year] a fair price for a school who’s graduates attend some of the country’s top colleges. The family’s five-bedroom home in [exclusive neighborhood]...”
- “His father was a defensive end for the Browns...now an assistant coach with the New Orleans Saints.”
- “A journeyman in the NBA and European leagues...the [family] is financially secure; they don’t need the NBA’s money.”
- “[Father] played nine seasons for the Buffalo Bills and Seattle Seahawks...”

Family-Structure Background

Two Parent Family

- “his mother, speaking from their home in [town name]...[father] built it 20 years ago at their home in [town name].”
- (father speaking) “My wife and I felt that during the summer it was better for him to...”
- “[Player] and his parents, [mother and father], spoke on parental involvement...”
- “I don’t know how I’d react if my mother was in a single-parent situation, if my father wasn’t around and I saw my mother struggle everyday...”

- “His parents wanted [player] to remain close to home.”
- “...from a two-parent home in a two-car-garage corner of suburbia”

Single Mother

- “as a single parent with two sons, she...”
- “She was a single parent.”
- “He never knew his real father, who left when [he] was 3...”
- “He grew up with his mom...after his dad left them.”
- “...[player] was raised by his mother...”
- “[His] father, who is not married to [mother] and has a separate life and family...”

Single Father

- [Player] was raised...by his father...has met his mother once.”
- “Growing up...his father, [father], supported five children...his mother died of breast cancer.”

Grandparent / Other

- “the grandmother who raised him from early childhood”
- “With his father dead and his mother in and out of prison...moved in with his select team coach.”
- “The two [player and brother] lived together at [Named] House for neglected and abused children, until [he] reached high-school age and moved to...another group home.”
- [Player]’s parents had been drug users with marital problems, and when he was 4 he moved in with his grandparents...The death of his parents to AIDS forced [player] to grow up quickly.”
- “...various problems forced him to [live] with his grandparents...and his uncle.”

In some cases indicators for social class and/or family background were blurry. Whenever descriptions did not fall clearly into a particular category, we did not code them, i.e. we treated that respondent (player) as missing data. The following description is an excellent example of how information can yield no useful codes:

“[player’s name] comes from the small town of [player’s hometown name], the second of eight children. His mother is on disability and his father is in ... prison. [Player’s name] thinks often of his grandparents’ farm in [place name], where he spent his first seven years. He still likes to return in the summer, doing chores and fishing with his cousins.”

Despite a rather detailed description of this player’s personal life and upbringing, parental occupation is not clear. We could assume that he belongs to the lower class, but we cannot be

sure: we do not know the respondent's father occupation prior to imprisonment, nor the mother's employment status prior to disability. Neither do we know the extent to which the grandparents provided support outside of summer-time invitations to the farm. In addition, missing information on when the father went to jail prevented us from making a family-structure decision. As such, this player is treated as a missing data case on both social class and family structure.

Appendix B: Sampling Based on Characteristics of the Dependent Variable

Analyses in this paper are performed on basketball players who have made it into the NBA. This raises the methodological issue of sampling based on the characteristics of the dependent variable, that is, as to whether we accurately account for the effects of the causal mechanisms on the dependent variable (Winship and Mare 1992; Stolenzberg and Relles 1997). Causality is defined here as a clear temporal chain of events where social origins (race, class and family structure) influence social destinations (attaining NBA status). Our causal model assumes an NBA sorting process as the mechanism for players' selection into the league. Following sorting, which occurs before teams select players into the NBA, players are selected into the league (1) or not (0). In our approach, sampling on the dependent variable is akin to what Winship and Mare (1992) refer to as „truncated sample explicit selection” (330): we only have cases of the dependent variable where selected into the league = 1; characteristics of those not selected are unknown.

To assess whether sampling on the dependent variable affects our substantive conclusions, we would need to know the characteristics of those who were eliminated by the NBA sorting process, and hence not selected into the league. According to our hypothesis, aspirants who come from advantaged social origins are more likely to be selected into the NBA than those who do not. If the majority of not selected players were from lower social origins, then disadvantage may decrease the odds of attaining an NBA position, in line with our expectation. If, however, most of those not selected were advantaged aspirants, then advantage may decrease the odds of attainment, a situation contrary to our hypothesis. We approximate the population of basketball aspirants as male children in the U.S. We analyze two data sources containing these potential aspirants – the U.S. Census 2000 and the General Social Survey (GSS) 1972-2002 – and compare the results with the newspaper data.¹³

¹³ The Census and the GSS data do not contain social class indicators directly comparable to ours; however, considering the ubiquity of income as an indicator of both socioeconomic status (SES) and social class, income

For the U.S. Census 2000 data on male children, we define lower social class as a category at or below 150 percent of the poverty line in 2000. In 2000, the average poverty threshold was \$17,608 for single person households and \$26,412 for a family of four. Using these data we estimate that 45 percent of African American male children live in a lower social class context, while only 34 percent of African American NBA players come from lower social class background ($z = -1.80, p < 0.05$).¹⁴ Similarly, 23 percent of white male children live in a lower social class context, while only 7 percent of white NBA players were from a lower social class background ($z = -1.42, p < 0.10$).

The differences between the relevant comparison group in the U.S. population, and NBA player, are less pronounced. Among African American male children, 62 percent come from non-two parent families (CPS 2002), while in the NBA sample 57 percent of African Americans do ($z = -0.95, p = 0.17$).¹⁵ Similarly, 26 percent of all white male children and 19 percent of white NBA players come from non-two parent families ($z = -0.63, p = 0.26$).

The results above, together with findings in Table 1 and 2, allow us to conclude the following: the difference between African American and white NBA players in terms of advantaged social origins is statistically significant, yet NBA players as a whole are less likely to come from disadvantaged backgrounds compared to the relevant comparison group in the US population.

For compounded disadvantage, we used the General Social Survey 1972 to 2002, including only males. For this data set, we defined lower social class as the category of

in these data sport a valid proxy (Iceland et al. 2005). Family structure background variables are reasonably comparable.

¹⁴ Current Population Survey (CPS). 2002. http://pubdb3.census.gov/macro/032001/pov/new22_005.htm Retrieved February 24, 2005.

¹⁵ http://www.nichd.nih.gov/publications/pubs/childstats/report2002_5.pdf Appendix A: Detailed Tables: Table Econ1A. Retrieved February 24, 2005.

respondents reporting below-average family income level at age 16. For family structure background, we created a dummy variable distinguishing two-parent family from the rest.

For African Americans, 28 percent of the NBA players come from compounded disadvantaged backgrounds, compared to 39 percent in the GSS sample ($z = -1.66$, $p = .05$, one tail test). None of the white players in our study come from compounded disadvantaged origins, compared to 30 percent in the GSS sample ($z = -2.17$, $p < .01$). These results indicate that NBA players as a whole are less likely to come from compounded disadvantaged backgrounds than their counterparts from the general population.

In sum, comparisons with regard to social class, family background, and the combination of the two, between potential basketball aspirants as defined earlier, and NBA players, indicate that the latter are significantly better off in terms of social origins than the former. This finding is in line with our research hypothesis that the advantaged are more likely to be selected into the league than the disadvantaged, and suggests that in this study sampling based on characteristics of the dependent variable did not substantially alter the substantive conclusions.

Appendix C

Following Woolley (2000: 164), word searches via Lexis-Nexis was a two-step process designed to reduce missing data. We first crafted words that were likely to be included in a player biography article: [player name], basketball, mother, mom, father, dad, grandparent(s), grandma, grandfather, parent(s), family, and class. We then refined the word search process based on our evaluation of the usefulness of the words and their combinations, noting that some words, like the player's name and the word „family,“ were more likely to yield social origin indicators. Variations on mother, such as “mom,“ were used, providing some articles that we would not have found otherwise but often times not yielding valid enough indicators of parental status. We used multiple word combinations in the search process for each player: for example, we used [player name], basketball, and mother in the same search, and in subsequent searches for that player, we replaced mother with father, parent, family, class and others we identified in the first step of the process.

Our data collection strategy is closest to what Woolley (2000: 157) describes as “record-coding”: coding to create an event record, or “count.” In practice, instead of gathering every article that referred to social origin for a particular player, we stopped searching for news articles if (a) the same information was repeated in multiple articles and (b) after exhausting all possible word searches.

Reliability checks are an important part of the process, and thus we used a common inter-rater reliability procedure. After gathering what appeared to be relevant articles, we each coded them independently. We then compared our codes, discussing similarities and discrepancies in light of our criteria for the coding categories. We resolved discrepancies through focused discussions. When discrepancies could not be resolved, the indicator was coded as missing. On average, we use two articles per player where social origin information could be coded; often, multiple indicators were found in the same article. For social class, the

inter-rater reliability percentage was 89, taking into account only those cases that had sufficient information with which to code ($n = 75$). Similarly, our inter-rater reliability percentage for family structure background was 94.