January 2019

A Social Space Approach to Testing Complex Hypotheses: The Case of Hispanic Marriage Patterns in the United States

Jeffrey A. Smith

University of Nebraska - Lincoln, jsmith77@unl.edu

Follow this and additional works at: http://digitalcommons.unl.edu/sociologyfacpub

Part of the Family, Life Course, and Society Commons, and the Social Psychology and Interaction Commons
A Social Space Approach to Testing Complex Hypotheses: The Case of Hispanic Marriage Patterns in the United States

Jeffrey A. Smith

Abstract
Where do individuals identifying as Hispanic fit in the racial landscape of the United States? The answer offered by past work is complex: The empirical results do not lend themselves to simple interpretation as no single hypothesis fits the Hispanic case very well. Instead, Hispanic integration is described as mixtures of different archetypical hypotheses, like panethnic formation, white assimilation, and racialized assimilation. My goal is to develop a formal framework to help make sense of this complex picture. I extend past work by showing which combination of integration processes (panethnic formation, white assimilation, etc.) best characterizes Hispanic marriage patterns. I make two analytical contributions. First, I organize past Hispanic hypotheses, both archetypical and blended, into a single theoretical framework defined by the salience of race and Hispanic ethnicity. Second, I parametrize this theoretical framework using latent social space models. In this way, I am able to specify a set of interconnected, complex hypotheses in a tractable manner. I follow past work and use marriage/cohabitation data to test the hypotheses. Using American Community Survey data (2010–2012), I find that Hispanic marriage/cohabitation patterns suggest high salience on both race and Hispanic ethnicity. Thus, categories like black-Mexican or white-Cuban represent relationally distinct social categories—distinct from both non-Hispanic racial categories (e.g., black or white) and Hispanic categories of a different racial identity.

Keywords
social space, social distance, marriage/cohabitation, race/ethnicity

Introduction
How do individuals identifying as Hispanic fit in the American racial landscape? The answer offered by past work is a complicated one as past results do not lend themselves to a single, simple interpretation (Feliciano, Lee, and Robnett 2011; Perez and Hirschman 2009). Words like complex, heterogeneous, and mixed are employed to make sense of a difficult empirical picture—one where the evidence rarely fits neatly into one hypothesis and does not generally hold for all groups (McConnell and Delgado-Romero 2004; Oropesa, Landale, and Greif 2008; Rodriguez and Cordero-Guzman 1992). As Alba, Jiménez, and Marrow (2014) conclude, “no single existing theoretical model will capture this diversity” (see also Telles 2010). The literature on marriage/cohabitation offers a clear example of these mixed results. Many studies find that Hispanics have high rates of intermarriage with the white majority, indicative of a classic assimilation trajectory (Lee and Bean 2004; Qian and Lichter 2007). The same studies, however, often find high rates of in-group marriage (among Hispanics), suggesting the panethnic term Hispanic is itself a meaningful social category (e.g., Fu 2007; Rosenfeld 2001). Still other studies find strong differences by racial identity within the Hispanic population (Qian 2002).

Recent work has grappled with this heterogeneity more directly, documenting the diversity of outcomes across Hispanic subgroups (Alba et al. 2014; Rumbaut 2009). A number of studies have considered differences in Hispanic marriage patterns across geographic regions (Choi and...
I approach the problem from a different angle, focusing on the challenge of specifying hypotheses in a context of mixed, or blended, results. Past work has described Hispanic marriage patterns in terms of white assimilation, panethnic formation (where Hispanic represents a racial/ethnic group in itself, like white or black), and racialized assimilation (where different Hispanic groups are incorporated into different racial groups) while suggesting that a blend of these ideas may in fact fit best, for example, because different hypotheses hold for different Hispanic subgroups. There is little consensus, however, on which combination of these processes (panethnic formation, white assimilation, etc.) best characterizes Hispanic marriage patterns in the United States. Thus, the question is not whether there are heterogeneous outcomes across Hispanic subgroups; rather, the question is which pattern of heterogeneity emerges based on the mixture of multiple processes of integration. Analytically, this is a difficult problem to tackle. The range of possible hypotheses is quite large as every combination of white assimilation, panethnic formation, and racialized assimilation yields a different set of implied marriage patterns. This makes it difficult to rely on traditional methods as we must consider a much wider range of hypotheses than in past work.

My goal is to develop a formal framework that captures the full range and complexity of Hispanic hypotheses. There are two key questions. First, how can we integrate all hypotheses into a single theoretical framework? And second, how can we parametrize this framework so it is possible to test the full range of Hispanic hypotheses?

The framework itself is based on a simple premise: that all Hispanic hypotheses, both archetypical (panethnic formation, white assimilation, racialized assimilation) and blended (hypotheses combining the archetypical hypotheses), can be represented in a theoretical space based on the salience of race and Hispanic ethnicity. This makes it easier to interpret difficult results as all hypotheses can be fit into the same two-dimensional space. What we think of as discrete, archetypical hypotheses are really just the extreme realizations of two underlying dimensions (salience of race and Hispanic ethnicity) or the corners of the theoretical space. A strong Hispanic dimension and a weak racial dimension characterize panethnic formation as Hispanic represents a single, cohesive category, with no internal, racial divisions. On the other extreme, a salient racial dimension and a weak Hispanic dimension characterize racialized assimilation. Blended hypotheses simply represent the points in between the theoretical extremes. Different relative strengths of race/ethnicity thus imply different substantive interpretations, capturing the range of Hispanic states implied (but not fully specified) by past work.

I parametrize this theoretical framework using latent social space models (Hoff, Raftery, and Handcock 2002). I take data on marriage/cohabitation and use that to map the social distance between racial/ethnic categories (where two categories are close if they have a high rate of intermarriage and far otherwise) (Bottero and Prandy 2003; Laumann 1969). This mapping of the racial/ethnic social space is then compared to the expectations under different Hispanic hypotheses, representing different combinations of racial/Hispanic salience. Thus, each Hispanic hypothesis is represented by a distinct theoretical social space that can be compared against the empirical social space. In this way, I can test a range of complex hypotheses in a tractable manner.

The empirical question is what the racial/ethnic landscape actually looks like: Where does it fall along the two specified dimensions, and what does that say about the racial/ethnic picture in the United States? I begin the paper by discussing past work on Hispanic integration. I then describe the theoretical framework before presenting results based on the American Community Survey, 2010–2012.

Race and Hispanic Ethnicity: Theoretical Approaches

Past work has offered three main hypotheses about the place of Hispanics in the racial stratification system (Frank et al. 2010). First, Hispanic could be a social category in itself, like black or white (Brown, Hitlin, and Elder 2007; Campbell and Rogalin 2006; Golash-Boza 2006). Here, Hispanic is seen as a panethnicity that cross-cuts racial or ethnic identification; or, Hispanic is the racial/ethnic identity ( Barrera 2008; Oropesa et al. 2008). Past work has, for example, demonstrated that Hispanic individuals often have difficulty placing themselves in the U.S. racial stratification system, as the racial categories do not necessarily translate well to their own past experiences (i.e., based on the racial distinctions in Latin American countries) ( Dowling 2014; Rodriguez and Cordero-Guzman 1992). For example, upward of 40 percent of people who identify as Hispanic identify as Other racially, suggesting that Hispanic is their understood racial/ethnic identity ( Brown et al. 2007; Hitlin, Brown, and Elder 2007). A panethnicity is likely to emerge when geographical and occupational concentration coincides with shared cultural factors, like language and religion ( Jones-Correa and Leal 1996; Lopez and Espiritu 1990; Okamoto 2003). Such factors can serve to connect an otherwise disparate population (McConnell and Delgado-Romero 2004). Government agencies (as well as social movements) themselves also play a role in creating a Hispanic category by making Hispanic a key question on forms and surveys ( Mora 2014; Okamoto and Mora 2014).
A second tradition views Hispanics in racial stratification terms (Denton and Massey 1989; Frank et al. 2010; Rumbaut 2009; Telles 2010). Here, Hispanics are seen as divided, or stratified, by racial identity (as well as national origin). The focus thus turns to the internal differentiation of the Hispanic population while the coherency of a Hispanic identity is downplayed or questioned (Alba et al. 2014; Lopez and Espiritu 1990). We can think of this hypothesis in terms of racialized assimilation (Bonilla-Silva 2004; Golash-Boza and Darity 2008). Some Hispanic individuals may be incorporated into the white majority, while others may be incorporated into nonwhite, non-Hispanic racial groups. Black-Hispanics (i.e., those identifying as black and Hispanic) may resemble non-Hispanic blacks, white-Hispanics may resemble non-Hispanic whites, and so on (Golash-Boza and Darity 2008). Racial differences are wrapped up intimately with class, language, physical appearance, and contextual differences across the Hispanic population (Logan 2003; Vargas 2015). Individuals with lighter skin, who are highly educated, live in predominately white neighborhoods, and speak English as the primary language are more likely to take on a white racial identity and less likely to experience racial discrimination (Frank et al. 2010; Golash-Boza and Darity 2008; Stokes-Brown 2012). These factors are often mutually reinforcing as individuals with different phenotypic features experience different levels of discrimination in the labor market as well as different levels of residential and educational segregation (Murguia and Telles 1996).

A third tradition points to the eventual incorporation of all Hispanics into the larger white category (Gallagher 2004; Lee and Bean 2004). Here, Hispanics are assimilated into mainstream culture and politics (see Alba and Nee 2003, although they also allow for the possibility of Hispanics changing the U.S. landscape itself). Here, all individuals identifying as Hispanic become white (and so are treated like they are white). Such an account draws on the still strong black/nonblack divide. The claim is that the real divide in the United States is black or not black (Yancey 2003). Thus, everything is relative to the black population, where black individuals are isolated not just from whites but from other racial groups as well (Alba 2009; Gans 1999; Roediger 1991). If individuals identifying as Hispanic are seen as not black, then a new dividing line will form that includes non-Hispanic whites, Asians, and Hispanics but excludes those who are black (Lee and Bean 2007).

**Heterogeneous Evidence for Different Theoretical Frameworks**

Past work has often used marriage/cohabitation data to test these hypotheses. The results, on the whole, are mixed as no single hypothesis adequately fits the data. A number of studies have found a mix between panethnic assimilation and white assimilation. For example, Fu (2007) concludes that Hispanics “fit both assimilation and panethnicity” as some Hispanics intermarry with whites and others with other Hispanics. Qian and Cobas (2004) find similarly mixed results: They find that white-Hispanics are being assimilated into the larger white category (see also Moran 2003), while the nonwhite Hispanic groups show high rates of intermarriage, indicative of a panethnicity. Qian and Lichter (2007) offer stronger evidence for a panethnicity as they find increasing rates of within-group marriage for Hispanics during the 1990s (see also Rosenfeld 2001). Similarly, Lichter et al. (2007) find decreasing rates of intermarriage between non-Hispanic whites and Hispanics in the 1990s (see also Lichter 2013; Lichter, Carmalt, and Qian 2011). Hispanics (as a whole) are, however, still closest to non-Hispanic whites than to any other racial group (Qian and Lichter 2007).

Other work finds stronger evidence for racialized assimilation, although here too the results point to a mix of hypotheses (e.g., Qian 2002). Quillian and Campbell (2003) find that racial identification strongly determines friendship choices among Hispanic adolescents, while Hispanic acts as a secondary but still important factor. Given the importance of both factors, the in-group bias for white-Hispanics, black-Hispanics, and Other-Hispanics is quite high, almost as high as non-Hispanic whites and blacks. Kao and Vaquera (2006) also find a mix between racialized assimilation and panethnic formation while looking at friendships ties, although the results point more strongly toward panethnicity (see also Kao and Joyner 2006). More generally, many studies in favor of a panethnicity also point to the importance of racial and/or national origin (Mexican, Cuban, etc.) subgroups within the larger Hispanic grouping (Rosenfeld 2001).

**Using Social Space to Specify a Complex Set of Hypotheses**

In sum, past results have been expressed as mixtures of discrete, archetypical hypotheses; additionally, different studies suggest combining different hypotheses. How can we make sense of such a complex, heterogeneous state of affairs? The answer comes in integrating past hypotheses into a single theoretical model, one that allows for mixed hypotheses to be specified in a precise manner. This makes it possible to see how different integration processes (white assimilation, panethnicity, racialized assimilation) combine to yield a particular pattern of intermarriage and more generally, a particular set of social boundaries. The basic idea is to develop a continuous representation of the Hispanic hypotheses. I argue that all Hispanic hypotheses can be characterized by the strength of two underlying dimensions: the salience of racial identity and the salience of Hispanic identity. Each hypothesis, both archetypical and blended, can be logically placed in this two-dimensional theoretical space.

The first question is how to tractably represent all of the hypotheses in the theoretical space. This is a difficult problem because the framework is inherently continuous, requiring a large number of interconnected hypotheses to be
specifying; namely, one must specify every possible hypothesis falling "in between" the archetypical hypotheses as well as the archetypal hypotheses themselves. One must also capture how the hypotheses shift as racial (or Hispanic) salience increases or decreases. I argue that a latent social space approach offers an ideal option (Blau 1977; Hoff et al. 2002; McPherson 1983).

A latent social space approach offers a relational mapping of a social system, here focusing on the racial/ethnic landscape in the United States. Categories are placed into locations in a multidimensional space defined by the observed frequency of contact between them, controlling for the size of different categories. See methods section for details on this multidimensional scaling-like analysis (Bottero and Prandy 2003; Laumann 1969). In this case, I use marriage/cohabitation data to measure the empirical social space. Categories are close together in the multidimensional space if marriage/cohabitation is likely and far apart if marriage/cohabitation is unlikely. Two categories occupying the same location, or with low social distance, thus have the same rate of marriage/cohabitation to all other categories and high frequency of intermarriage/cohabitation with each other. Thus, if white and black are socially close, then white individuals and black individuals face the same social boundaries and can be considered as a single, coherent social group (at least in terms of whom one marries). Note that I interpret marriage rates as the end result of a number of factors, such as geographic segregation, occupational sorting, and socioeconomic status (SES) differences (as well as individual preferences), that create and maintain social boundaries in a population (Smith, McPherson, and Smith-Lovin 2014).

A social space approach is ultimately useful because it makes it easier to specify a set of interconnected, complex hypotheses in a tractable manner. Each hypothesis can be represented by a distinct social space and compared to the empirical social space to find the best fit. Empirically, a latent social space model captures the social distances between categories (here, based on the frequency of intermarriage). My hypotheses can be specified in an analogous way, showing the expected distance between categories under a given hypothesis. These distances can be easily altered to reflect different hypotheses, representing different combinations of racial and Hispanic salience. In that sense, I represent each hypothesis as a picture, or map, and test which map best approximates the actual data (for a related approach, see Levine, Klein, and Mathews 2001). It would be difficult to specify and test hundreds of different hypotheses using more traditional regression frameworks (i.e., most log-linear models or case control logistic regression; Smith et al. 2014).

I begin by describing each archetypical hypothesis in social space terms. Each hypothesis is placed on the two-dimensional theoretical space, representing the salience of race and Hispanic ethnicity. I then turn to the "in-between" hypotheses. Figure 1 plots this two-dimensional representation of the Hispanic hypotheses. The x-axis represents the salience of a Hispanic ethnicity, while the y-axis represents the salience of race. I assume for Figure 1 that there are data on both racial identity and Hispanic identity for respondents and their partners. Other-Hispanics, for example, are those identifying as Other racially and Hispanic ethnically. I use a set of combined racial-Hispanic categories (Other-Hispanic, white-Hispanic, etc.) as this makes it possible to test if racial and Hispanic identities are socially salient. For example, if Hispanic is really just a government-imposed label with no social reality (Rodriguez 2000), then I would not expect a Hispanic identity to be very important in terms of marriage/cohabitation patterns. The key is specifying what social space will look like under different Hispanic hypotheses. I can then ask which racial/Hispanic combination best fits the empirical social space.

Figure 1 is based on a simple aggregate measure of Hispanic ethnicity (yes/no), but it is straightforward to consider a more disaggregated set of categories based on national origin (e.g., Mexican or Cuban). The categories would then have the form of black-Mexican, white-Cuban, and so on. My actual analysis includes both sets of results, one with the simple coding and one including national origin. 

**Salient Hispanic Dimension and Weak Racial Dimension: Panethnic Formation**

First, it may be the case that Hispanic acts a salient demographic dimension while race does not. In this case, there is an overarching panethnicity, where Hispanic represents a unified.

---

1 Racial/ethnic categories are defined as the set of possible identities one can claim within a survey.

2 Marriage/cohabitation is an appropriate metric because it captures the observed divisions in a population (e.g., Rosenfeld 2008; Schwartz 2013). Marriage/cohabitation rates reflect both the physical boundaries separating groups (or the opportunity to interact) as well as status differences between groups (i.e., do relationships form, given the opportunity?) (Haller 1981; Qian and Lichter 2007).

3 More generally, a social space approach is useful because it offers a holistic view of the racial/ethnic social structure. The meaning of Hispanic is not simply determined by within-group relations but also by relations to other racial/ethnic groups as well as the relations between those groups (see also Abascal 2015). For example, how white-Hispanics relate to black-Hispanics is dependent on how white-Hispanics relate to non-Hispanic whites and how non-Hispanic whites relate to black-Hispanics. A social space approach naturally captures these interdependencies (McPherson 2004): The approach captures the position of all categories relative to all other categories in a system of interrelations. A social space approach thus makes it explicit that the categories only find meaning and can thus be understood in relation to other social categories; an idea that is empirically and theoretically lost if we simply focused on the rates of marriage between each pair of categories in isolation.
social group. We would see high rates of intermarriage/cohabitation between those identifying as Hispanic regardless of racial identity or national origin. All the Hispanic categories would also have the same pattern of marriage/cohabitation to other racial/ethnic categories; for example, the rate of marriage with non-Hispanic whites would be the same for black-Hispanics as white-Hispanics. This hypothesis is represented in social space terms in the bottom right corner of Figure 1. All of the Hispanic categories occupy the same social location and are thus relationally identical as Hispanic is the only identification that really matters. This aggregate Hispanic category also occupies a location that is far from non-Hispanic whites. Hispanic is a salient divide, and we would not see assimilation into the white majority. The picture is effectively the same if we consider national origins, like Mexican or Cuban, as all Hispanic categories (black-Mexican, white-Cuban, etc.) occupy a single location.

**Salient Race Dimension and Weak Hispanic Dimension: Racialized Assimilation**

Alternatively, race may be the salient dimension structuring social space while Hispanic identification plays no role at all. Here, the racial identification of an individual is crucial, and we see racialized assimilation rather than a panethnicity. A racialized hypothesis emphasizes racial differences in class and education that make cross-race marriages unlikely. This is reflected in high rates of marriage/cohabitation between Hispanics and non-Hispanics of the same racial identity—as Hispanic identification does not matter. Hispanic and non-Hispanic categories of the same race will occupy the same social location in social space. Looking at the top left of Figure 1, white-Hispanic and white occupy the same social location, black-Hispanic and black occupy the same social location, and so on. This means that white-Hispanics (for example) will have high rates of intermarriage with non-Hispanic whites; they will also have the same rate of marriage/cohabitation to black-Hispanics, Other-Hispanics, and so on.

---

4Marriage/cohabitation rates can be also influenced by the process of convergence, where one spouse “takes on” the racial or ethnic identity of the other. This is more likely to occur along racial lines for Hispanics, who may not be as familiar with the racial distinctions being made in the United States (Dowling 2014). Thus, spouses may converge on racial identity over time, mapping race more strongly onto the observed social groups.
as non-Hispanic whites. The basic idea would be the same if we considered national origins, like Cuban or Mexican. Here, black-Mexican, black-Cuban, and so on will occupy the same location as non-Hispanic black.

**Neither Race nor Hispanic as Salient**

**Demographic Dimensions: White Assimilation**

It is also possible that neither race nor Hispanic identification is salient for those identifying as Hispanic. Here, the racial-Hispanic identity claimed on the survey is not salient in shaping marriage patterns. The racial-Hispanic categories will then be incorporated into a larger racial category. If Hispanics follow the path of some prior immigrant groups, or straight-line assimilation, then they would be incorporated into the aggregate white category. Under white assimilation, those identifying as Hispanic (regardless of racial identity or national origin) will have high rates of marriage/cohabitation with non-Hispanic whites; they will also have the same rate of marriage/cohabitation to other racial/ethnic groups as non-Hispanic whites. Thus, the rate of marriage with non-Hispanic blacks is the same for non-Hispanic whites as black-Hispanics. In social space terms, all of the racial-Hispanic categories will occupy the same location as white. See the bottom left of Figure 1. Here, there is nothing relationally distinguishng black-Hispanics, Other-Hispanics, and so on from non-Hispanic whites, and they are effectively the same even though they identify racially and ethnically different. The story is the same when we consider more finely measured categories based on national origin: Here, black-Mexicans, black-Cubans, Other-Mexicans, Other-Cubans, and so on are all the same as non-Hispanic whites.

**Race and Hispanic as Equally Salient Dimensions: Racial-Hispanic Differentiation**

Finally, race and Hispanic ethnicity may be equally salient. This is consistent with past work pointing to the possibility of Hispanics remaking the racial landscape (Alba and Nee 2003). Here, a number of new, distinct racial/ethnic groups emerge, and we do not see an overarching panethnicity or assimilation into existing racial groups. Under racial-Hispanic differentiation, black-Hispanics have distinct marriage patterns from non-Hispanic blacks, non-black Hispanics (e.g., white-Hispanics, Other-Hispanics), and non-Hispanic whites, as both race and Hispanic ethnicity map onto observed social boundaries. The top right corner of Figure 1 plots a social space where race and Hispanic are equally important demographic dimensions. Black-Hispanics, for example, occupies a location that is equally close to the location occupied by black (as race matters) and the other Hispanic categories (as Hispanic matters) but distinct from both. Thus, black-Hispanics will have about the same chance of being married to (or cohabitating with) non-Hispanic blacks as non-black Hispanics. This differs from racialized assimilation, where black-Hispanic is black; panethnic assimilation, where black-Hispanic is Hispanic; and white-assimilation, where black-Hispanic is white. Each racial-Hispanic category will occupy an analogous, distinct social location.

The picture of social space is similar if we consider national origin. Black-Mexicans, for example, would be distinct from non-Hispanic blacks and non-black Hispanics. The key difference is that with the disaggregated categories, there is no assumption that categories like black-Hispanic or white-Hispanic are themselves socially coherent. We can then ask if this is the case. Do we see white-Hispanic, for example, emerge as a meaningful category, incorporating white-Cuban, white-Mexican, and so on into a distinct, cohesive social group? Or do white-Mexicans have distinct marriage patterns from both white-Cubans (for example) and non-Hispanic whites, thus occupying a distinct location in social space? This would suggest that white-Hispanic does not in fact constitute a clear social group.

**Blended Hypotheses: Combinations of the Archetypical Hypotheses**

The hypotheses laid out so far describe the extreme poles of the theoretical space, defined by the salience of race and Hispanic ethnicity. It is possible, however, that none of these hypotheses fit very well; instead, Hispanic marriage patterns may be represented by some combination, or blend, of the traditional hypotheses. Note that a blended hypothesis does not necessitate having a different set of categories (i.e., a set of blended racial/ethnic categories); rather, a blended hypothesis simply means that the marriage/cohabitation patterns do not fit neatly into one of the four archetypical hypotheses (racialized assimilation, panethnic formation, etc.). These blended hypotheses can be naturally expressed in the given framework. Blended hypotheses represent points in the two-dimensional space that fall in between the archetypical hypotheses. We can thus derive what the in-between states look like within a single theoretical system.

For example, Hispanic marriage patterns may fall in between racialized assimilation and white assimilation. Here, the Hispanic dimension is quite weak and the racial dimension strong but not so strong as to lead to racialized assimilation. The social space for this in-between state will have low (but greater than 0) distance between the racial-Hispanic categories and their respective racial categories. Thus, Filipino-Hispanic will be close to Filipino but will not occupy the same location. The distance between the racial-Hispanic categories and white will, in contrast, be lower than under racialized assimilation.

---

6 Note that this hypothesis is distinct from racialized assimilation for those who identity as white racially: Under white assimilation, those who are white-Hispanic are the same as black-Hispanic, Other-Hispanic, and so on. This is not the case under racialized assimilation.
This social space falls between the top left and bottom left corners in Figure 1. Alternatively, the social space could be in between panethnic formation and racial-Hispanic differentiation (i.e., between the bottom right and top right corners of Figure 1). Race and Hispanic ethnicity are both important demographic dimensions, but Hispanic serves as the more salient of the two. Here, the racial-Hispanic categories are closer to each other than to their respective racial categories but do not form a single cohesive category.

Thus, we can represent the full range of racial-Hispanic states within a single, unified framework. I have discussed a very small number of possibilities here, but the framework is flexible enough to capture any combination of racial/Hispanic salience. The question is what the racial/ethnic social space actually looks like and where it falls in the theoretical space.

**Data**

The data come from three 1 percent samples of the American Community Survey (ACS; Public Use Microdata; Ruggles et al. 2010). The data cover the 2010–2012 samples. Each data set includes information on race and Hispanic identification for all household members as well as the relationship between household members. Here, I am only interested in relationships defined by marriage or cohabitation. Cohabitation is defined in the ACS as unmarried partners. All existing relationships (either marriage or cohabitation) are included in the analysis. The analysis is thus based on the aggregate pattern of current social connections. The analysis makes no restrictions on the age of those in the relationship, save for being over 18.

Since 2000, individuals have been allowed to identify as more than one race on census surveys. Individuals are allowed to identify as white, black, Japanese, Chinese, Hawaiian, Filipino, Native American, Other, as well other smaller Asian categories, and they are allowed to mark two races (a limited number of people mark three). For the main analysis, I code the racial data so that individuals are placed into a single race category. Multirace individuals are placed into a single race based on what they are most likely to have selected if offered only one racial option. I use the coding scheme of Ingram et al. (2003) as a guide. For example, individuals identifying as black and another race are generally coded as black, while those identifying as Native American and white are generally coded as white. I have also repeated the analysis, treating the multirace categories as categories in their own right, and the results are very similar to those reported here.

In addition to the racial questions, individuals were asked if they identified as Hispanic, denoted by Mexican, Cuban, Puerto Rican, or other Hispanic identification. There are thus separate questions about racial and Hispanic identity. I construct a set of racial/ethnic categories from these two questions. The full set of categories includes the cross between racial identity and Hispanic identification.

I run two separate analyses. The first analysis uses a broad Hispanic measure, ignoring national origin. There are five racial-Hispanic categories in this analysis: white-Hispanic (collapsing white-Cuban, white-Mexican, . . . ), black-Hispanic, Other-Hispanic, Native American Hispanic, and Filipino-Hispanic. Individuals who identify as white and Hispanic are labeled as white-Hispanic while those that identify as white but not Hispanic are labeled as white. Some racial categories had too few individuals to construct a separate Hispanic hyphenated category. For example, Filipino is the only Asian ethnicity with sufficient Hispanic population to sustain a separate Hispanic category. The second analysis incorporates national origin into the Hispanic categories. I consider the following national origins: Cuban, Dominican, Puerto Rican, Mexican, and Spaniard, as these represent the largest categories in the data (I also include Other Hispanic as a residual category). The Hispanic categories are thus: white-Cuban, white-Dominican, white-Mexican, white-Puerto Rican, white-Spaniard, white-Other Hispanic, black-Cuban, black-Mexican, and so on.

The second analysis, which incorporates national origin, has the advantage of not forcing individuals into an aggregate category (e.g., defining white-Mexican, white-Cuban, etc. as white-Hispanic) but introduces the practical difficulty of incorporating a large number of categories into the analysis—complicating the presentation of results. I thus use the first analysis to paint a basic picture of social space and walk through the general findings. I use the second analysis to address more specific questions, for example, whether categories like white-Hispanic or black-Hispanic constitute meaningful social groupings, a question that cannot be answered with the first analysis.

It is important to recognize that an individual’s racial/ethnic identity may shift over time as their economic and social conditions change (Penner and Saperstein 2008). Such individual fluidity, while interesting in its own right, is not a concern for the analysis. The goal of this paper is to use marriage/cohabitation rates to describe the observed social boundaries in a population. I am concerned with the pattern of social connections (measured as marriage/cohabitation rates) at a given moment in time. The currently held identity is thus the identity of interest, sufficient to map out the existing social boundaries.

**Methods**

I draw on latent social space models to test the Hispanic hypotheses. The basic idea is to compare the empirical social space, based on marriage/cohabitation data, to the hypothesized social spaces. The question is which spot in the theoretical space (based on racial/ethnic salience) best

---

6Note that some categories, such as black-Spaniard, had too few individuals to form a separate category.
fits the empirical data. The first step in answering the question is to estimate the empirical racial/ethnic social space, showing the social distances between racial/ethnic categories. Categories are far apart in the space if the frequency of marriage/cohabitation is low and close otherwise. I use the Hoff et al. (2002) latentnet model to estimate the empirical social space (see also Krivitsky et al. 2009). I can write the model as:

$$\log(F_{ij}) = \mu + \mu_{i}^{R} + \mu_{j}^{C} - \|Z_{i} - Z_{j}\|,$$

where \(F_{ij}\) is the frequency of marriage/cohabitation between category \(i\) and \(j\); \(\mu\) captures the overall mean, \(\mu_{i}^{R}\) and \(\mu_{j}^{C}\) are (fixed effects) factors for the rows and columns, and \(Z_{i}\) is the unobserved latent position of category \(i\) in a multidimensional Euclidean space (Krivitsky et al. 2009). The model estimates the location of each category such that the distance between \(i\) and \(j\), over all \(i\) and \(j\), best predicts the frequency of marriage/cohabitation between categories, or \(F_{ij}\). As the distance between \(i\) and \(j\) increases, the frequency of marriage/cohabitation, relative to the marginals, decreases. Note that \(\mu_{i}^{R}\) and \(\mu_{j}^{C}\) control for the size of different categories. The estimates are thus net of chance expectations (i.e., by chance we would not expect many ties between two small categories). Note also that \(F_{ij}\) captures the frequency of marriage/cohabitation between racial/ethnic categories measured in the ACS. For the first analysis, this includes: black, black-Hispanic, Filipino, Filipino-Hispanic, Native-American, Native American-Hispanic, Other, Other-Hispanic, white, white-Hispanic, as well as a number of Asian ethnic identities.

It is necessary to specify the dimensions of social space before estimating the final social locations. I use the data itself to determine how many dimensions are necessary to define the racial/ethnic space. I first estimate the latent space model using a two-dimensional space. I repeat the estimation with an increasing number of dimensions and select the best fitting model. I determine model fit by predicting the frequency of contact between all \(i, j\) pairs using the estimated locations for that model. For example, given a value of 10 for \(\mu\) (controlling the overall frequency of marriages), 0 for \(\mu_{i}^{R}\), 0 for \(\mu_{j}^{C}\) (adjusting for the relative sizes of the two categories), and a distance of 2 (i.e., \(\|Z_{i} - Z_{j}\| = 2\)), the predicted \(F_{ij}\) would be \(\exp(10 + 0 + 0 - 2) = 2980\). I then compare the predicted values to the true values, summarizing this as a deviance score.\(^{7}\) Larger values indicate a worse fit. I use a Bayesian Information Criterion (BIC) score to penalize more complicated models, here models with a larger number of dimensions.

Table 1 presents the latentnet model results for the 2010–2012 data. Each row in the table presents the fit statistics for a model with different dimensions. The models grow increasingly complex as one moves down the rows, going from 2 dimensions all the way up to 14 dimensions. The results clearly show that the best fitting model, relative to the number of parameters, is the model with 7 dimensions. This is clear as the BIC score is minimized with 7 dimensions. I thus use the social locations based on a 7-dimensional solution. I use this model to define the empirical distance matrix, \(D_{\text{empirical}}\). \(D_{\text{empirical}}\) captures the observed distances between all \(ij\) categorical pairs (i.e., I take the inferred locations and calculate the distance between each pair of categories, \(\|Z_{i} - Z_{j}\|\)). The matrix will be \(N \times N\), where \(N\) is the number of categories in the analysis.

\(^{7}\)Deviance is defined as: \(2\sum F_{ij} * \log\left(\frac{F_{ij}}{\hat{F}_{ij}}\right)\), where \(F_{ij}\) is observed frequency count of marriages between categories \(ij\) and \(\hat{F}_{ij}\) is the predicted count based on the model.
Testing Hypotheses

I use the empirical social space to test the Hispanic hypotheses. Each hypothesis is specified in social space terms and is then compared to the empirical social space to find the best fit. The difference between the empirical distance matrix and the expected distance matrix (under each hypothesis) is summarized using a simple total sum of squares measure: $\sum (D_{\text{empirical}} - D_{\text{hypothesis}})^2$. The degrees of freedom under each hypothesis are the same, making the summary measures directly comparable. The hypothesis with the lowest total sum of squares (TSS) offers the best fit. The difference between the empirical distance matrix and the expected distance matrix (under each hypothesis) is summarized using a simple total sum of squares measure:

$$\sum (D_{\text{empirical}} - D_{\text{hypothesis}})^2.$$  

The degrees of freedom under each hypothesis are the same, making the summary measures directly comparable. The hypothesis with the lowest total sum of squares (TSS) offers the best fit.

The key is to specify what social space will look like under each hypothesis. Specifically, one must define the distance matrix that would arise if each hypothesis were true. Thus, there will be one distance matrix (or hypothesis) corresponding to each coordinate in the theoretical space. I begin with the archetypical hypotheses before moving to the rest of the theoretical space. In each case, all of the categories are held fixed at their observed social location except for the racial-Hispanic categories (Other-Hispanic, white-Hispanic, etc.). The racial-Hispanic categories are moved to be consistent with the specified hypothesis given the (fixed) locations of the other racial categories. I then recalculate the distance matrix with the new locations of the racial-Hispanic categories. The question is which movement of the racial-Hispanic categories causes the least damage to the observed social space.

Under a hypothesis of white assimilation, all of the racial-Hispanic categories will occupy the same location as non-Hispanic white. I thus take the observed location of white (defined along seven dimensions) and impose that location on all of the racial-Hispanic categories. I then define a distance matrix, $D_{\text{white assim}}$, that is calculated from these new locations. This distance matrix will have 0 distance between the racial-Hispanic categories and white. Similarly, the racial-Hispanic categories will be relationally white relative to other categories: If black (for example) is far from white, it will also be far from the racial-Hispanic categories as they occupy the same location as white.

Under racialized assimilation, the racial-Hispanic categories will occupy the location of their analogous racial category. Here, each racial-Hispanic category is moved to their respective racial category location. For example, Other-Hispanic is moved to the observed location of Other. I then calculate a distance matrix, $D_{\text{race assim}}$, based on this new set of locations. The distance between the Hispanic categories and the respective racial category will be 0; they (e.g., Other and Other-Hispanic) will also have the same distance to all other racial categories.

With a panethnic formation hypothesis, all racial-Hispanic categories will occupy the same location. Here, I first calculate the median location of the racial-Hispanic categories. For each dimension, the locations of the racial-Hispanic categories are gathered, and the median over those values is calculated. Each racial-Hispanic category is given this new set of locations. The racial-Hispanic categories are thus assumed to occupy the same location, one that falls in the middle of their observed locations. The distance matrix based on these locations is defined as $D_{\text{panethnic}}$. The distance between the racial-Hispanic categories, for example, will be 0. Additionally, under a panethnic formation hypothesis, Hispanic is a category in its own right and is not incorporated into the aggregate white category. Thus, I set the distance between white and the racial-Hispanic categories to be greater than 0, specifically, at the mean distance between white and all other categories (as Hispanic is as much a racial category as black, Native American, etc.).

I can use $D_{\text{race assim}}$, $D_{\text{white assim}}$, and $D_{\text{panethnic}}$ to specify the distance matrices for the rest of the theoretical space, including racial-Hispanic differentiation (top right corner in Figure 1). The remaining hypotheses represent blends or mixes of the archetypical hypotheses. I formally capture this notion by representing the blended (or in-between) hypotheses as weighted summations of the distance matrices for white assimilation, racialized assimilation, and panethnic formation. Each location in the theoretical space can thus be represented by placing an appropriate set of weights onto the archetypical distance matrices. A high salience on one dimension will correspond to a high weight for the analogous distance matrix; for example, a strong race dimension means a high weight on $D_{\text{race assim}}$. In the extreme case, the weights would be $1 \times D_{\text{race assim}} + 0 \times D_{\text{white assim}} + 0 \times D_{\text{panethnic}}$, corresponding to racialized assimilation, or $0 \times D_{\text{race assim}} + 0 \times D_{\text{white assim}} + 1 \times D_{\text{panethnic}}$, corresponding to panethnic formation. Low salience on both dimensions will correspond to a high weight on the distance matrix for white assimilation (as neither race nor Hispanic ethnicity is important under white assimilation; i.e., $0 \times D_{\text{race assim}} + 1 \times D_{\text{white assim}} + 0 \times D_{\text{panethnic}}$ for white assimilation).

The blended hypotheses extend this logic in a straightforward manner. For example, consider a hypothesis that corresponds to a moderate racial dimension and a weak Hispanic dimension. Assume this hypothesis falls halfway

---

8This is the case because the number of categories that move location are the same under each hypothesis (the number of racial-Hispanic categories).
between racialized assimilation and white assimilation in Figure 1. I can write the distance matrix as follows: 
\[
.5 \times D_{raceassim} + .5 \times D_{whiteassim} + 0 \times D_{panethnic}. 
\]
There is 0 weight on the panethnic distance matrix because the Hispanic dimension is weak. There is equal weight on the racialized assimilation and white assimilation matrices because the salience of race is only moderate, falling between racialized assimilation and white assimilation. Similarly, a hypothesis based on a weak racial dimension and a moderate Hispanic dimension (assuming it is halfway between panethnic formation and white assimilation) would yield the following formula: 
\[
0 \times D_{raceassim} + .5 \times D_{whiteassim} + .5 \times D_{panethnic}. 
\]

Consider a strong Hispanic, moderate race hypothesis. Assume that this hypothesis falls directly in between a hypothesis of panethnic formation and racial-Hispanic differentiation. This yields the following set of weights:
\[
.25 \times D_{raceassim} + .5 \times D_{whiteassim} + .75 \times D_{panethnic}. 
\]

These calculations are simply examples of how to represent different points in the theoretical space as a set of weights. The question is what the actual social space looks like. To that end, I will produce a fit over the entire theoretical space, capturing all possible hypotheses, both archetypical and blended.

Note that the analysis does not include controls for other variables, such as income or education. Such factors are important in structuring marriage patterns. My core question is, however, focused on the overall rates of intermarriage between racial/ethnic categories, making it inappropriate to include a heavy list of controls in the analysis. More generally, intermarriage rates are the end result of a number of structural factors (e.g., residential sorting and income inequality) that create divisions in the population. I focus on the end result (the intermarriage rates) of all such factors rather than the effect of any single variable.

**Results**

Figure 2 presents a formal test of the Hispanic hypotheses. The results are presented as a contour plot in three dimensions. The x-axis corresponds to racial salience, and the y-axis corresponds to Hispanic salience. The z-axis, providing the surface’s height, corresponds to the fit statistics, comparing the true distance matrix (i.e., the observed distance between all \(ij\) categories) to the distance matrix under that hypothesis. Each point in the surface represents a different hypothesis. The figure thus captures hypotheses resulting from all potential combinations of racial/Hispanic salience. Lower numbers indicate a better fit as the true distances are close to the distances implied by that hypothesis.

It is clear from Figure 2 that the best fitting hypotheses is a blend of racialized assimilation and racial-Hispanic differentiation, with the results closer to differentiation than any other archetypical hypothesis. This corresponds to a location in the theoretical space at the top of the x-axis (the racial dimension) and nearly to the top of the y-axis (the Hispanic dimension). There is high salience on both the race and Hispanic dimensions, with the racial weight slightly higher; or, more formally, the best fitting hypothesis corresponds to a .620 weight on racialized assimilation, a .380 weight on panethnic formation, and 0 weight on white assimilation.

What does the social space actually look like, given it is best described as racial-Hispanic differentiation, with a hint of racialized assimilation? This question is answered in Figure 3, which presents a three-dimensional representation.
of the empirical racial/ethnic social space.\textsuperscript{15} Categories are close together if they have a high frequency of intermarriage/cohabitation (and similar rates of marriage/cohabitation to other racial categories) and far apart if they have low rates of intermarriage/cohabitation (and dissimilar patterns of marriage/cohabitation).

One can see, most clearly, from Figure 3 that all of the racial-Hispanic categories occupy distinct locations in the social space. This means that each racial-Hispanic category has a unique profile of marriage/cohabitation different from any other racial/ethnic category. The racial-Hispanic categories are thus not only distinct from white, black, and so on, they also exhibit distinct patterns from each other. For example, black-Hispanics have different marriage/cohabitation patterns than non-Hispanic blacks and non-black Hispanics, where black-Hispanics are socially closer to white-Hispanics and non-Hispanic blacks than they are to each other.

The racial-Hispanic categories thus cannot be treated as a single, unified category (as the racial-Hispanic categories are not collapsed into a single location), but nor can they be easily incorporated into white or other racial, non-Hispanic categories (as they occupy different locations than their respective racial categories). If panethnic formation or white assimilation were true, then all Hispanic individuals, regardless of racial identity, would find the same barriers difficult to cross. This is clearly not the case as the Hispanic categories occupy very different social locations. There are thus particularly poor fits in Figure 2 for white assimilation and panethnic formation, as well as all hypotheses that are a blend of these two.

It is also clear from Figure 3 that the closest category to the racial-Hispanic categories is usually the respective racial category, thus, the mixture between racialized assimilation and racial-Hispanic differentiation. White-Hispanic is closest to white. Black-Hispanic is closest to black, and so on. The exception is Other-Hispanic. This is not surprising, however,

\textsuperscript{15}The actual results are based on a seven-dimensional solution, but unfortunately, seven dimensions are difficult to visualize. I thus present a three-dimensional representation of the results based on principal component analysis.
as those identifying as Other non-Hispanic represent a diverse set of racial backgrounds. Other need not mean the same thing for those identifying as Hispanic as those not identifying as Hispanic. In general, however, the racial-Hispanic categories are closer to their respective non-Hispanic racial category than other Hispanic categories. Thus, white-Hispanics have the highest frequency of out-group marriage/cohabitation to non-Hispanic whites. This means that white-Hispanic is closer to white than black-Hispanic, Other-Hispanic, or Native American-Hispanic. The distance to Other-Hispanic is, for example, about 45 percent more than the distance to white.

It is important to emphasize that the racial-Hispanic categories are not easily incorporated into their respective racial categories, despite the relative closeness between black-Hispanic and black, white-Hispanic and white, and so on. Individuals identifying as black-Hispanic (for example) are more akin to those identifying as black than Filipino-Hispanic, white-Hispanic, and so on but are still very much distinct from non-Hispanic blacks. This is the case as Hispanic ethnicity does exert influence over marriage/cohabitation patterns. Native American-Hispanic is, for example, 43 percent closer to other Hispanic categories (using the median Hispanic location\(^{16}\)) than all categories combined (i.e., averaging over all distances).\(^{17}\) And more generally, Native American-Hispanic is 43 percent closer to other Hispanic categories (using the median Hispanic location) than all categories combined (i.e., averaging over all distances).\(^{17}\)

\(^{16}\)The median Hispanic location is simply the location falling in the middle of the Hispanic categories along each dimension.

\(^{17}\)Note that I do exclude Native American-Hispanic in the calculation of the median location.
Smith

American-Hispanics are closer than non-Hispanic Native Americans to white-Hispanic, Other-Hispanic, and so on.

In sum, the racial-Hispanic categories remain distinct categories in their own right, and the space most closely approximates racial-Hispanic differentiation (although the closest category tends to be the respective racial category). Those identifying as white and Hispanic, for example, have a relatively high rate of marriage/cohabitation (and similar patterns) with non-Hispanic whites. White-Hispanics are, however, more likely to marry/cohabit with other white-Hispanics than with non-Hispanic whites (i.e., there is an in-group bias); they are also more likely than non-Hispanic whites to marry/cohabit with other Hispanic groups.

**Testing Hispanic Hypotheses using National Origins**

The results thus far have presented an aggregate picture of the racial/ethnic social space. The results, have, however, also neglected potentially important differences within the Hispanic categories (Monk 2015; Okamoto and Mora 2014). Specifically, the analysis ignored any differences based on national origin when describing the marriage/cohabitation patterns. This means that white-Hispanic, black-Hispanic, etc. were assumed to be meaningful categories in their own right, ignoring any distance between Cubans, Mexicans, etc. of the same racial identity. In this analysis, I treat each racial, national origin combination (black-Cuban, white-Mexican) as distinct. It thus becomes a question whether Hispanic categories of the same racial identity but different national origin are socially close enough to be considered a coherent social group.

I begin with the main question posed throughout the paper: What blend of hypotheses best captures the experience of Hispanics in the United States? I present the results in Figure A1 (in the appendix). The figure is organized as in Figure 2, with racial salience on the x-axis, Hispanic salience on the y-axis, and the fit under each hypothesis on the z-axis. The results are similar to what was seen in the first analysis except the Hispanic dimension comes out more strongly here. The best fitting hypothesis is with high (equal) salience on both the racial and Hispanic dimensions, with .50 weight on racialized assimilation, .50 weight on panethnic formation, and 0 weight on white assimilation. This corresponds to strong salience on both dimensions, or the top right corner in the theoretical space.

The results suggest, as before, that the Hispanic categories are socially distant from non-Hispanic white, other Hispanic categories, and their respective racial category, lending support to a differentiation hypothesis. The Hispanic categories are, in addition, not always closer to their respective racial category than to the other Hispanic categories, as in the first analysis. While some categories (black-Mexican, Native American-Mexican) are in fact socially closer to a non-Hispanic racial category, other Hispanic categories (white-Puerto Rican, white-Cuban) are closer to the median Hispanic location. For example, white-Cuban is 17 percent closer to the median Hispanic location than to non-Hispanic white. In many cases, the distances are not statistically distinguishable (Other-Dominican, white-Mexican).

The results thus point strongly to a differentiation hypothesis: Relationally, black-Mexicans are not white, Hispanic, or black. But how far does this differentiation extend? Are the racial-Hispanic categories, like black-Hispanic, socially coherent, so there is little social distance between black-Cuban, black-Mexican, and so on? Or do black-Mexicans (for example) exhibit distinct marriage patterns even from other Hispanic groups of the same racial identity?

I answer this question in Table 2, which presents the distances between the key categories of interest. Each row represents a different Hispanic category (black-Cuban, white-Cuban, etc.). The columns capture the distance to different categories of interest. The distances in Table 2 are reported as 95 percent credible intervals to capture the uncertainty in the estimates. For example, a distance of 2 corresponds to an 86 percent decrease in the expected number of marriages (compared to two categories at distance 0). The first column shows the distance to the respective racial category; the second column shows the distance to the median Hispanic location (the location falling in the middle of the Hispanic categories), the third column shows the distance to the non-Hispanic white location, and the fourth column shows the median distance to categories of the same racial identity but different national origins (e.g., for black-Cuban, one would take the distances to black-Mexican, black-Dominican, black-Puerto Rican, etc. and calculate the median). I also report the category closest to each Hispanic category.

It is clear from Table 2 that the Hispanic categories are typically closest to another Hispanic category of the same racial identity but a different national origin. Black-Dominican is closest to black-Puerto Rican, Other-Dominican is closest to Other-Puerto Rican, white-Dominican is closest to white-Puerto Rican, and so on. The distance to other racial/ethnic categories is often quite high in comparison. For example, black-Dominican is 59 percent closer to black-Puerto Rican than to non-Hispanic black and 78 percent closer to black-Puerto Rican than to non-Hispanic white.

These results raise the possibility that aggregate categories like white-Hispanic, black-Hispanic, etc. constitute coherent social groups (with low social distance between categories of the same racial identity but different national origins). I examine this possibility more closely in column 4 of Table 2, which captures the median distance between the category of interest and Hispanic categories of the same racial identity.

The results, perhaps surprisingly, are largely inconsistent with a racial-Hispanic group hypothesis: There are often considerable social divides between groups of the same racial identity but different national origins. For example, while black-Cuban is quite close to black-Puerto-Rican, it is
Table 2. Social Distance by Racial Identity and National Origin.

<table>
<thead>
<tr>
<th>Distance from Hispanic Categories (Rows) to Different Categories of Interest (Columns)</th>
<th>Non-Hispanic Racial Category</th>
<th>Median Hispanic Location 1</th>
<th>Non-Hispanic White Location 2</th>
<th>Hispanic Categories of Same Race Location 3</th>
<th>Closest Category Location 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Black]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[Black-Puerto Rican]</td>
</tr>
<tr>
<td>Black-Dominican</td>
<td>(4.648, 4.923)</td>
<td>(4.399, 4.986)</td>
<td>(5.632, 5.862)</td>
<td>(3.226, 3.887)</td>
<td>(2.419, 2.856)</td>
</tr>
<tr>
<td>[Black]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[Black-Puerto Rican]</td>
</tr>
<tr>
<td>[Black]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[Black-Other Hispanic]</td>
</tr>
<tr>
<td>Black-Other Hispanic</td>
<td>(3.433, 3.628)</td>
<td>(3.922, 4.498)</td>
<td>(5.024, 5.225)</td>
<td>(2.921, 3.553)</td>
<td>(2.699, 3.196)</td>
</tr>
<tr>
<td>[Black]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[Black-Puerto Rican]</td>
</tr>
<tr>
<td>Black-Puerto Rican</td>
<td>(2.829, 3.000)</td>
<td>(3.952, 4.365)</td>
<td>(4.544, 4.73)</td>
<td>(2.764, 3.365)</td>
<td>(2.419, 2.856)</td>
</tr>
<tr>
<td>[Black]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[Black-Dominican]</td>
</tr>
<tr>
<td>Filipino-Mexican</td>
<td>(3.020, 3.716)</td>
<td>(4.817, 5.694)</td>
<td>(4.101, 4.663)</td>
<td>(2.430, 4.056)</td>
<td>(2.43, 4.056)</td>
</tr>
<tr>
<td>[Filipino]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[Filipino-Other Hispanic]</td>
</tr>
<tr>
<td>Filipino-Other Hispanic</td>
<td>(3.136, 4.006)</td>
<td>(5.662, 6.865)</td>
<td>(4.735, 5.353)</td>
<td>(2.430, 4.056)</td>
<td>(2.43, 4.056)</td>
</tr>
<tr>
<td>[Filipino]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[Filipino-Mexican]</td>
</tr>
<tr>
<td>[Native American]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[Native American]</td>
</tr>
<tr>
<td>Native American-Other Hispanic</td>
<td>(4.482, 5.119)</td>
<td>(3.687, 4.325)</td>
<td>(4.953, 5.228)</td>
<td>(3.054, 4.136)</td>
<td>(3.214, 3.859)</td>
</tr>
<tr>
<td>[Native American]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[Native American-Mexican]</td>
</tr>
<tr>
<td>[Native American]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[Native American-Other Hispanic]</td>
</tr>
<tr>
<td>Other-Cuban</td>
<td>(5.134, 5.801)</td>
<td>(4.086, 4.788)</td>
<td>(5.01, 5.286)</td>
<td>(3.448, 4.575)</td>
<td>(3.00, 3.449)</td>
</tr>
<tr>
<td>[Other]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[Other-Puerto Rican]</td>
</tr>
<tr>
<td>Other-Dominican</td>
<td>(4.721, 5.213)</td>
<td>(4.575, 5.089)</td>
<td>(5.68, 5.859)</td>
<td>(3.938, 4.58)</td>
<td>(2.392, 2.586)</td>
</tr>
<tr>
<td>[Other]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[Other-Puerto Rican]</td>
</tr>
<tr>
<td>Other-Mexican</td>
<td>(4.344, 4.613)</td>
<td>(3.752, 4.207)</td>
<td>(3.995, 4.039)</td>
<td>(3.834, 4.325)</td>
<td>(2.546, 2.62)</td>
</tr>
<tr>
<td>[Other]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[Other-Other Hispanic]</td>
</tr>
<tr>
<td>Other-Other Hispanic</td>
<td>(4.18, 4.485)</td>
<td>(3.213, 3.744)</td>
<td>(4.226, 4.293)</td>
<td>(3.268, 3.487)</td>
<td>(2.546, 2.62)</td>
</tr>
<tr>
<td>[Other]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[Other-Mexican]</td>
</tr>
<tr>
<td>Other-Puerto Rican</td>
<td>(3.359, 3.641)</td>
<td>(3.653, 4.118)</td>
<td>(4.255, 4.346)</td>
<td>(3.00, 3.449)</td>
<td>(2.392, 2.586)</td>
</tr>
<tr>
<td>[Other]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[Other-Dominican]</td>
</tr>
<tr>
<td>Other-Spaniard</td>
<td>(3.860, 4.75)</td>
<td>(4.675, 5.427)</td>
<td>(4.322, 4.616)</td>
<td>(3.834, 4.638)</td>
<td>(3.448, 4.638)</td>
</tr>
<tr>
<td>[Other]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[Other-Cuban]</td>
</tr>
<tr>
<td>White-Cuban</td>
<td>(4.024, 4.095)</td>
<td>(3.146, 3.718)</td>
<td>(4.024, 4.095)</td>
<td>(3.539, 3.891)</td>
<td>(2.727, 2.836)</td>
</tr>
<tr>
<td>[White]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[White-Other Hispanic]</td>
</tr>
<tr>
<td>White-Dominican</td>
<td>(5.116, 5.267)</td>
<td>(3.701, 4.294)</td>
<td>(5.116, 5.267)</td>
<td>(3.700, 4.018)</td>
<td>(2.500, 2.670)</td>
</tr>
<tr>
<td>[White]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[White-Puerto Rican]</td>
</tr>
<tr>
<td>White-Mexican</td>
<td>(3.032, 3.054)</td>
<td>(2.947, 3.495)</td>
<td>(3.032, 3.054)</td>
<td>(3.713, 3.835)</td>
<td>(2.723, 2.781)</td>
</tr>
<tr>
<td>[White]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[White-Other Hispanic]</td>
</tr>
<tr>
<td>White-Other Hispanic</td>
<td>(3.226, 3.264)</td>
<td>(1.938, 2.54)</td>
<td>(3.226, 3.264)</td>
<td>(2.727, 2.836)</td>
<td>(2.653, 2.75)</td>
</tr>
<tr>
<td>[White]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[White-Puerto Rican]</td>
</tr>
<tr>
<td>[White]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[White-Dominican]</td>
</tr>
<tr>
<td>White-Spaniard</td>
<td>(2.665, 2.763)</td>
<td>(2.957, 3.358)</td>
<td>(2.665, 2.763)</td>
<td>(3.428, 3.624)</td>
<td>(2.665, 2.763)</td>
</tr>
<tr>
<td>[White]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[White]</td>
</tr>
</tbody>
</table>

Note. The numbers in parentheses reflect the distance between the row category and the column category. The distances are reported as 95 percent credible intervals. The names in the brackets below the distances show which category (if ambiguous) the distance corresponds to. For example, the distance in the top left cell is the distance between black-Cuban and non-Hispanic black. Note that the median Hispanic location excludes the row category of interest.

relatively far from other black-Hispanic categories, like black-Dominican and black-Mexican. Overall, the distance between black-Cuban and the black-Hispanic categories corresponds to a 98 percent decrease in the frequency of marriages (compared to two categories at distance 0). This distance is no lower than the distance between black-Cuban and non-Hispanic black (see columns 1 and 4 in Table 2). For black-Mexican, the distance to non-Hispanic black is actually
lower than the distance to the black-Hispanic categories. Black-Mexican is not particularly close to black-Cuban, black-Dominican, or black-Puerto Rican. We see similar results with the white-Hispanic categories. White-Cuban is socially closer to white-Hispanic categories than non-Hispanic white, but the opposite holds for white-Mexican, and there is no clear difference for white-Puerto Rican (see column 1 and 4 in Table 2).

The Hispanic categories thus tend to be socially distinct, with low rates of out-group marriage and distinct patterns of marriage to other categories, even Hispanic categories of the same racial identity. This shows the importance of national origins in creating social boundaries (see also Rosenfeld 2001).\(^1\) White-Cubans are not simply white, but nor are they simply Hispanic: We do not see the emergence of a larger Hispanic category, even when looking only at individuals with the same racial identity. White-Hispanic does not clearly emerge as a coherent social group.

Putting this together, the results suggest that social boundaries form first around national origins. Individuals identifying as Mexican have distinct marriage/cohabitation patterns from Cubans, Dominicans, and so on. Some divides are, however, less salient than others. For example, Dominicans and Puerto Ricans are separated by relatively weak social boundaries. Individuals also marry (or cohabit with) people of a similar class/language background, which tends to map onto racial identity. Additionally, individuals may converge on racial identity over time, creating clearer racial boundaries. Given the strong divides along national origins (some stronger than others), the end result is a set of social groups with distinct racial and national identities.

**Conclusion**

What place do Hispanics hold in the racial stratification system? There is no simple answer. Traditional hypotheses of white assimilation, racialized assimilation, and panethnic formation do not fit the evidence very well alone. Past work has consequently described Hispanic outcomes using the language of blends, mixtures, and heterogeneity as no single theory captures the Hispanic experience (Alba et al. 2014; McConnell and Delgado-Romero 2004). Here, I develop an analytical framework to help make sense of this complex picture. I integrate all Hispanic hypotheses into a single theoretical framework. I argue that all Hispanic hypotheses, both archetypical and blended, can be represented in a theoretical space defined by the salience of race and Hispanic ethnicity. I parameterize this theoretical space using a social space framework—where each hypothesis (archetypical and blended) is specified as a distance matrix, showing the social distance between racial/ethnic groups. Each hypothesis is thus represented as a map of sorts to be compared against the actual data. This formalization makes it easier to test a complex set of hypotheses.

Using ACS marriage/cohabitation data (2010–2012) and latent social space models, I find that the best fitting hypothesis puts strong weight on both the racial and Hispanic dimensions. My results suggest that the social space is best represented by a differentiation hypothesis, where the Hispanic categories are not easily incorporated into a racial category or an aggregate Hispanic category. Thus, categories like black-Puerto Rican have distinct marriage/cohabitation patterns: with high in-group bias, relatively high rates of marriage to non-Hispanic blacks (relative to other Hispanic groups), and relatively high rates of marriage to other Hispanic groups (relative to non-Hispanic blacks).

Moreover, it is clear that racial-Hispanic categories, like white-Hispanic and black-Hispanic, do not represent coherent social groups. The distance between categories of the same racial identity and different national origin are often quite high. For example, while white-Dominican is close to white-Puerto Rican, white-Dominican is not particularly close to white-Mexican. And in fact, white-Mexicans are socially closer to non-Hispanic whites than other white-Hispanic groups. Thus, even among groups with the same racial identity, the tendency toward a larger Hispanic grouping is weak.

Overall, the results suggest that racial/ethnic categories, like white-Cuban, black-Mexican, etc. occupy distinct locations in social space or represent distinct social groups, with boundaries between sets of individuals with different racial identities or national origins. White-Cuban is, for example, distinct from non-Hispanic white, black-Cuban, and white-Mexican. This does not mean that individuals will strongly identify as a hyphenated racial identity, such as white-Cuban; it does, however, mean that an individual who identifies as white and Cuban is likely to be married to someone else who identifies as white and Cuban. Such relationally defined groups, like white-Cubans, thus have the potential to become strong identities for individuals, assuming such social divides continue over long periods of time and are reinforced by existing differences in material and cultural resources.

These results offer an important update to past work on Hispanic heterogeneity. Many studies argue that different subgroups experience different kinds of integration. For example, past work points to a mix of white assimilation and panethnic formation: where some Hispanic subgroups become white while all others come together to form a distinct panethnicity (e.g., Fu 2007). The results here suggest that neither of these processes occur: None of the Hispanic groups occupy the same location in social space, nor is white-Hispanic incorporated into the non-Hispanic white location. More generally, there is little evidence that some subgroups follow one process (white assimilation) while others follow another (panethnicity). All Hispanic groups follow the same basic “logic” of differentiation, occupying

\(^{1}\)It is important to note that the results are not solely driven by national origin as it is not the case that categories of the same national origin but different racial identity cluster together. For example, on average, white-Cuban is 28 percent closer to Hispanic categories of the same race but different national origin (white-Mexican, white-Puerto Rican, etc.) than to Cuban categories of different racial identity (black-Cuban, Other-Cuban).
distinct locations in social space, with distinct marriage patterns from all other groups (including other Hispanic groups).

What do the results suggest for the future of racial stratification in the United States? First, we may continue to see (relationally) distinct categories, like white-Mexican and black-Dominican. This would suggest that Hispanics have greatly reshaped the racial stratification system, as they do not fall easily into traditional categories or simply fall into a “new” category of their own. Second, it is possible that an aggregate Hispanic identity will become stronger. Distinctions between Cubans, Mexicans, etc. fall away in favor of a panethnicity (with weaker boundaries, like Dominican/Puerto Rican, falling first). Racial boundaries may still exist, however, making it more likely to see white-Hispanic, black-Hispanic, and Other-Hispanic rather than an overarching Hispanic category. Finally, the tendency toward racialized assimilation may increase. Here, we would see the eventual incorporation of the Hispanic categories into larger racial aggregates. Even if racial identity did become more important, the incorporation of the Hispanic categories into larger aggregates is unlikely to be complete: for those identifying as Other and Hispanic are unlikely to ever be incorporated into another racial category. Other-Mexican, for example, is much farther from Other than white-Mexican is to white (black-Mexican is to black, etc.). Thus, Hispanic is unlikely to disappear completely as a racial/ethnic category even under conditions of increasing racial salience. Other-Hispanic categories would simply become the Hispanic categories, and the rest would slowly assimilate into larger racial groups.

Methodologically, this paper offers a new approach for specifying and testing sociological theories. I introduce a framework based on a theoretical space. Here, the researcher specifies the competing theories as a N-dimensional grid. In that sense, what we think of as discrete hypotheses are really points in a larger theoretical continuum. The approach will be particularly useful in cases where the core hypotheses can be characterized in a table, but there exists many logical possibilities in between the “edge” hypotheses. Here, we have four basic hypotheses (i.e., white assimilation, racialized assimilation, etc.) capturing two underlying dimensions (racial/Hispanic salience). Similar examples can be found in effectively every subfield in the discipline; recent examples include such diverse topics as network structure among adolescents and labor markets in the pre–civil war south (e.g., McFarland et al. 2014; Ruef 2012; Smith and Faris 2015). The hope is that this approach will make it easier to characterize theoretical tables in continuous terms: where it becomes easier to specify (and test) complicated hypotheses, those representing blends, or mixtures, of the core hypotheses.

Appendix

![Figure A1](image-url)

**Figure A1.** Formally testing the Hispanic hypotheses: Splitting the Hispanic categories by national origin.

Note. The axes are scaled from 0 to 1, corresponding to the relative weight put on racialized assimilation or panethnic formation as one moves away from white assimilation (holding the other dimension at 0 weight). The z-axis corresponds to the model fit for each hypothesis, measured as the total sum of squares (TSS). The TSS compares the true distance matrix to the distance matrix under that hypothesis. Lower values imply a better fit.
Acknowledgments

The author would like to thank Robin Gauthier, Jacob Fisher, and Regina Werum for helpful comments on earlier drafts of this paper. The author would also like to thank the HAAS faculty award program at the University Nebraska-Lincoln for providing financial support during the writing of this manuscript.

References


Author Biography

Jeffrey A. Smith is an assistant professor as the University of Nebraska-Lincoln. His research interests fall at the intersection of network analysis, social stratification, and traditional statistical methods. He has done methodological work on network sampling and missing data, developing new techniques to make inference about network properties form incomplete data. His more substantive work has explored network structure in adolescence, status hierarchies, as well as changes in discussion networks across time, with a focus on shifts in homophily.